Aging-Aware Design Methods for Reliable Analog Integrated Circuits using Operating Point-Dependent Degradation

Dem Fachbereich Physik, Elektrotechnik und Informationstechnik
der Universität Bremen

zur Erlangung des akademischen Grades eines
DOKTOR-INGENIEUR (Dr.-Ing.)
genehmigte Dissertation

von
Dipl.-Ing. Nico Hellwege
aus Bremerhaven

Referent:	Professor Dr.-Ing. St. Paul
Korreferent:	Professor Dr.-Ing. W. Nebel
Eingereicht am:	29.09.2015
Tag des Promotionskolloquiums:	11.12.2015

To Anneliese.

Abstract

Reliability of integrated circuits fabricated using deep-submicrometer technology nodes is significantly affected by variations in process parameters, supply voltage and ambient temperature (PVT). More recently, also time-dependent degradation modes in complementary metal oxide semiconductors (CMOS) have risen as a key aspect in a reliability-driven analog integrated circuit design, as aging mechanisms additionally threaten a reliable operation over lifetime. These so called PVTA variations need to be considered during circuit design, in order to satisfy given specification limits and reliability constraints. Whereas in the digital domain methods handling these shifts have already been established over the past decades, the analog domain is still in need of aging-aware design methods, which help to solve the challenges caused by degrading circuit characteristics.

In practice, most analog designs start with hand calculated estimates of circuit characteristics, which need to be proven and refined by SPICE equivalent circuit simulations. However, the impact of degradation modes on circuit characteristics is first investigated after all design targets have been set. Special reliability simulators use complex models of degradation effects in order to compute expected deviations in circuit behavior based on a series of different SPICE simulations. This scheme results in long iterative simulation runs and leaves no insight to the designer on how to optimize towards higher reliability.

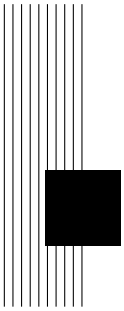
The focus of this thesis is on the development and implementation of aging-aware design methods, which are suitable to satisfy current needs of analog circuit design. Based on the well known $g_\mathrm{m}/I_\mathrm{D}$ sizing methodology, an innovative tool-assisted aging-aware design approach is proposed, which is able to estimate shifts in circuit characteristics using mostly hand calculation schemes. The developed concept of an operating point-dependent degradation leads to the definition of an aging-aware sensitivity, which is compared to currently available degradation simulation flows and proves to be efficient in the estimation of circuit degradation. Using the aging-aware sensitivity, several analog circuits are investigated and optimized towards higher reliability. Finally, results are presented for numerous target specifications.

Kurzfassung

Die Zuverlässigkeit von integrierten Schaltkreisen, die unter Verwendung von Technologieknoten im Submikrometerbereich hergestellt werden, unterliegen signifikanten Variationen der Prozessparameter, Versorgungsspannung und Temperatur (PVT). Unlängst erhalten auch zeitabhängige Degradierungsmoden eine Schlüsselrolle im zuverlässigkeitsorientierten Entwurf analoger integrierter Schaltungen, da diese Alterung zusätzlich den Betrieb über die Lebenszeit beeinflusst. Diese sogenannten PVTA Variationen müssen im Entwurf berücksichtigt werden, um gegebene Spezifikationen der Zuverlässigkeit einzuhalten. Während in der digitalen Domäne bereits seit Jahrzehnten Verfahren zur Einhaltung von Spezifikationen unter Berücksichtigung von PVTA Variationen existieren, werden diese Methoden in der analogen Domäne noch vermisst.
Im Allgemeinen beginnt der Entwurf von analogen Komponenten durch Berechnung von einfachen Gleichungen ohne Rechnerunterstützung. Die abgeschätzten Charakteristiken werden basierend auf einer Netzliste durch eine SPICE äquivalente Simulation gestützt und verfeinert. Die Auswirkung der Degradationsmoden auf Schaltungseigenschaften wird erst nach der Festlegung aller Entwurfsziele betrachtet. Spezielle Zuverlässigkeitssimulatoren verwenden komplexe Alterungsmodelle zur Berechnung der erwarteten Abweichung in den Schaltungseigenschaften, aufbauend auf einer Reihe unterschiedlicher SPICE Simulationen. Hierdurch entstehen lange Simulationszeiten und es ergeben sich keine Einblicke für den Entwurf der Schaltung, durch welche Änderungen eine Erhöhung der Zuverlässigkeit erzielt werden kann.
Der Schwerpunkt dieser Arbeit ist die Entwicklung und Implementierung einer Entwurfsmethode für analoge Schaltungen, die die Berücksichtigung von Alterungsmechanismen bereits im Handentwurf ermöglicht. Basierend auf der bekannten $g_\mathrm{m}/I_\mathrm{D}$ Dimensionierungsmethode wird ein programmgestützter Entwurfsansatz vorgestellt, der auf Basis von Handberechnungen die alterungsabhängige Verschiebung von Schaltungseigenschaften bestimmen kann. Das entwickelte Konzept der arbeitspunktabhängigen Degradierung führt zur Definition von alterungsabhängigen Sensitivitätsfeldern, die mit bestehenden Simulationsabläufen verglichen werden und sich als effiziente Methode zur Approximation von Alterungseinflüssen herausstellt. Mit Hilfe dieser Methodik werden verschiedene Schaltungen untersucht und auf höhere Zuverlässigkeit hin optimiert. Abschließend werden Ergebnisse für unterschiedliche Schaltungseigenschaften präsentiert.

Foreword

This work originated as part of my research activities as a doctoral student and research assistant at the Institute of Electrodynamics and Microelectronics (ITEM) at the university of Bremen at the department of mobile communication. I would like to thank Prof. Steffen Paul for the supervision of this work and also for the comments, hints and methodical advices I was given over the years. Also I have to thank Prof. Wolfgang Nebel of the university of Oldenburg for being the second reviewer of this thesis. I thank Prof. Walter Lang and Prof. Alberto Garcia for the examination at the colloquium.
I have to thank all of my colleagues, which I have had the pleasure working with over the past years. With some of them I had worked very closely and had numerous discussions and debates, which enriched this thesis as well as my personal development. Specifically, I want to thank Nils Heidmann, for sharing his scientific abilities with his officemate and also for his ongoing encouragement during our stay at the ITEM. Secondly, I have to thank Jonas Pistor for his ideas, advices and inspiring talks. I want to thank Dagmar Peters-Drolshagen for her support and encouraging advices and also Valerie Gerdes, for being her my omniscient language consultant and for always helping me out. Great thanks also to the Michael Peeper and the technical staff, which I have worked very closely with.
All of this would not have been possible, without the continuous support from my family, especially my parents Monika and Jürgen, and also from my partner Sandra, who had always taken care of me with her love, strength and understanding.

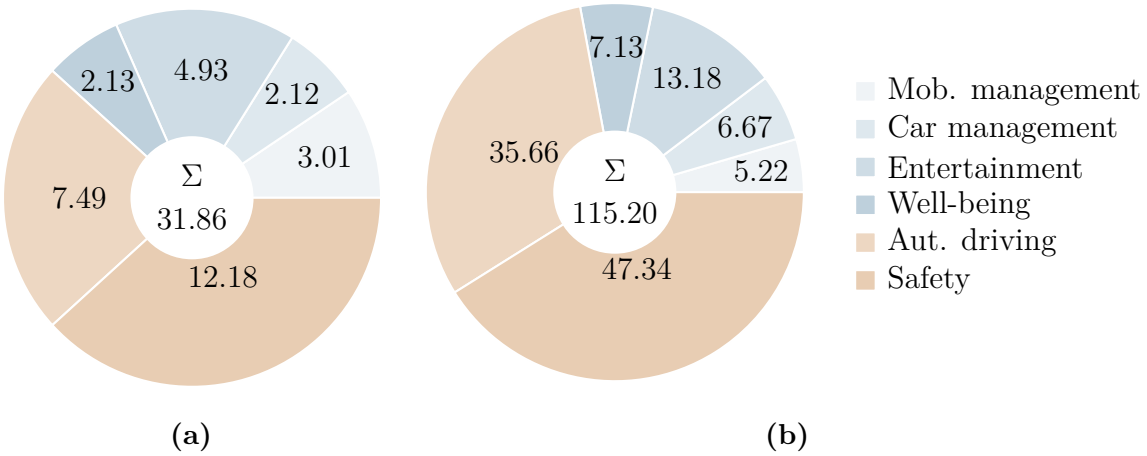


Contents

1 Introduction

Design and development of electronic components with reliable operational lifetimes of several years or even decades has become a key challenge in several fields of safety-critical applications, such as automotive, avionic or medical. The need for reliable electronics is seen for all three of these fields. In the case of automotive applications, two fields are emerging at a rapid development pace: safety systems and components for autonomous driving. A recent study [117] shows a huge increase in the potential market share for automotive electronics from 2015 to 2020, which is almost entirely driven by an increase for safety systems and components for autonomous driving, as seen in Fig. 1.1. The total market share increases from 31.86 billion Euro to 115.20 billion Euro, where the difference of 83.34 billion Euro is almost completely build on the increase in those to sectors (63.33 billion Euro). Avionic manual systems have been replaced by electronics over the past decades, e.g. in fly-by-wire systems [37]. Obviously, those systems need to be highly reliable during flight operation. This not only includes the algorithmic core, but also sensors and measurement systems, which are distributed throughout the aircraft. Most of these exterior components are composed of mixed-signal circuits with a high percentage of analog circuitry. A main relief on the constraints of reliability for automotive and avionic electronic systems is the ability to replace parts as needed (or demanded by service intervals). In medical applications, such as a pacemaker or an intracortical neural measurement system, a change of components is not easily performed, as this mostly involves medical surgeries. These surgeries put patience in danger, which should be avoided under any circumstances. All of the above fields of electronic applications build upon reliable systems and circuits, which need to satisfy high standards.

The design of integrated circuits usually incorporates the use of transistor models, which describe the behavior of a device directly after fabrication. However, there exist several physical degradation mechanisms, which alter transistor parameters over time, leading to degraded circuit characteristics. One of the most discussed modes are negative bias temperature instability (NBTI) and hot carrier degradation (HCD), which both lead to a shift of the threshold voltage of a transistor. These physical degradation modes are a major source of reliable issues, which need to be addressed, in order to assure a proper operation of the system.



(a) **(b)**

Figure 1.1: Developement of potential market share considering the automotive sector in billion Euro. Estimates for (a) 2015 and (b) 2020 [117].

When these reliability issues first occurred, the design flow for developing analog integrated circuits had not been adopted to the upcoming challenges. In the digital domain an impact of degradation is mostly seen as a shift in a gate delay, which could already be analyzed with existing yield and timing analysis tools. The analog domain is based on a completely different development flow and the impact on circuit level is seen in many different performance characteristics. However, decades ago the measured influence on analog circuits seemed little, so not attempt was made to account for degradation in the design flow. As smaller technology nodes evolved, the impact of degradation effects on analog circuits increased. Since the design flow has not changed, an estimation of reliability (or lifetime) has only been performed at the very end and has more been considered as a post-schematic analysis. This is mostly due to the complicated nature of the degradation modes, for which specific simulators are needed to calculate the estimated deviations in behavior. Typically, a simulation addressing transistor reliability effects is at least two to three time slower, than a regular circuit simulation. Both of these factors, the need of additional simulators and the enormous increase in the computational runtime of the simulators, has led to the situation that even today, no effort is made to put these steps to earlier design phases.
However, as degradation and reliability become more and more a key aspect in circuit design, these issues should be dealt with as early as possible in the design flow, as major decisions are made in the very beginning of the design flow: the hand calculation phase. Although modern research approaches try to cover the aspect of aging circuits in early design phases, all of them basically still rely on complete circuits, which again puts the investigation of degradation to end of the design flow. An approach, where the estimation of reliability effects is brought to the hand calculation phase, has not been thoroughly investigated so far.

1.1 Contributions by the Author

The core contribution of this work is an aging-aware design methodology for analog integrated circuits, which provides methods and tools allowing a consideration of degradation-induced deviations of circuit characteristics at the very beginning of the design flow. This is achieved by using normalized small signal parameters, which degrade in dependence on the operating point of the device [51]. The notion of an operating point-dependent degradation leads to the definition of a so called aging-aware sensitivity (AAS) [54] [48], enabling designers to take degradation into account as an analytical parameter.
In order to allow an investigation on reliability effects at an early design phase, a proper design method needs to be chosen and evaluated first. In general, there exist several different design methods, which account for simple hand calculations results. The majority of these methods are based on Shockley's quadratic MOS equations, which have been used for years. However, modern technology nodes do not follow these models anymore, which is why innovative charge-based MOS models, such as the EKV MOS model, have led to a different approach: the g_m/I_D sizing methodology. Within this work the g_m/I_D methodology is chosen as the basis for a degradation analysis on transistors. The normalized small signal parameters, which are taken from the g_m/I_D method, are expanded to account for a transistor degradation, which is dependent on the operating point of the device. This phenomenon leads to the definition of a new operating tuple, which accounts for the aging behavior and still is fully compatible with the already known sizing methodology. The methods developed around the operating point-dependent degradation have been implemented into a math-assisted tool, which provides easy access to all transistor parameters within the g_m/I_D regime at a fresh and also at a degraded state. This GMID-Tool is also able to calculate an estimated degradation of circuit characteristics by using analytic performance expressions [55].
The results of the operating-point degradation analysis are used to estimate deviations in circuit characteristics over time, by solely relying on hand calculation approximations. Several circuit characteristics are investigated on different analog circuits. Further, a complete analysis method is developed, which is able to predict the degradation for any analytical performance expression, as it is usually provided in analog circuit design [54] [48]. The definition of the aging-aware sensitivity is directly built upon the concept of the operating point-dependent degradation.
Also other aging-aware design methods for analog circuits have been explored. A topology driven design method for increasing the reliability of amplifiers has been proposed in [50]. In this study, different bias circuits and current mirror topologies are compared in regards of their response to certain degradation modes. The results are used in a folded cascode amplifier design and lead to a considerable increase in reliability.
In order compute reliability figures out of a statistically generated batch of samples, an efficient simulation flow has been developed, which accounts for process variability and

degradation mechanisms simultaneously. The variability-aware gradual aging (VAGA) approach [52] reduces the overall simulation effort compared to a standard Monte Carlo flow.

1.2 Structure of this Work

In chapter 2 the fundamental terms yield and reliability are introduced. Also, a nomenclature is presented, which is used throughout this work. After that, different aging mechanisms and the according modeling approaches are described, introducing hot carried degradation as well as bias temperature instability. At last, a brief overview of current aging simulation tools is given.
Chapter 3 begins with a study on different current mirror and bias circuit topologies. Multiple combinations are investigated in regards of their response to degradation modes. After that, an innovative variability-aware gradual aging simulation flow is presented, which generates reliability figures from a set of distributed samples. The result of the structural bias circuit and current mirror study are used to increase the aging behavior of folded cascode amplifier, as used in a neural measurement system.
In chapter 4 the concept of operating point-dependent degradation and the underlying g_m/I_D sizing methodology is introduced. These are used to investigate aged performance characteristics of a common source amplifier. After that, the GMID-Tool is presented and used to demonstrate an aging-aware design of a Miller amplifier.
Chapter 5 presents the aging-aware sensitivity (AAS) method for analog circuit design. An analysis on a common source amplifier demonstrates the AAS method by generating so called AAS-maps. A validity analysis is shown to explore the extend of the AAS and its boundaries. At last, optimum aging-aware operating points are found through the use of the AAS method.
In the last chapter a summarized conclusion of each of the main contributions of this work is given. Also, an outlook to the development of the AAS method is presented and discussed.

2 Reliability in CMOS Circuits

A system based on integrated circuits is designed to perform a specific function. The range, in which the system operates, is limited by specifications and only if these are met, the system properly fulfills its task. A designer is responsible for the design of the system and its ability to operate within these limits. These designs are based on models and assumptions, which describe the behavior of the circuit and its inherent transistors in a most ideal way. However, imperfections within the CMOS process scheme cause deviations from the model, leading to small changes in geometries or other process related metrics. The influence of these imperfections act spatially as well as temporally, some are systematic and others random, but all of them limit the functionality of the design and may even cause failures, which make a system unreliable.
The analysis of these failures on circuit or system level demands a fairly deep understanding of the underlying unreliability issues on transistor and process level. Not all of these unreliability effects apply to all levels of design, e.g. some effects need to be addressed at layout level, others at schematic level, and the impact even differs in dependence on the technology node. It is therefore important to understand the extend and applicability of each effect and how these alter characteristics of an integrated circuit design.
This chapter first introduces a terminology around the notion of reliability. After that, the most common degradation modes are compared using a spatial basis, which nicely describes the extend of each effect from atomic to die level. Each of these effects is then described in subsequent sections. A brief introduction to common reliability simulators is given. At last compact models for the most important aging mechanisms are described.

2.1 Fundamentals

The analysis of unreliability issues in CMOS environments is mathematically described through stochastic definitions and expressions. The formalism, which is described in this section, derives a basic understanding of the terms yield and reliability and how

these relate to metrics found in the area of integrated circuits. Both terms are linked to the reliability function and failure rate, which is discussed afterwards.

2.1.1 Yield and Reliability

Yield and reliability are strongly connected and, as it will be shown, may be used interchangeably, within a certain extend. Whereas reliability is a function of time, the term yield is usually not connected to a transient behavior. The definition of yield is used in the semiconductor processing in order to provide a metric on how many samples of a circuit passed the post-fabrication test successfully. Optimizing the yield of a batch involves many different aspects, but the following two major sources can be regarded as the most important ones: a well defined and calibrated technology node, which provides accurate MOS models for the designer, and also circuit designs, which are able to withstand variabilities, which occur during the fabrication process. Calibration and modeling of the CMOS process is usually provided by the processing semiconductor foundry and is of no concern for the designer. In this work, the focus is given from the designer's view, leaving processing aspects aside.
In order to distinctively define the term yield, a special nomenclature is used, which will be introduced in the following. Let a specific design of a circuit be defined as τ, which is, e.g., a schematic. A **sample** c is a specific incarnation of τ, where process related variations are taken into account. The vector containing m samples of a topology τ shall be defined as

$$\mathbf{c}_\tau^\top = [c_1 \; c_2 \; \cdots \; c_m]. \tag{2.1}$$

A topology consists of multiple transistors, each defined by various parameters (V_{th}, g_{m}, g_{ds}, i_{ds}, etc), which are dependent on the operating point tuple $\mathbb{T}_{\mathrm{OP}}$. Note that the placement of a transistor in a schematic, is bound to the topology τ, whereas the transistor parameter are dependent on the actual process corner and therefore bound to a sample. The vector containing all **transistor parameters** of a specific transistor k is defined as

$$\mathbf{d}^k_{\mathbb{T}_{\mathrm{OP}}}{}^\top = [d_0 \; d_1 \; \cdots \; d_w]. \tag{2.2}$$

Each circuit holds specific characteristics such as gain or bandwidth. A **performance value** is a scalar value, which represents a specific characteristic of a sample of a circuit. Ultimately, the characteristic of each sample is given by the transistor parameters $\mathbf{d}^k_{\mathbb{T}_{\mathrm{OP}}}$. The vector of n performance values is defined as

$$\mathbf{p}(c_j)^\top = [p_1(c_j) \; p_2(c_j) \; \cdots \; p_n(c_n)] \tag{2.3}$$

$$\text{with } c_j = c_j\left(\mathbf{d}^k_{\mathbb{T}_{\mathrm{OP}}}\right). \tag{2.4}$$

A **specification** sets limits, which a performance value needs to achieve. The set of n specification limits is defined as

$$\mathbf{s}_{\mathbf{p}}^{\top} = [s_{p1}\ s_{p2}\ \cdots\ s_{pn}]. \tag{2.5}$$

An **evaluation function** $f_{s_j}(c_i)$ probes whether the j-th performance value of sample c_i fulfills its specification limit s_j. It is defined as

$$f_{s_j}(c_i) = \begin{cases} 1, & \text{if } p_j(c_i) \text{ fulfills } s_{p_j} \\ 0, & \text{else.} \end{cases} \tag{2.6}$$

A specific sample c_i is considered to be a **good sample**, if all performance values fulfill their specification limits, such that

$$f_{s_j}(c_i) = 1\ \forall\ j = \{1, 2, \ldots, n\} \tag{2.7}$$

or in equivalence, a good sample fulfills

$$\prod_{j=1}^{n} f_{s_j}(c_i) = 1. \tag{2.8}$$

The quotient of good samples to all samples is considered to be the yield of the circuit (topology), which shall be denoted as Y_τ:

$$Y_\tau = \frac{\sum_{i=1}^{m} \prod_{j=1}^{n} f_{s_j}(c_i)}{m}. \tag{2.9}$$

The definition from Eq. (2.9) describes how many samples have been processed, which are unable to fulfill certain specification limits. The reason of failure is a design, which is unable to compensate for variability. Once a sample is processed and operates, time-dependent issues arise, which change the behavior of the sample over time and failures or dysfunctions will occur eventually. These in-field failures are not taken into account by the definition of yield in Eq. (2.9). Therefore, a new metric needs to be introduced.
The probability, that a given sample fulfills all its specifications for a certain operating time t is defined as reliability [36]. The definition of reliability can directly be determined from (2.9), if c_i (and therefore $f_{s_j}(c_i)$) is considered to be time-dependent. Since the performance value of a sample is dependent on the transistor parameters, these need to be dependent on the operating time

$$p_i(c_j(t)) = f(\mathbf{d}^k_{\mathbb{T}_{\mathrm{OP}}}, t). \tag{2.10}$$

From this, the reliability R_τ is defined as

$$R_\tau(t) = \frac{\sum_{i=0}^{m} \prod_{j=0}^{n} f_{s_j}(c_i(t))}{m}. \tag{2.11}$$

The expressions in Eq. (2.9) and (2.11) show, that reliability may be interpreted as a time-dependent yield and correspondingly, yield as the reliability at time $t = 0$:

$$R_\tau(t) = Y_\tau(t) \tag{2.12}$$
$$Y_\tau = R_\tau(0) \tag{2.13}$$

As the difference for yield and reliability is its dependence on time, both are often referred to as fresh and aged yield respectively [79].

2.1.2 Reliability Function and Failure Rate

The definition from (2.11) links the term reliability to a time-dependent yield, which is useful for determining the fresh and aged yield of a specific design, e.g. through simulation and evaluation. In addition to this definition, there exist fundamental principles when modeling reliability. Aside from reliability, failure rate is an important indicator for time-dependent device properties.
In order link both terms, a different approach on defining reliability than the one in (2.11) has to be used. Let X_{life} be a random variable, which denotes the lifetime of a device and $F(t)$ shall correspond to its cumulative distribution function, which returns the probability of a device not exceeding a specific operating time t,

$$F(t) = \text{prob}(X_{\text{life}} \leq t), \tag{2.14}$$

then the reliability $R(t)$ may be written as

$$R(t) = \text{prob}(X_{\text{life}} > t) \tag{2.15}$$
$$= 1 - F(t). \tag{2.16}$$

The term $R(t)$ is referred to as the *reliability function* and $F(t)$ is the *lifetime distribution function* [75] [126]. The failure rate (or hazard function) $h(t)$ may be interpreted as the frequency at which a device fails (e.g. once per decade). If the failure of the device is non-recoverable, then $h(t)$ is an indicator for how fast the device approaches its end of lifetime. The definition can be derived through the use of $R(t)$:

$$h(t) = -\frac{\dot{R}(t)}{R(t)} = \frac{f(t)}{R(t)}, \tag{2.17}$$

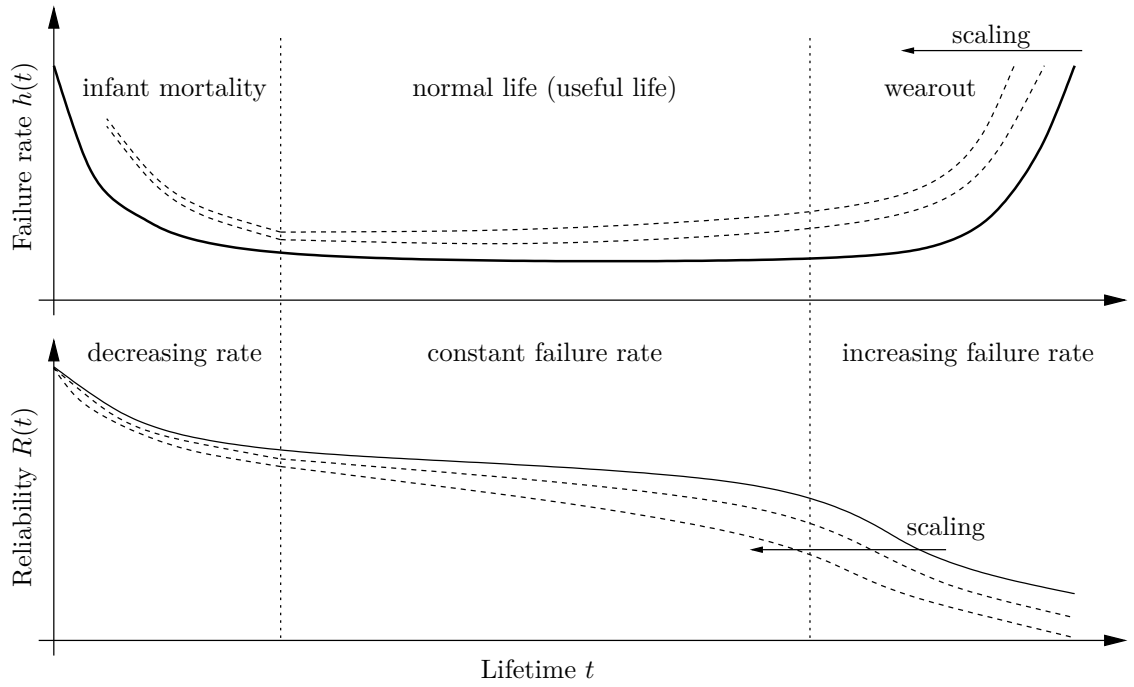


Figure 2.1: Failure rate (or hazard) function $h(t)$ and the corresponding reliability function $R(t)$ given for the lifetime of a device. Plot is divided into three regions: infant mortality, normal life and wearout, each with its corresponding failure rate behavior.

where $f(t)$ is the lifetime probability density function, which is defined as

$$f(t) = \frac{\mathrm{d}F(t)}{\mathrm{d}t} = \dot{F}(t). \tag{2.18}$$

The most common type of the failure rate function is the *bathtub curve* [70] [1]. The name is given by the shape of $h(t)$, which is depicted Fig. 2.1 (top). The corresponding reliability $R(t)$ function is shown underneath $h(t)$. Three distinct regions can be identified:

Infant mortality

Within this period the device fails due to initial defects or imperfections. The failure rate decreases until a steady state is reached. Each failing sample also decreases the reliability, consequently also $R(t)$ decreases.

Useful life

This is the normal mode of operation and the failure rate is approximately constant. The region is also known as intrinsic failure period. In general the value of $h(t)$ is very low in this period. As only a small amount of samples fails, the reliability function only decreases little.

Wearout

This last period is dominated by a steep increase in the failure rate. The raise in

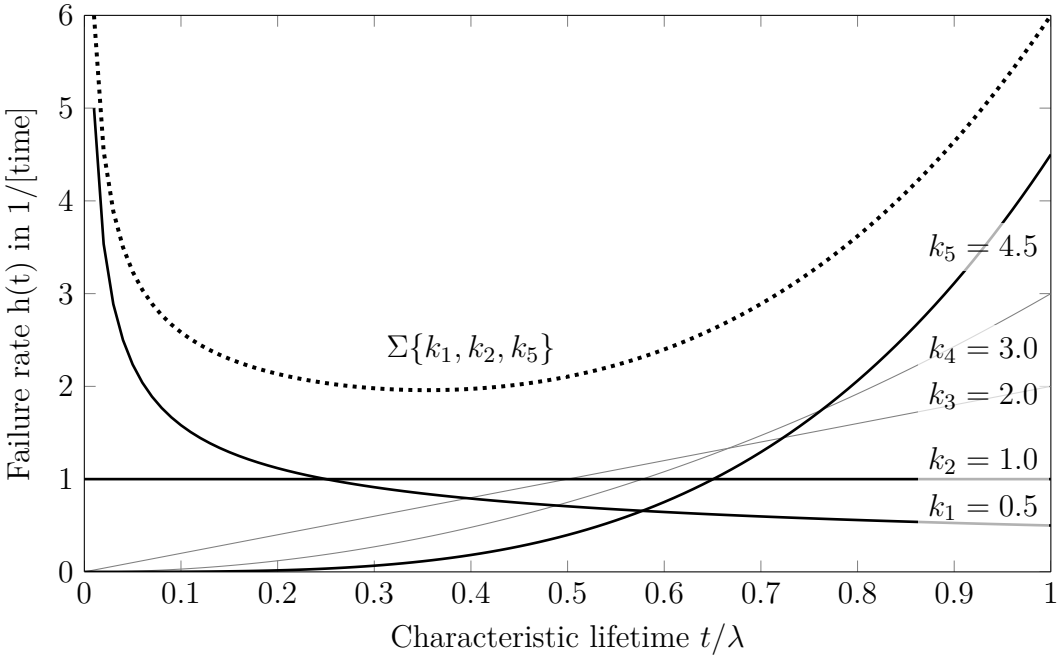


Figure 2.2: Weibull failure rate function $h(t)$ for different shape factors k over the characteristic lifetime t/λ. The dotted line is the sum of three Weibull functions generated by $k_1 = 0.5$, $k_2 = 1.0$ and $k_5 = 4.5$ and is also known as the bathtub curve. Weibull functions generated from $k_3 = 2.0$ and $k_4 = 3.0$ show how the shape factor influences the shape of the Weibull curve.

failure rate forces reliability to steadily decrease until the last sample failed (approaching $R(t) \to 0$). The main cause of failures are time-dependent unreliability issues.

There exist different methods to model the failure rate. One of the most common models is the use of an exponential failure distribution

$$F(t) = \int_0^t \lambda e^{-\lambda\tau} \, \mathrm{d}\tau = 1 - e^{-\lambda t}, \tag{2.19}$$

which according to (2.17) and (2.16) results in a reliability and failure rate function of

$$R(t) = e^{-\lambda t} \tag{2.20}$$

$$h(t) = \lambda. \tag{2.21}$$

In order to describe the lifetime and reliability of electronic components, the Weibull distribution is used throughout literature. The distribution is given as

$$F(t) = 1 - \exp\left[-(\lambda t)^k\right], \tag{2.22}$$

where λ is the scale parameter and k is the shape parameter. The hazard function of a Weibull distribution results to

$$h(t) = \lambda k(\lambda t)^{k-1}. \tag{2.23}$$

The bathtub curve from 2.1 can be modeled using a superposition of several Weibull failure functions [126], which is illustrated in Fig. 2.2. Here, the sum of three failure functions results in a representation of the bathtub curve, where $k < 1$ is used to model infant mortality, $k \approx 1$ models the constant failure rate and $k > 1$ mimics the wearout. The time line can be adjusted to fit any lifetime using the scaling factor λ.
As seen in Fig. 2.1, scaling has a huge impact on device reliability. This phenomenon consists of several individual mechanisms. Scaling reduces the minimum channel length of transistors and therefore leads to steadily decreasing geometries. During the process of CMOS circuits, defect particles cause errors in the lithography, which highly depend on the size of the defect particle. The impact of a defect particle of a specific size increases, as scaling reduces the size of the transistors. This mainly results in an increase of infant mortality.
Secondly, wearout mechanisms also increase, as scaling not only decreases device dimensions, but also the thickness of the gate dielectric. A thin dielectric is more likely to conduct a higher leakage current, than a thick equivalently stressed one under the same voltage conditions. This increase in leakage induces effects like time-dependent dielectric breakdown (TDDB) and ultimately bias temperature instability (BTI). Due to scaling, these time-dependent wearout modes cause a device to fail earlier for each new technology node [121].
Both scaling induced mechanisms, an increase of infant mortality and the impact of wearout modes, worsen the reliability of upcoming technology nodes. As safety-critical applications cannot tolerate high failure rates, these applications are covered by established nodes. However, even if the device is fabricated using a rather large CMOS process, there still exist unreliability issues, which need to be addressed. The most common are discussed in the following.

2.2 Classification of Reliability Issues

A summary of physical degradation and unreliability modes is illustrated in Fig. 2.3. In contrast to other illustrations [27] [17] [79], the modes are ordered by their spatial influence, which can be thought of as a spatial wavelength or characteristic length, ranging from the largest to the smallest. The classification in order of spatial influence is chosen in accordance to Pelgrom's mismatch law [87], which is further discussed in section 2.2.3. The largest spatial influence indicates an effect which acts on wafer level, whereas the smallest represents an influence on atomic scale. The center of action is virtually placed inside the channel of a transistor, such that effects with a spatial range

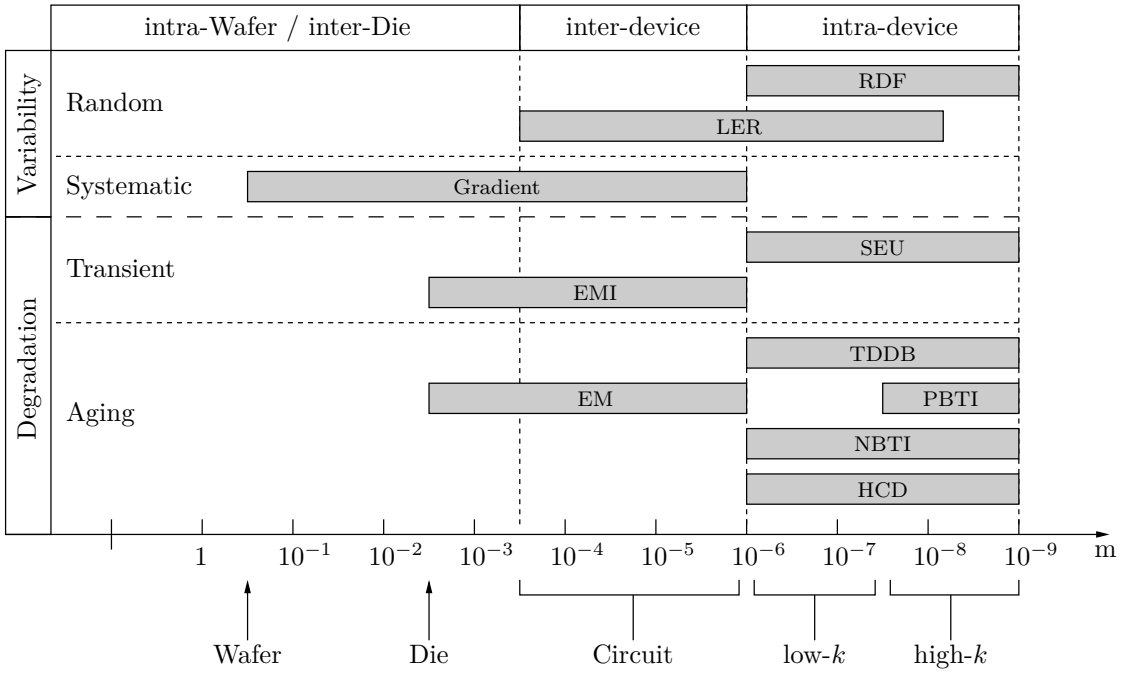


Figure 2.3: Spatial view on CMOS variability and reliability effects.

less than a channel length are considered to act inside the channel area. This scheme helps to understand which mode has to be regarded at what scale, e.g. the separation of inter- and intra-device modes. Depending on the used technology nodes, the spatial classification depicted in Fig. 2.3 may change due to different wafer or device sizes. The illustration is chosen to follow an artificial technology node with a channel length of around 1 µm and a wafer diameter of about 8 inch.

As shown in the last section, yield and reliability are strongly linked, since since $Y_\tau(t)$ is just a time dependent yield, which is exactly the definition of $R_\tau(t)$. In equivalence, the terms variability and degradation are linked in the same way. Variability describes effects, which are responsible for a statistical spread of process and transistor parameters after fabrication. Degradation shall be defined as all effects, which act during the lifetime of a circuit.

Variability effects occur immediately after production and do not depend on the operating time. This mode consists of random components (e.g. random dopant fluctuations (RDF) and line edge roughness (LER)) and systematic components (e.g. process gradient effects). These effects highly depend on the actual layout geometry and are mostly due to a non-idealistic lithography. Minor differences in the projection from the mask result in small deviations of transistor and circuit characteristics and thus determines the fresh yield of a circuit. Degradation effects are time-dependent and further separated into transient modes (e.g. electromagnetic interference (EMI), single event upsets (SEU)) and aging modes (electromigration (EM), bias temperature instability (BTI), etc.). Transient modes occur once in time and lead to a sudden deviation or failure and usually do not leave a permanent change in the circuit. Aging modes are

operating-time-dependent and may result in a permanent shift of circuit characteristics. The random and systematic part is equivalent to the separation of degradation into transient and aging effects, as these either occur randomly or systematically in time. Note that the classification borders given in Fig. 2.3 are not considered to be sharp, since relative dimensions of a wafer or die change over different technology nodes.
In the following sections, variability and degradation effects are described. Aging effects are discussed in a separate section, as the remaining chapters of this work highly focus on these degradation modes.

2.2.1 Variability

Process variability (or static reliability) becomes worse for decreasing geometries in newer technology nodes. As complexity in lithography increases, e.g. through the use of double pattering [85] or stepper motion [127], it is more and more difficult to ensure proper development of CMOS layers. Variability is usually distinguished in local (or intra-die) variations and global (or inter-die) variations. Intra-die variability causes mismatch in transistors or circuits, which originally were designed to be identical. The spatial impact only affects neighboring devices or short distanced circuits. Inter-die variability addresses effects, which act from die to die, even across a whole wafer. This occurs, if dies or wafers are processed at a different time or in a different manufacturing chain. Intra-die variation changes the mean value of process related metrics (e.g. dopant concentration or threshold voltage). Most issues in unreliability originate from local variations [7], which will be further discussed in the following.
Local variability is further differentiated into random and systematic effects. As shown in Fig. 2.3, these effects act on different spatial ranges. From a variability perspective, both may be seen as the same statistical process variable at different spatial frequencies. This idea has first been formulated in [87], but outside the context of degradation. The illustration in Fig. 2.3 expands this view by the addition of time-dependent degradation effects. In the following the most important variability sources are discussed.

2.2.2 Systematic Inter-device Effects

Most of these effects are more general and are described as offsets of gradients. A gradient describes the change of process parameters in dependence on the area coordinate across a wafer, which affects e.g. the oxide thickness or dopant concentration. These effects can be seen as the systematic time-independent part of each individual unreliability effect. As channel lengths decrease in modern technology nodes, these effects will become more apparent and a major limitation to circuit characteristics [2]. Systematic effects like gradients in layout are addressed by symmetric transistor positioning, e.g. a common centroid layout for transistors [46].

2.2.3 Random Intra-device Effects

Random variability has a short spatial characteristics length and therefore acts on an inter-device level. Most described phenomena are random dopant effects, line edge and width roughness, trapped charges at the gate-channel interface and also variations in oxide thickness resulting from interface roughness [79] [2]. Intra-device effects add mismatch to circuit elements, which are designed to be identical and are positioned closely in the layout. As the spatial range is limited to the boundaries of the device, these effects are statistically independent from one device to another. A general formalism describing how area-dependent variability influences the mismatch of transistors is given by Pelgrom [87].
Let $P(x, y)$ be an arbitrary process parameter, which is dependent on the die coordinates x and y, then the mismatch ΔP of two identical parameters located at different coordinates defined by (x_1, y_1) and (x_2, y_2) is given by [87]

$$\Delta P(x_{12}, y_{12}) = \frac{1}{A}\left(\iint_{A(x_1,y_1)} \rho(x', y')\ \mathrm{d}x\,\mathrm{d}y - \iint_{A(x_2,y_2)} \rho(x', y')\ \mathrm{d}x\,\mathrm{d}y\right). \tag{2.24}$$

Using a two-dimensional Fourier transformation, it is possible to separate the geometry-dependent part and the mismatch source[1]:

$$\Delta\mathcal{P}(\omega_x, \omega_y) = \mathcal{G}(\omega_x, \omega_y)\mathcal{P}(\omega_x, \omega_y) \tag{2.25}$$

According to [87], these parts can be regarded as a mismatch source $\mathcal{P}(\omega_x, \omega_y)$, which generates spatial frequencies, and a geometry dependent filter function given by $\mathcal{G}(\omega_x, \omega_y)$. The concept of spatial frequencies is the basis for the diagram in Fig. 2.3, where each unreliability effect is treated as a mismatch source with different spatial bandwidths, ordered by the spatial wavelength.
The description in (2.25) can be used to determine the variance of ΔP. The solution leads to Pelgrom's model:

$$\sigma^2(\Delta P) = \frac{A_P^2}{WL} + S_P^2 D_x^2 \tag{2.26}$$

where A_P is a technology dependent area proportionality constant for P and S_P is the variation of parameter P with respect to the spacing D_x, WL is the are of the device. Unreliability modes, which only act within a device, are mostly uncorrelated to the distance and therefore can be described by a first order approach [79]:

$$\sigma^2(\Delta P) \approx \frac{A_P^2}{WL} \tag{2.27}$$

[1] The integral in (2.25) may be interpreted as the convolution of double box functions formed by the integral boundaries [87].

At device level, most variations are seen in the gate length L, the threshold voltage V_{th} and the current factor β [129], which is modeled using (2.27)

$$\sigma(\Delta V_{\text{th}}) = \frac{A_{V_{\text{th}}}}{\sqrt{WL}}. \tag{2.28}$$

The mismatch constant $A_{V_{\text{th}}}$ scales down with each new technology node, but is mostly limited by the accuracy of the lithography and also other factors. There exist two major effects, which need to be discussed: random dopant fluctuation and line edge (and width) roughness.

Random Dopant Fluctuations

Variation in the threshold voltage partly results from a nonuniform distribution of dopant atoms. As technology scales down, the influence and variation of the number of dopant atoms increases (because the total number of channel atoms also decreases) [74] [109]. In [74] it is shown, that random dopant fluctuations is responsible for more than 60 % of the V_{th} mismatch of identically designed NMOS transistors in a 65 nm process. These results have been verified further in a 45 nm technology node. The variation on the threshold voltage based on RDF is mostly modeled by [106]

$$\sigma_{\text{rdf}}(\Delta V_{\text{th}}) = \Phi \cdot \frac{t_{\text{OX}}}{\epsilon_{\text{OX}}} \frac{\sqrt[4]{N}}{\sqrt{W_{\text{eff}} L_{\text{eff}}}}, \tag{2.29}$$

where Φ is a technology and temperature dependent constant, t_{OX} is the gate oxide thickness, ϵ_{OX} is the oxide permittivity, N is the number of channel dopants, and W_{eff} and L_{eff} are the effective width and length of the channel respectively.

Line Edge and Width Roughness

The effect of line edge roughness (LER) or line width roughness (LWR) is a result of sub-wavelength lithography, which essentially describes the fact that the wavelength, which is used in the lithography process, is larger than the developed device area. Sub-wavelength lithography is used since the 0.25 µm technology node, where a lithography wavelength of $\lambda_{\text{lit}} = 248$ nm is used. Modern nodes below 130 nm use a light source achieving a $\lambda_{\text{lit}} = 193$ nm, which is still used in 45 nm and below. LER describes the roughness of the edge of line, whereas LWR is the change in width of a line due to LER. Both mechanisms are depicted in Fig. 2.4.

2.2.4 Transient Degradation Effects

Transient effects occur after production, but do not depend on the actual operating time of the circuit but are transiently random. In general these modes are induced by

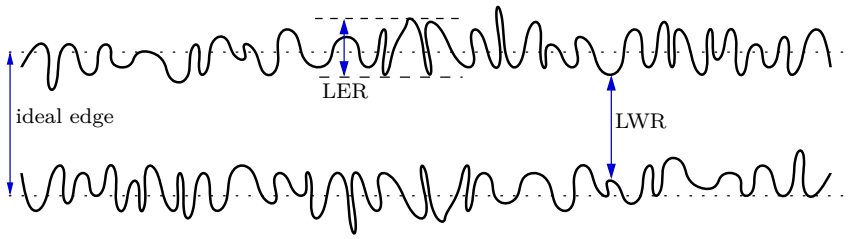


Figure 2.4: Line edge roughness and line width roughness [79].

external events, which affect the circuit at that specific time. Two major effects in this class are single event upsets (SEU) [66] and electromagnetic interference (EMI) [67]. An SEU is caused by single ionizing particles, such as ions, electrons, photons, etc. These particles enter a sensitive area in a CMOS device and change its state, e.g. flipping a bit. This change is induced through the charge, which is created by the ionization. The error seen at the device output is called an SEU or a soft error. EMI is a disturbance which is caused by emitted electromagnetic induction or electromagnetic radiation from an external source. In contrast to SEU, EMI can affect multiple devices at a time. Both SEU and EMI are not considered to permanently damage transistors and each of the effects can effect digital as well as analog circuits, but clocked gates are more prone to these modes.

2.2.5 Other Variability Sources

Despite the above mentioned sources of variability effects, several others have been identified in modern technology nodes, e.g. random discrete dopant (RDD) [123], metal gate granularity (MGG) [14] or layout dependence effect (LDD) [107]. No list would be able to accumulate every existing effect and the importance for each of them is highly dependent on the used technology node, circuit topology and the runtime environment.

2.3 Aging

In contrast to variability effects, degradation occurs after production and is a function of time. Degradation mechanisms, which depend on the operating time (despite being randomly distributed over time) are classified as aging. As seen in Fig. 2.3, there exist multiple aging mechanisms. Each of them acts with different spatial wavelength, or in other words, is active for different minimum channel lengths. In this section, each of these aging mechanism is explained by their physical origin and also how these are modeled in simulation environments.

2.3.1 Physical Origin of Aging Mechanisms

Aging or operating time-dependent degradation is perceived as a shift in a transistor parameter, where the magnitude of the shift changes over time. Degradation modes, which follow this definition are hot carrier degradation (HCD), bias temperature instability (BTI) (and its discrete variant random telegraph noise (RTN)) and time-dependent dielectric breakdown (TDDB). Each effect has different modeling approaches and varies in the impact on transistor parameters, but the physical origin is often overlapping or even identical. The most common physical degradation modes are interface state generation, electron or hole trapping and deep oxide traps, which are discussed in more detail in the following.

Interface State Generation

During the fabrication process of a MOS transistor, at some point the gate oxide is deposited over the channel area. The oxide consists of silicon dioxide (SiO_2), which is often referred to as silicon oxide, and the bulk material consists of doped silicon (Si). Both of these crystalline structures differ in their geometrical configuration, e.g. different lattice constants. This leads to crystallographic defects along the interface of channel and gate. Consequently, some chemical bounds are not satisfied, which results in free carriers and also in so-called dangling bonds. Each dangling bond provides an extra energy state between the valence and conducting band, which are referred to as interface states. The number of interface states is given by N_{it} and the change is given by the interface state generation ΔN_{it}.

Electron and Hole Trapping

If potential energy is supplied (e.g. through the gate potential V_{G}), free carriers will interact with interface states and get trapped. Depending on the type of carrier, this effect either is referred to as electron or hole trapping. The number of trapped charges in interface states is given by N_{et} or N_{ht}. As the channel reaches the state of inversion, even more free carriers are supplied, which fill up remaining traps. These traps accumulate a parasitic charge in the gate, which directly causes a shift in the threshold voltage. The number of production-induced interface states can be reduced, if the channel is exposed to hydrogen gas (H_2) before the SiO_2 is grown on top. The hydrogen fills up dangling bonds and reduces the number of remaining interface states after the oxide is deposited.

Deep Oxide Traps

Crystallographic defects do not only occur at interfaces, but also within the material. During the epitaxial growth of the SiO_2 dielectric, deep oxide traps are generated. These deep traps are well beyond the reach of channel carriers and are only reached through carriers with high kinetic energy.

The following time-dependent degradation modes are all based on the above discussed physical mechanisms. Each mode differs in its excitation and also in its ability to be reversible or not.

2.3.2 Hot Carrier Degradation (HCD)

If a carrier is accelerated under an electric field and gains a higher kinetic energy than the lattice in thermal equilibrium, it is considered *hot* [21]. These hot carriers can cause a significant damage to a transistor, if injected into certain parts of the structure. Hot carriers, which are generated inside the channel of a MOS device, are referred to as channel hot carrier (CHC) [125]. Hot particles have enough energy to induce impact ionization, generation of interface states or even cause defects in the oxide structure of the insulator at the gate-channel interface. Some of these carriers are also trapped by other interface states. All of these mechanisms lead to a degradation of electrical parameters of a transistor, such as the threshold voltage V_{th} or different transconductances like g_{m} or g_{ds}.
Hot carrier degradation is distinguished into four different mechanisms: channel hot electron (CHE) injection effect, drain avalanche hot carrier (DAHC) injection, secondary generated hot electron (SGHE) injection and substrate hot electron (SHE) injection [108] [29]. More recently these modes are separated into conductive and non-conductive hot carrier degradation [21], but the physical phenomenas are identical.

Channel Hot Electron (CHE)

The gate and drain potential V_{G} and V_{D} shall be equal in voltage, as depicted in Fig. 2.5a. In this state, only lucky-electrons are able to surpass the gate oxide and leave the channel, since for low voltages, no carriers are attracted to the gate and in the case of a high electric field at the drain DAHC is generated. Lucky-electrons are hot carriers, which have not been part of a particle collision inside the channel and are therefore able to continuously gain kinetic energy along the channel. These channel hot carrier result in a gate current I_{G} or are eventually trapped in deep oxide traps. As holes have a larger effective mass than electrons, NMOS devices are more susceptible to CHE.

Substrate Hot Electron (SHE)

The exciting configuration for this mode is shown in Fig. 2.5b. If the bulk voltage V_{B} is significantly positive or negative, depending on the device type, carriers in the channel gain energy and are more likely to pass the oxide barrier of the gate channel interface. This mechanism is independent of the drain and source potential and therefore uniformly distributed over the channel. This effect is mostly seen in circuits with a nonzero body bias. NMOS devices suffer from substrate hot electron injection, whereas in PMOS devices this effect is known as substrate hot hole (SHH) injection.

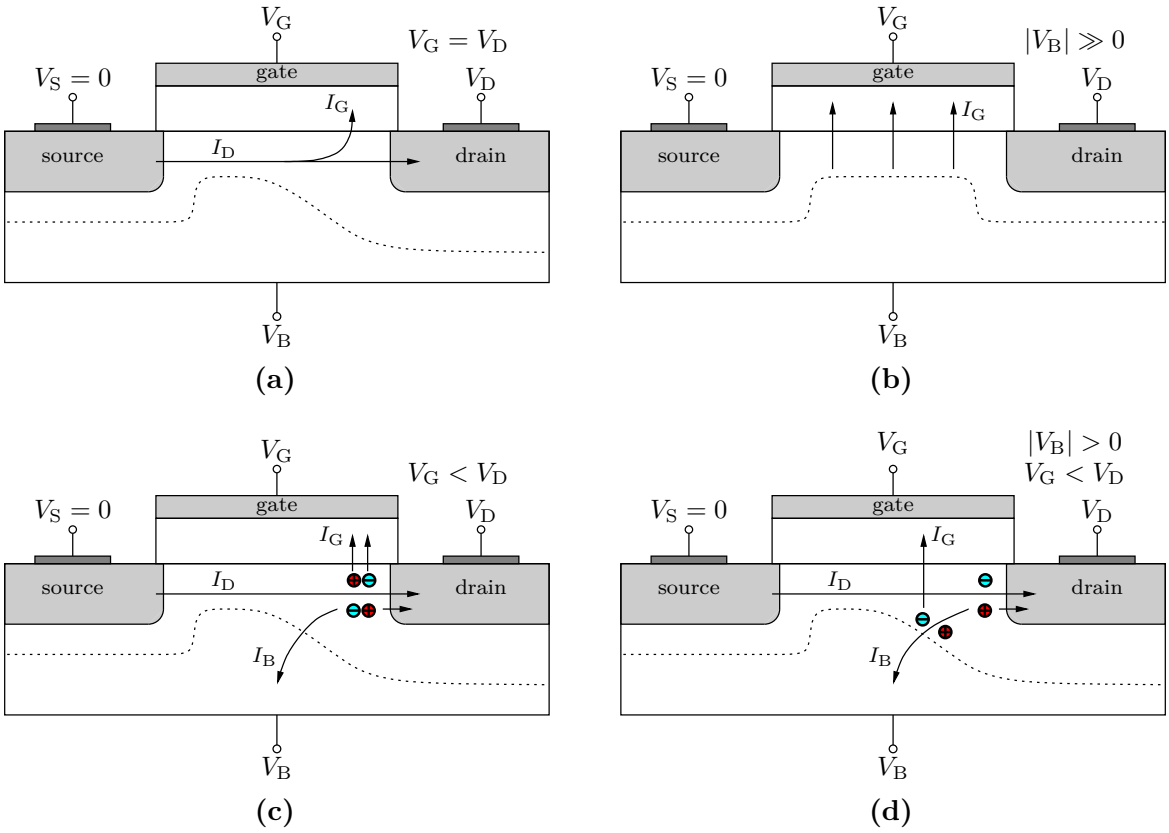


Figure 2.5: Different mechanisms of hot carrier degradation: (a) Channel hot electron (CHE). (b) Substrate Hot Electron (SHE). (c) Drain Avalanche Hot Carrier (DAHC). (d) Secondary Generated Hot Electron (SGHE).

Drain Avalanche Hot Carrier (DAHC)

If the transistor is operated in strong saturation, such that the drain voltage is high compared to the gate voltage, impact ionization occurs at the drain side of the channel, as shown in Fig. 2.5c. The generated electron hole pairs are now driven by the electric field inside the channel. Most carriers are attracted to the gate and cause further ionization or interface state generation. The remaining particles are accelerated along the channel according to their charge. Only few are drawn to the bulk and add to the bulk current I_B.

Secondary Generated Hot Electron (SGHE)

Carriers, which have been generated from high energetic collisions, such as in the case of DAHC, are also able to cause further impact ionization events. The now generated particles are called secondary generated hot carrier. This mechanism shown in Fig. 2.5d only contributes in small amounts to the overall HCD effect.

Long Term Modeling

Each of the described mechanisms contributes to the overall hot carrier degradation, but each is only active at a specific state of the transistor. Recently, DAHC has

been shown to be the most influencing contributor [29]. The long term shift in threshold voltage due to HCD is generally modeled using a power law [60] [72] [80] [57]

$$\Delta V_{\mathrm{th}} = A_{\mathrm{hcd}} \cdot t^{n}, \tag{2.30}$$

where n has a value of about 0.5 to 0.7. The generation of interface states is exponentially dependent on the oxide electric field E_{OX}. The factor A_{hcd} is further dependent on the maximum lateral electric field E_{lat}, the temperature T and also the channel length L [60] [80] [120]:

$$A_{\mathrm{hcd}} \propto \frac{1}{\sqrt{L}} \exp\left(\alpha_1 \cdot E_{\mathrm{OX}}\right) \exp\left(\frac{\alpha_2}{E_{\mathrm{lat}}}\right) \tag{2.31}$$

where α_1 and α_2 are technology dependent parameters.

2.3.3 Bias Temperature Instability (BTI)

BTI induced degradation is usually perceived as a shift in V_{th}, which is the result of high a bias voltage applied at the gate of a transistor. The impact of this effect is increased with a rise in temperature. Under normal conditions, the shift in threshold voltage can yield up to 30 mV, if a bias voltage is applied for more than 5 years (45 nm technology node [65]). BTI can also be responsible for a degradation in the carrier mobility μ [100], however, this is negligible compared to the shift in V_{th}.
There exist two types of BTI degradation: negative BTI (NBTI), which only occurs in PMOS devices and positive BTI (PBTI), which only occurs in NMOS devices. Both effects are controversially discussed in regards of the physical explanation and also the model to describe the degradation on transistor level. A general consensus has not been agreed on so far, however, several physical mechanisms have been identified to be responsible for NBTI and PBTI, which seem to be acknowledged and widely accepted by most authors. NBTI shall be the result of a combination of hole trapping ΔN_{ht} due to process related preexisting defects [6] [61] [92] and generation of interface states at the channel oxide interface ΔN_{it} [116] [71] [3] [19]. PBTI is believed to result from electron trapping in oxide traps in combination with trap generation [100] [64] [41]. Whereas NBTI is active for simple oxide gate dielectrics, PBTI only seems to be observed in high-k NMOS devices, where the impact is comparable or even worse than NBTI [42]. High-k metal gates (HKMG) NMOS and PMOS devices are used for technology nodes of 45 nm and beyond [82], which is also illustrated in Fig. 2.3 as a difference in NBTI and PBTI.
In contrast to HCD, BTI has the property to allow a relaxation or recovery of the shift in threshold voltage, if the applied gate potential is reduced. This phenomena is immediately revealed after the stress voltage is lowered [64]. Relaxation is unique to BTI and complicates modeling and extrapolation of BTI degradation. So far, it has not

been shown whether BTI is fully reversible or a permanent degradation remains [40], but most authors favor the idea of a permanent component [43].

Long term BTI model

In general BTI degradation is modeled following the approach of a recoverable and permanent component [78] [41] [79]:

$$\Delta V_{\mathrm{th}} \propto \left[\underbrace{\exp(\alpha_1 V_{\mathrm{GS}}) t^{n_{\mathrm{P}}}}_{\text{permament part}} + \underbrace{V_{\mathrm{GS}}^{\alpha_2}(C_{\mathrm{R}} + n_{\mathrm{R}} \log_{10}(t))}_{\text{recoverable part}} \right] \cdot \exp\left(-\frac{E_{\mathrm{a}}}{kT}\right), \tag{2.32}$$

where α_1 and α_2 are technology-dependent voltage scaling parameters, C_{R}, n_{R} and n_{P} are time exponents for the recoverable and permanent component respectively, E_{a} is the activation energy, and k is the Boltzmann constant. Note that the recoverable part in Eq. (2.32) is dependent on the duty cycle of the applied stress signal, which is modeled within C_{R} and n_{R}. BTI degradation has not shown to be dependent on frequency for measurements up to 3 GHz [97] [91].

2.3.4 Time Dependent Dielectric Breakdown (TDDB)

The insulator at the gate channel interface needs to separate charges accumulated in the channel and the gate plate. Each dielectric can sustain a certain maximum electric field. If the dielectric is exposed to higher fields, a dielectric breakdown occurs. This is referred to as a hard break down. At lower fields a time dependent wearout occurs, which eventually results in a time dependent dielectric breakdown (TDDB).
TDDB is a statistical process. Before the event of a breakdown a stress-induced leakage current (SILC) is observed through the gate [122] [18]. As the SILC increases, also the trap density increases until a critical limit is reach, which exhibits the breakdown. This behavior can be modeled through a Weibull probability distribution [122] [124]:

$$F(t_{\mathrm{bd}}) = 1 - \exp\left[-\left(\frac{t_{\mathrm{bd}}}{\alpha_{\mathrm{bd}}}\right)^{\beta_{\mathrm{bd}}} \right], \tag{2.33}$$

where $F(t_{\mathrm{bd}})$ is the cumulative density function for time-to-breakdown event, α_{bd} and β_{bd} are technology-dependent parameters.

2.3.5 Electromigration (EM)

The effect of material migration is caused by various transport processes in solids. The most common are chemical diffusion due to concentration gradients, temperature induced material migration, migration caused by mechanical stress, and material migration due to high electric fields [114]. The last one is often referred to as electromigration

and is supposed to be the most severe material process for integrated circuits. Electromigration usually occurs at grain boundaries of crystalline structures, where the crystal is inhomogeneously structured. At high currents, conducting electrons interact with atoms of the metal line structure and transport part of the line material in the direction of current flow, along the grain boundary [77]. A common measure for the degradation of a wire is the mean time to failure (MTTF), which in case of electromigration is modeled by Black's law [13]:

$$\mathrm{MTTF} = \frac{A}{J^n} \exp\left(\frac{E_\mathrm{a}}{kT}\right), \tag{2.34}$$

where A is a constant dependent on the cross section area of the conducting wire, J is the current density, E_a is the activation energy, k is the Boltzmann constant, T is the temperature, and n a scaling factor. Electromigration is extremely layout dependent. The relation in Eq. (2.34) gives several advices on how to reduce this degradation behavior, e.g. by lowering the temperature or allowing large cross sectional areas. Conducting wires should avoid sharp bends, as the current density is not uniform in cornered wires. Recent approaches have shown, that the effect of wide wires for avoiding EM is limited, which leads to innovative approaches in the layout, such as slotted wires [77] or even octahedron metal tracks [69].

2.4 Aging and Reliability Simulation

All of the above discussed effects influence the behavior of a circuit. The actual impact on circuit characteristics is computed using compact models, as it will be discussed in section 2.5. However, these models are most likely not solvable through standard solvers, such as SPICE, since the aging model equations are quite different from existing transistor compact models. Therefore, special simulators evolved within the last decades, which purely serve the need for simulating degradation effects in integrated circuits.
The demand for reliability simulators has already started in the beginning of the 1990's. Many authors have proposed simulators, e.g. RELY [101], BERT [59] or HOTRON [9]. All of these simulators mainly covered HCD issues and were not commercially available. With the shrinking feature sizes of modern technology nodes and new degradation effects like BTI and TDDB, EDA manufacturers included reliability simulators into their portfolio. Three major reliability simulators have evolved, which are briefly discussed in the following sections.

2.4.1 Mentor Graphics ELDO

This simulator is based on the work from [86]. In contrast to most other reliability simulation schemes, this simulator is build directly into the Mentor SPICE engine called ELDO. The ELDO simulator provides information about circuit performance shifts due

to gradual transistor aging effects, specifically BTI and HCD. Sudden effects, such as TDDB are not supported. ELDO does not use predefined models, therefore the user (or the foundry) has to implement the according models using the user-defined reliability model (UDRM) interface.
In order to simulate the degradation of a circuit, at least a transient (`.tran`) or steady state (`.sst`) analysis has to be performed. The final operating time t_{age} is split into a natural number of smaller time steps t_i, which may be distributed in either linear or logarithmic form. During simulation the voltage at every node is used to compute a stress vector S_i for each transistor. The actual shift of transistor parameters, e.g. ΔV_{th}, is computed based on S_i and extrapolated for the next simulation time step t_i. Both models, the computation of the stress vector S_i and the relation to the actual shift in transistor parameters, have to be provided within the UDRM interface. The simulation is repeated until the all time steps t_i have been simulated and the final operating time t_{age} has approached.

2.4.2 Synopsys MOSRA

The MOS reliability analysis (MOSRA) is included in HSPICE and CustomSim [115]. MOSRA is capable of simulating degradation caused by BTI and HCD only, but includes the possibility to account for a recovery effect in BTI. The simulation flow supports the use of user-defined models, but also offers predefined models, which can be adjusted through various reliability parameters. The use of predefined models reduces the effort for the foundry to support aging models, as only the specific reliability parameters have to be provided instead of a whole aging model.
A reliability simulation is performed within two phases. In the pre-stress phase, MOSRA calculates the electrical stress of selected transistors, which is based on the MOSRA aging models. The total stress for the operating time t_{age} is computed by extrapolation of the simulation result. The resulting degraded circuit is evaluated in the post-stress phase.

2.4.3 Cadence RelXpert

The Cadence reliability expert (RelXpert) is included within the Virtuoso IC development suite. RelXpert supports Ultrasim and the analog design environment (ADE) and offers the simulation of BTI and HCD degradation modes by default. Other degradation modes as well as user-defined models for existing modes can be added via the user-defined reliability interface (URI). The BTI model does not support any recovery effect and is therefore comparable to implementation done in ELDO. The RelXpert simulator is basically a commercial implementation of the Berkley reliability tools (BERT) [59].

Throughout this work degradation simulation is performed using the RelXpert reliability simulator, which therefore is discussed in detail in the following sections. In section 2.4.4 the reliability simulation flow is covered, which is based on the calculation of a transistor specific *AGE* parameter. The calculation of this parameter is described in 2.5.3. RelXpert supports two different methods to model degradation modes for a specific technology: aged model files and the proprietary AgeMOS model.

Aged Model Files

Transistor model parameters (e.g. BSIM) are extracted for the fresh state and also at a number of different stress intervals. These fresh and aged model parameters form a set of aged model files. Each file represents the behavior of a transistor for a dedicated stress time, which in case of RelXpert, is expressed as certain *AGE*. The calculation of this term is described in section 2.5.3. During simulation, the *AGE* of each transistor in a circuit is calculated. The final aged model parameters are computed through regression and interpolation of the aged model files.

AgeMOS

In this analytical approach, each degradation effect is separately modeled. This model must be supplied by the foundry, as the model generation requires special tools, which are usually not available to the user. However, the user may also describe their own model using URI input mode. In case of a default AgeMOS simulation, the shift in transistor parameter ΔP is described by

$$\Delta P = f(P_{\text{fresh}}, AGE, \vec{a}), \tag{2.35}$$

where P_{fresh} is the parameter at its fresh state, *AGE* is the transistor age as described in 2.5.3 and $\vec{a}$ is a vector containing AgeMOS parameters. The model itself is not public and only the computation of the circuit lifetime parameter *AGE* is publicly available.

2.4.4 RelXpert Simulation Flow

The simulation flow of RelXpert is depicted in Fig. 2.6. In order to compute the degradation of a circuit, RelXpert needs two input data: the fresh netlist of a circuit and the degradation models, which in this case is provided through the internal AgeMOS models. A first preprocessor prepares the fresh netlist, parses the circuit and collects the information on the used degradation model. The circuit is simulated using a transient SPICE simulation. In this phase, the voltage and current signal on each node is saved and passed to post-processor This post-processor calculates the degradation (*AGE*) for each individual transistor and outputs a degradation table. This table is used by the second preprocessor to create an aged netlist. This aged netlist contains an individual degraded model set for each transistor. The preprocessor-2 additionally runs a final

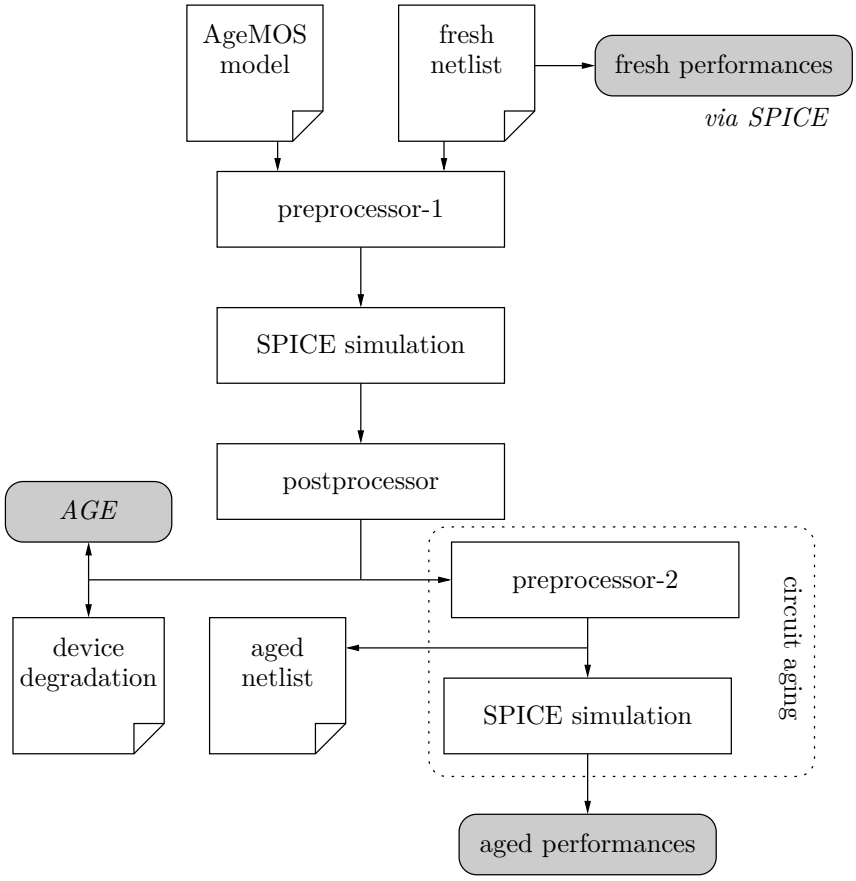


Figure 2.6: Simulation flow of RelXpert for circuit degradation evaluation using the proprietary AgeMOS model [93].

SPICE simulation to simulate the degraded circuit performance after HCD and BTI stress.

2.5 Compact Modeling of Aging Mechanisms

In physical simulators, such as the *Technology Computer-Aided Design* (TCAD) from Synopsys [68] or the *COMSOL Multiphysics* toolbox [23], atomic simulation models can be used to identify a transistor's behavior concerning BTI, HCD or TDDB. Although the results of these simulations are accurate in terms of physical modeling, the computational efficiency is very low and also time consuming. Using these accurate models only allows a simulation time of several nano- or microseconds. The usual simulation time of interest in regards of reliability covers the whole life span of a device, which is likely to be in the range of several years. Therefore, in the case of a compact modeling approach, which is the basis for all SPICE related simulators, different model sets and simulation schemes have to used. The following described common compact modeling approaches are used in reliability simulators, specifically in RelXpert. Compact models for NBTI

and HCD are reviewed and also their implementation in RelXpert through a circuit lifetime model.

2.5.1 Reaction-Diffusion Model for NBTI

The Reaction-Diffusion (RD) model describes NBTI as a thermally activated reaction of holes with Si–H bonds at the interface between substrate and gate-dielectric. A silicon bulk material is prone to unbound bonds at the surface, as it is shown in Fig. 2.7a. These traps catch free hydrogen and oxygen ions. Hydrogen is intentionally added through the CMOS production, in order to fill up open bonds. As hydrogen diffuses away from the interface, dangling bonds are generated at the interface, resulting in interface states at the substrate to dielectric barrier and also in an accumulation of positively charged hydrogen within the gate, which is shown in Fig. 2.7c. The mechanism describing the RD model has first been described in [62]. In its initial version the model was used to explain NBTI gate oxide field and temperature dependency. The model has been subsequently updated to include NBTI saturation for long stress times and also the independence on the input signal frequency [5] [4].
Within the RD model, NBTI is described as a hole-assisted breaking of Si–H (and also Si–O) bonds at the channel gate interface. This reaction is dependent on the electrical field and temperature. The rate at which traps at the interface are generated (*reaction*) is described by

$$\frac{\mathrm{d}N_{\mathrm{it}}}{\mathrm{d}t} = k_{\mathrm{F}}(N_0 - N_{\mathrm{it}}) - k_{\mathrm{R}} N_{\mathrm{H}}(0) N_{\mathrm{it}}, \tag{2.36}$$

where N_{it} is the number of interface traps, which in this case are caused by broken Si–H bonds, N_0 is the number of initially existing Si–H bonds, k_{F} represents the oxide-field dependent forward dissociation constant, k_{R} is the annealing rate constant and $N_{\mathrm{H}}(0)$ is the concentration of hydrogen at the interface for $x = 0$. The first term in Eq. (2.36) shows that the rate of N_{it} is dependent on the number of broken bonds (at rate k_{F}), whereas the second terms describes the annealing of broken bonds with already released hydrogen atoms (at rate k_{R}). Released hydrogen atoms may not only anneal to broken bonds, but also diffuse into the gate dielectric (*diffusion*), which is described by

$$\frac{\mathrm{d}N_{\mathrm{H}}}{\mathrm{d}t} = D_{\mathrm{H}} \frac{\mathrm{d}^2 N_{\mathrm{H}}}{\mathrm{d}x^2}, \tag{2.37}$$

where N_{H} is the total hydrogen concentration in the oxide and D_{H} is the diffusion constant. A closed form solution is obtained using the following assumptions:

1. The initial trap generation rate is small:

$$\frac{\mathrm{d}N_{\mathrm{it}}}{\mathrm{d}t} \approx 0. \tag{2.38}$$

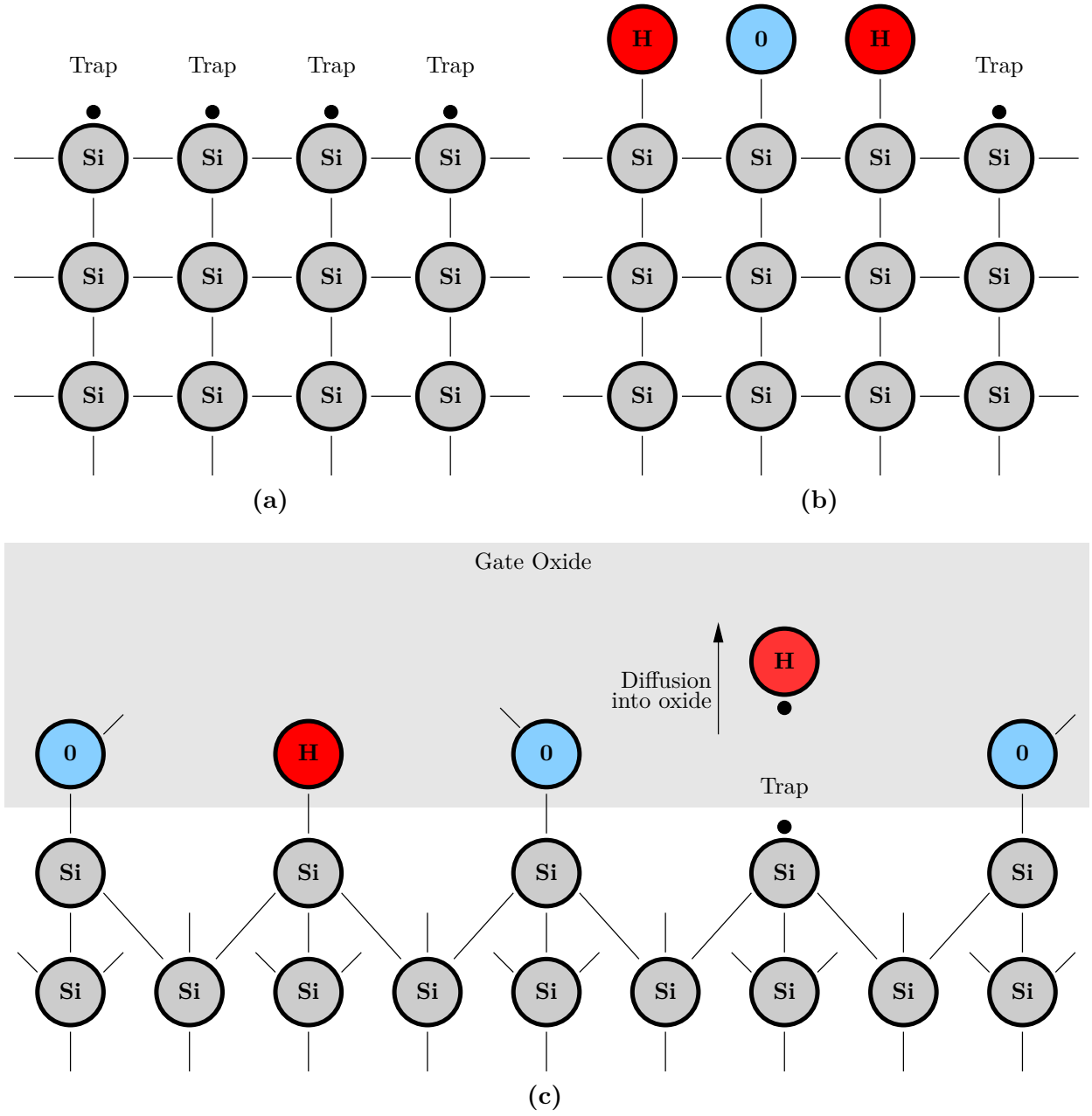


Figure 2.7: Silicon surface and dangling bonds. (a) Simple silicon structure as bulk material The surface atoms have unsatisfied bonds, which form traps. (b) Traps catch other carriers, such as hydrogen (H) or oxygen (O). (c) Silicon 111 structure at the bulk oxide interface. Difference in lattice geometry leaves spots, which are filled with H during process. Diffusion of H into oxide is part of BTI.

2. Initially, the number of generated traps N_{it} is small compared to the number of available bonds to brake N_0:

$$N_{\mathrm{it}} \ll N_0 \approx 5 \cdot 10^{12}\,\mathrm{cm}^{-2}. \tag{2.39}$$

Therefore, Eq. (2.36) may be written as

$$N_{\mathrm{H}}(0) N_{\mathrm{it}} \approx \frac{k_{\mathrm{F}}}{k_{\mathrm{R}}} N_0. \tag{2.40}$$

After the initial startup, mainly the diffusion of hydrogen contributes to the trap generation process. The depth of diffusion (the distance from the channel interface into the gate, which has been covered by H atoms) is given by

$$X_{\mathrm{DF}}(t) = \sqrt{D_{\mathrm{H}} \cdot t}. \tag{2.41}$$

Since the number of diffusing H atoms increases with time, also the diffusion depth X_{DF} is dependent on time. Additionally, the number of hydrogen atoms within the diffusion zone encountered between $0 < x < X_{\mathrm{DF}}$ is equal to the number of generated interface traps N_{it}, such that

$$N_{\mathrm{it}} = \int_0^{\sqrt{D_{\mathrm{H}} \cdot t}} N_{\mathrm{H}}(x,t)\,\mathrm{d}x = \frac{1}{2} N_{\mathrm{H}}(0) \sqrt{D_{\mathrm{H}} \cdot t}. \tag{2.42}$$

The solution from Eq. (2.42) is inserted into Eq. (2.36) and finally yields

$$N_{\mathrm{it}} = \left(\frac{k_{\mathrm{F}}}{k_{\mathrm{R}}} \frac{N_0}{2} \right)^{1/2} (D_{\mathrm{H}} \cdot t)^{1/4}. \tag{2.43}$$

The expression in Eq. (2.43) yields the number of generated interface traps N_{it} for a given stress time t. Every diffused hydrogen causes a shift in threshold voltage, which in sum is equal to voltage on the oxide capacitance C_{OX} induced by the total charge of the hydrogen:

$$\Delta V_{\mathrm{th}} = \frac{q N_{\mathrm{it}}}{C_{\mathrm{OX}}}. \tag{2.44}$$

2.5.2 Lucky Electron Model for HCD

The impact of HCD can be modeled in dependence on the substrate current I_{sub} and the gate current I_{g} for hot carrier degradation in NMOS and PMOS devices respectively, which is both supported by the RelXpert simulator. As HCD is most severe in NMOS devices, only modeling of the substrate current is further investigated. A correlation of I_{sub} and HCD exists, as both are driven by the maximum channel electrical field E_{m} [60]. This field occurs at the drain side of the channel.

The substrate current is modeled using a well known approximation as shown in [30]

$$I_{\mathrm{sub}} = \alpha_i \cdot E_{\mathrm{m}} l_{\mathrm{c}} I_{\mathrm{D}} \exp\left(-\frac{\phi_{\mathrm{i}}}{q\lambda E_{\mathrm{m}}}\right) = C_1 \cdot I_{\mathrm{D}} \exp\left(-\frac{\phi_{\mathrm{i}}}{q\lambda E_{\mathrm{m}}}\right), \tag{2.45}$$

where α_i is a process related constants, l_{c} is the characteristic length of the saturation region, λ is the hot-electron mean free path and ϕ_{i} is the minimum energy required for impact ionization. The distance $\phi_{\mathrm{i}}/qE_{\mathrm{m}}$ has to be covered by en electron within E_{m}, in order to gain the energy ϕ_{i}. Therefore, $\exp(-\phi_{\mathrm{i}}/q\lambda E_{\mathrm{m}})$ describes the probability of an electron covering a sufficient distance (without collision) gaining an energy equal or greater than ϕ_{i}. If I_{D} is considered to be the rate of cold electrons within the transistor channel, then $I_{\mathrm{D}} \cdot \exp(-\phi_{\mathrm{i}}/q\lambda E_{\mathrm{m}})$ describes the rate of hot electrons possessing energies above ϕ_{i}. This concept is known as the *lucky electron* principle is first introduced by Shockley [104] to describe bulk phenomena.
HCD can be expressed by calculating the generated interface traps ΔN_{it}. The expression for ΔN_{it} is gained similarly to the lucky electron principle [60]. For a given transient simulation time t the number of generated interface traps results to

$$\Delta N_{\mathrm{it}} = C_2 \left[\frac{I_{\mathrm{D}}}{W} \cdot \exp\left(\frac{\phi_{\mathrm{it}}}{q\lambda E_{\mathrm{m}}}\right)\right]^n t^n, \tag{2.46}$$

where C_2 is a process related constant and ϕ_{it} is the critical energy for an electron to create an interface trap. The exponent n is experimentally determined. In order to compute ΔN_{it} degradation with easily measurable characteristics, such as the substrate current I_{sub}, (2.46) can be rewritten. Expression (2.45) is used to calculated the relation of I_{sub} to I_{D}

$$\frac{I_{\mathrm{sub}}}{I_{\mathrm{D}}} = C_1 \cdot \exp\left(-\frac{\phi_{\mathrm{i}}}{q\lambda E_{\mathrm{m}}}\right). \tag{2.47}$$

The relation from (2.47) is inserted into (2.46) with $m = \phi_{\mathrm{it}}/\phi_{\mathrm{i}}$ and $H = C_1^m \cdot C_2^{n-1}$, which yields

$$\Delta N_{\mathrm{it}}(I_{\mathrm{sub}}) = \left[\frac{I_{\mathrm{DS}}}{WH} \cdot \left(\frac{I_{\mathrm{sub}}}{I_{\mathrm{DS}}}\right)^m\right]^n t^n. \tag{2.48}$$

2.5.3 Circuit Lifetime Model

The circuit lifetime model used in RelXpert is based on the work by Hu on the BERT simulator [59]. The idea of the lifetime model is to use a single generic aging parameter AGE, which represents the aging of a transistor. This generic AGE is extrapolated to the final operating time t_{age}, which results in an intermediate degradation metric ΔD. The actual shift in transistor parameters, e.g. ΔV_{th}, is generated as a function of ΔD by individual models. Each degradation mode uses an individual model to compute AGE and ΔD.

The degradation ΔD is defined as the generation of interface traps, which within RelXpert is generically modeled as

$$\Delta D = f_{\mathrm{N}}(A_{\mathrm{rate}} \cdot t) = (A_{\mathrm{rate}} \cdot t)^n = AGE^n, \tag{2.49}$$

where A_{rate} is the aging rate, at which device degradation occurs and $AGE = A_{\mathrm{rate}} t$ is an abstract circuit age, which is used by $f_{\mathrm{N}}(x) = x^n$ to compute the final degradation ΔD. The value of ΔD has a physical equivalent, which is the number of generated interface traps N_{it}

$$\Delta D = \Delta N_{\mathrm{it}}. \tag{2.50}$$

Comparing Eq. (2.49) with Eq. (2.48) for HCD and with Eq. (2.43) for NBTI, the aging rate for both effects results to

$$\text{HCD: } A_{\mathrm{rate}} = \frac{I_{\mathrm{DS}}}{WH} \cdot \left(\frac{I_{\mathrm{sub}}}{I_{\mathrm{DS}}}\right)^m. \tag{2.51}$$

$$\text{NBTI: } A_{\mathrm{rate}} = \sqrt[4]{\left(\frac{k_{\mathrm{F}}}{k_{\mathrm{R}}}\frac{N_0}{2}\right)^{1/2} \cdot D_{\mathrm{H}}}. \tag{2.52}$$

Although it is possible to derive aging rates for both effects, the exact expression in case of NBTI has not been published, therefore the value given in Eq. (2.52) is not necessarily the one used within RelXpert. The value of A_{rate} is computed during a SPICE simulation. In case of a DC simulation, this value is computed once and then used to extrapolate for the wanted stress time $t = t_{\mathrm{age}}$. Varying signals, e.g. v_{gs}, ultimately cause a variant substrate current i_{sub}. Therefore, the aging rate A_{rate} is also changing over the simulation time and has to be evaluated for each time step in a transient simulation. The final degradation ΔD for varying stress can then be obtained using

$$\Delta D = \left[\left(\frac{1}{t_{\mathrm{sim}}}\right) \cdot \int_{t_i} A_{\mathrm{rate}}(t_i)\,\mathrm{d}t\right]^n. \tag{2.53}$$

RelXpert calculates the actual shift of each transistor parameter based on the degradation ΔD, e.g. the shift in threshold voltage is given as

$$\Delta V_{\mathrm{th}} = f_{\mathrm{deg}}(\Delta D), \tag{2.54}$$

where f_{deg} is a mapping function, which is determined experimentally.

2.6 Summary

This chapter has discussed basic reliability issues for integrated circuits. A mathematical nomenclature for defining reliability based on a set of circuit samples was introduced. This nomenclature will be used throughout this work in the upcoming chapters. The

reliability and failure rate function led to the famous Weibull distribution for analyzing reliability as a failure rate over time. Different variability and specifically aging mechanisms were described and how these affect transistor parameters. Physical and compact modeling approaches were shown for NBTI and HCD and also how aging simulators exploit these models for reliability simulations.

3 Degradation-Oriented Analog Circuit Design

Designing circuits with a high lifetime requires knowledge about the mechanisms, which influence the behavior of a circuit, and also needs design countermeasures, in order to harden the device against these mechanisms. A possible countermeasure is a structural change in the circuit topology, which involves selecting a specific structure of less susceptibility to aging. A structural change within a circuit is not allowed to affect the main characteristics of a design, such as gain or bandwidth of the system, therefore these parts can only be substituted by functional equivalents. Once the system is hardened, the gain in reliability needs to be determined. Estimation of reliability figures under the consideration of aging is not a straight forward flow within current standard EDA tools. Since the estimation of the time-dependent yield usually requires numerous SPICE and aging simulation runs, more efficient simulation flows need to be considered. An innovative modified Monte Carlo analysis can be used for this estimate, which reduces the number of simulation runs while still providing accurate results.
In this chapter a folded cascode operational amplifier is hardened with respect to its response to aging. This is achieved on a structural basis, where specific bias circuit and current mirror configurations are investigated in regards of NBTI and HCD degradation mechanisms. The hardened amplifier is evaluated using an adapted Monte Carlo analysis, which determines an estimate for the expected reliability.

3.1 Single Transistor Degradation Behavior

As discussed in chapter 2, aging mechanisms such as NBTI and HCD cause a degradation in transistor parameters, which is mainly seen as a degradation in the threshold voltage V_{th}. NBTI and HCD both share similar physical principles (generation of interface states), but are also induced through unique effects (hydrogen diffusion for NBTI or ionization for HCD). Therefore, in order to excite each aging effect, different setups have to be chosen.
The shift in threshold voltage ΔV_{th} given in percentage for an PMOS and NMOS type transistor based on a RelXpert simulation is depicted in Fig. 3.1.

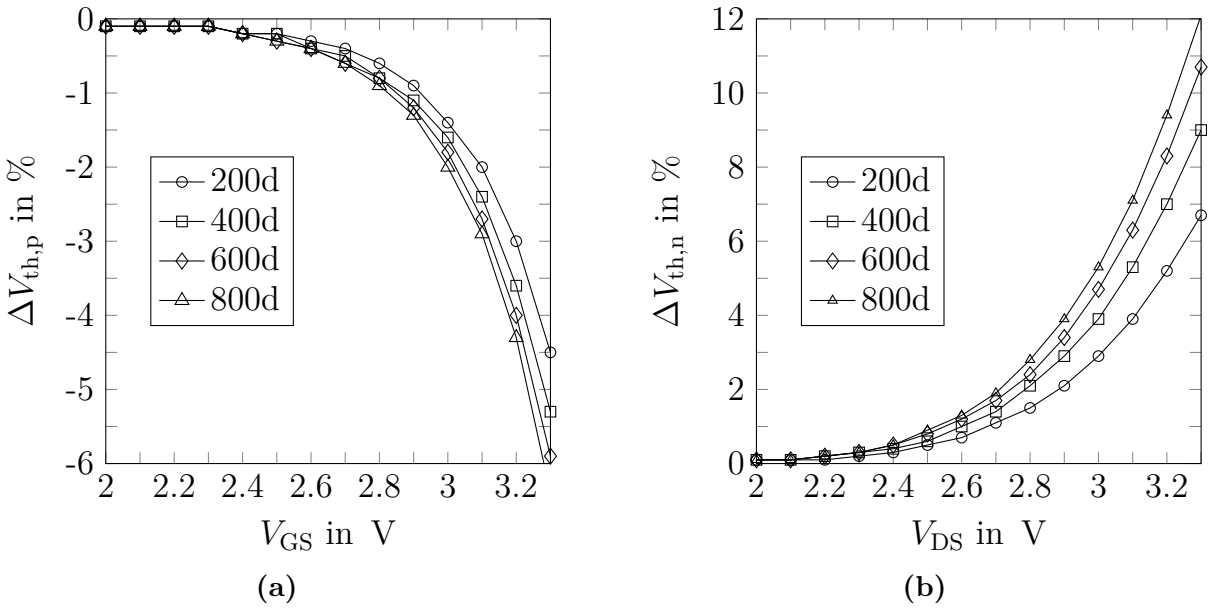


Figure 3.1: Degradation of threshold voltage in a 3.3 V 150 nm technology node for different stress times from 200 to 800 days in a step size of 200 days [50]. (a) Shift $\Delta V_{\mathrm{th,p}}$ in PMOS transistor ($W = 12.5\,\mu\mathrm{m}$, $L = 1\,\mu\mathrm{m}$) due to NBTI at a constant $V_{\mathrm{DS}} = 1.65\,\mathrm{V}$. (b) Shift $\Delta V_{\mathrm{th,n}}$ in NMOS transistor ($W = 5\,\mu\mathrm{m}$, $L = 1\,\mu\mathrm{m}$) due to HCD at a constant $V_{\mathrm{GS}} = 1.0\,\mathrm{V}$.

The PMOS transistor is tressed through a variant gate-source voltage $2 \leq V_{\mathrm{GS}} \leq 3.3\,\mathrm{V}$ and the drain-source voltage V_{DS} is held constant at 1.65 V, which corresponds to $1/2 \cdot V_{\mathrm{dd}}$, whereas in case of the NMOS transistor a variant drain-source voltage is swept from $2 \leq V_{\mathrm{DS}} \leq 3.3\,\mathrm{V}$ and the gate-source voltage is kept constant at $V_{\mathrm{GS}} = 1\,\mathrm{V}$. Note that NBTI manifests itself as a change in the PMOS $V_{\mathrm{th,p}}$ and is dependent on V_{GS}, whereas HCD is responsible for shift in the NMOS $V_{\mathrm{th,n}}$ and dependent on V_{DS}.
The shift ΔV_{th} can reach a value of over 6 % for the n-type transistor and 10 % for the p-type equivalent. Considering fresh values of $V_{\mathrm{th,p}} = -0.622\,\mathrm{V}$ and $V_{\mathrm{th,n}} = 0.561\,\mathrm{V}$, this results in absolute shifts in the range of several mV. There are two main observations, which can be derived from the simulation results in Fig. 3.1:

1. The shift in threshold voltage ΔV_{th} is exponentially dependent on the stress voltage (V_{DS} for NMOS and V_{GS} for PMOS).

2. The stress time worsens the impact of the degradation modes (HCD for NMOS and NBTI for PMOS).

The design of integrated circuits relies on the exact knowledge of component parameters, which specifically includes the threshold voltage V_{th}. Moreover, many other parameters, such as g_{m} or g_{ds}, are dependent or derived from V_{th} and are therefore influenced as well. It becomes clear, that such deviations cannot be neglected and may lead to severe influences on whole circuits.

In order to comprehend the impact of the discussed degradation modes on complex circuits, a detailed analysis is carried out on how specific circuit topologies suffer from aging. The following sections investigate the aging behavior of biasing circuit, mirroring structures, composed current mirrors and also how the results help to improve the design of complete operational amplifiers.

3.2 Current Mirrors

Current mirrors structures are widely used throughout various circuit types in analog circuit design and are responsible for important system performances. A current mirror is designed to accept an input current I_{in} and to provide an output current I_{out}, where

$$I_{\text{out}} = \kappa \cdot I_{\text{in}} \tag{3.1}$$

with κ being a constant ratio of both currents. In its most simplest form a current mirror consists of two transistors of the same type. The input transistor in configured as a gate-drain-connected diode ($V_{\text{D}} \equiv V_{\text{G}}$) and converts I_{in} into a bias voltage $V_{\text{bias}} = V_{\text{GS}}$, which is fed to the output transistor where the output current is generated through

$$I_{\text{out}} = g_{\text{m,out}} \cdot V_{\text{bias}} = g_{\text{m,out}} \cdot V_{\text{GS}}. \tag{3.2}$$

In this case, the ratio κ is given by the ratio of the shape factors $\kappa = {}^{S_{\text{f,in}}}/_{S_{\text{f,out}}}$, with $S_{\text{f}} = {}^{W}/_{L}$ of the corresponding transistor. This basic functionality of a current mirror can be split into two sub-circuits: a biasing and a mirroring circuit.

3.3 Bias Circuits

The function of a bias circuit is to convert a given bias current I_{bias} into a bias voltage V_{bias}, which is used by a mirror circuit. Ideally, the associated bias current should also be generated within the circuit. The generated bias voltage should be independent of the supply voltage and should also be stable against process variability. There exist numerous biasing circuit variants in literature, each with specific advantages and disadvantages [39] [47] [28]. The following analysis investigates two different biasing circuits: a MOS only reference (MOR) in section 3.3.1 and a beta matching reference (BMR) in section 3.3.2.

3.3.1 Mos Only Reference (MOR)

A schematic representation of a MOR is shown in Fig. 3.2a. The circuit can generate a voltage $V_{\text{bias,n}}$ or $V_{\text{bias,p}}$, which is used to bias an NMOS or PMOS to drive the current I_{bias} or any κ-multiple of it. The MOR has to be specifically sized in order to either

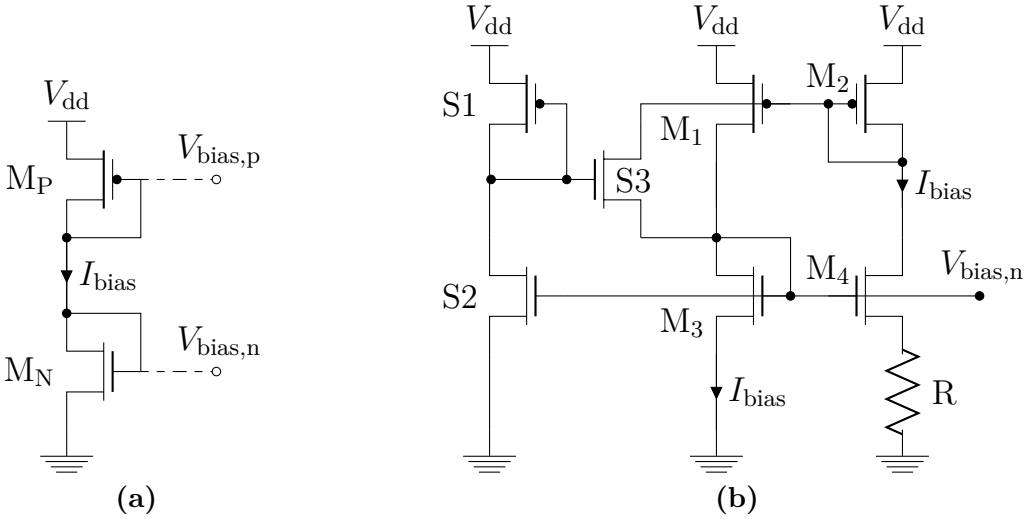


Figure 3.2: Schematic view of different biasing circuits. (a) MOS only reference. (b) Beta matching reference.

generate $V_{\mathrm{bias,n}}$ or $V_{\mathrm{bias,p}}$. Assuming the MOR should generate $V_{\mathrm{bias,n}}$, then the transistor $\mathrm{M_P}$ in Fig. 3.2a needs to be sized accordingly, such that the voltage drop of $\mathrm{M_P}$ fits $V_{\mathrm{bias,n}}$. A sizing rule is obtained as follows. Since both transistors are configured as a drain-connected diode, Kirchoff's law yields

$$V_{\mathrm{dd}} = V_{\mathrm{GS,P}} + V_{\mathrm{GS,N}}. \tag{3.3}$$

The gate-source voltage V_{GS} can be obtained using an quadratic approach for I_{DS}

$$I_{\mathrm{DS}} = \frac{\mu_0 C_{\mathrm{OX}}}{2} \frac{W}{L} (V_{\mathrm{GS}} - V_{\mathrm{th}})^2 \tag{3.4}$$

$$\Rightarrow V_{\mathrm{GS}} = \sqrt{\frac{2 I_{\mathrm{bias}}}{\frac{\mu_0 C_{\mathrm{OX}}}{2} \left.\frac{W}{L}\right|_{\mathrm{M_P}}}}. \tag{3.5}$$

By replacing the according V_{GS} in Eq. (3.3) with the expression from Eq. (3.5) it follows that

$$V_{\mathrm{dd}} = \sqrt{\frac{2 \cdot I_{\mathrm{bias}}}{\frac{\mu_{0,\mathrm{P}} C_{\mathrm{OX}}}{2} \left.\frac{W}{L}\right|_{\mathrm{M_P}}}} + V_{\mathrm{th,p}} + \sqrt{\frac{2 \cdot I_{\mathrm{bias}}}{\frac{\mu_{0,\mathrm{N}} C_{\mathrm{OX}}}{2} \left.\frac{W}{L}\right|_{\mathrm{M_N}}}} + V_{\mathrm{th,n}}. \tag{3.6}$$

The expression from Eq. (3.6) is solved for the $\frac{W}{L}$ ratio of the generating device, which is $\left.\frac{W}{L}\right|_{\mathrm{M_N}}$ in case for $V_{\mathrm{bias,n}}$ and $\left.\frac{W}{L}\right|_{\mathrm{M_P}}$ in case for $V_{\mathrm{bias,p}}$. The bias current I_{bias} and $\frac{W}{L}$ ratio is first found for one device using Eq. (3.5) and then passed to Eq. (3.6) in order to solve for the remaining $\frac{W}{L}$ ratio. The MOR circuit is sized according to Table A.1 in the appendix. Typically, the transistor, which operates as a resistor, has a very narrow

aspect ratio, leading to long channel lengths. These long channel lengths are unlikely to be supported by the technology node and should be avoided, in order to allow a dense and compact layout of the circuit. These long transistors are split into a series of unit transistors, where all gate and all bulk terminals form a common terminal.

3.3.2 Beta Matching Reference (BMR)

Another approach for a biasing circuit is accomplished by using a beta multiplier reference, as depicted in Fig. 3.2b [10]. This reference circuit features two major advantages: The bias voltage is widely independent of the supply voltage and the BMR can also function as constant g_m bias source. The benefit of a constant g_m bias can be observed from the following derivation. Kirchhoff's voltage law is applied to the main circuit in Fig. 3.2b, such that

$$V_{\mathrm{GS},3} = V_{\mathrm{GS},4} + I_\mathrm{bias} \cdot R. \tag{3.7}$$

Equation (3.7) only holds if $V_{\mathrm{GS},3} > V_{\mathrm{GS},4}$, as otherwise the voltage $I_\mathrm{bias} \cdot R$ would be negative. Using a quadratic approach for drain current, V_GS can be expressed as

$$V_\mathrm{GS} = \sqrt{\frac{2I_\mathrm{D}}{\beta}} + V_\mathrm{th}. \tag{3.8}$$

Before inserting this expression into Eq. (3.7), we propose that β_3 and β_4 shall be matched by an arbitrary constant K

$$\beta_4 = K \cdot \beta_3, \tag{3.9}$$

with $K > 1$. Using the above expressions, the bias current I_bias can be evaluated to

$$I_\mathrm{bias} = \frac{2}{R^2 \beta_3} \left(1 - K^{-\frac{1}{2}}\right)^2. \tag{3.10}$$

An interesting result for $g_{\mathrm{m},3}$ is obtained for a scale factor of $K = 4$:

$$g_{\mathrm{m},3} = \sqrt{2\beta_3 \cdot I_\mathrm{bias}} = \frac{1}{R}. \tag{3.11}$$

The result from Eq. (3.11) shows, that the transconductance of M_3 is only dependent on the value of R. In this configuration, process shifts, which affect transistor parameters, do not alter the transconductance of M_3. Since process related shifts do not affect $g_{\mathrm{m},3}$, this configuration is also known as a constant-g_m bias circuit. However, depending on how the resistor R is processed (e.g. n-well or poly-silicon), the resistance may not

be fabricated precisely. In that case, the resistor needs to be tunable, e.g. through a switchable series resistance.
The specific characteristic of a constant process-independent g_m makes the circuit very useful when designing over process corners. However, this independence is not observed in combination with NBTI and HCD degradation modes. The constant g_m from Eq. (3.11) relies on the fact, that the threshold voltages of M_3 and M_4 are identical. Since M_3 and M_4 experience different stress ($V_{\mathrm{GS},3} \neq V_{\mathrm{GS},4}$, $V_{\mathrm{DS},3} \neq V_{\mathrm{DS},4}$), the according threshold voltages $V_{\mathrm{th},3}$ and $V_{\mathrm{th},4}$ are also likely to degrade differently. This introduces a mismatch, denoted as

$$\Delta V_\mathrm{th} = V_{\mathrm{th},3} - V_{\mathrm{th},4}. \tag{3.12}$$

Considering this mismatch and $K = 4$, Eq. (3.10) results to

$$I_\mathrm{bias} = \frac{\Delta V_\mathrm{th}}{R} \pm \frac{\sqrt{8\beta_3 R \Delta V_\mathrm{th} + 1} + 1}{4\beta_3 R^2}, \tag{3.13}$$

which leads to a transconductance of

$$g_{\mathrm{m},3} = \frac{\sqrt{4\beta_3 R \Delta V_\mathrm{th} + \sqrt{8\beta_3 R \Delta V_\mathrm{th} + 1} + 1}}{\sqrt{2} R}. \tag{3.14}$$

As it can be observed in Eq. (3.14), in the case of a degradation induced mismatch, the transconductance is still dependent on process shifts as well as on the mismatch ΔV_th itself. If no degradation occurs ($\Delta V_\mathrm{th} = 0\,\mathrm{V}$), the expression from Eq. (3.14) leads to the result from Eq. (3.11). The worsened response due to degradation is further investigated in section 3.5.1.

3.4 Mirror Circuits

A mirror circuit uses the voltage generated by a bias circuit, e.g. the MOS only reference or the beta matching reference, and transforms it into a constant current. Ideally the output current should be independent of the load. The simplest form of a mirror circuit is constructed using a single NMOS or PMOS transistor, depending on the generated bias voltage. This simple mirroring is very compact and efficient in layout size, but only offers a limited output resistance. For single MOS mirror circuits, the output resistance is highly dependent on the channel length L, which results in rather long transistors for highly resistive outputs. Therefore, more complex mirroring strategies have been developed to overcome this issue.
Two main types of mirror circuits are distinguished: sinks and sources. Given a single rail configuration (single supply voltage), a current sink drives a current from a given node to ground, whereas a current source drives a current from a given node to V_dd. The study in section 3.5.1 investigates three different mirror structures, each in either

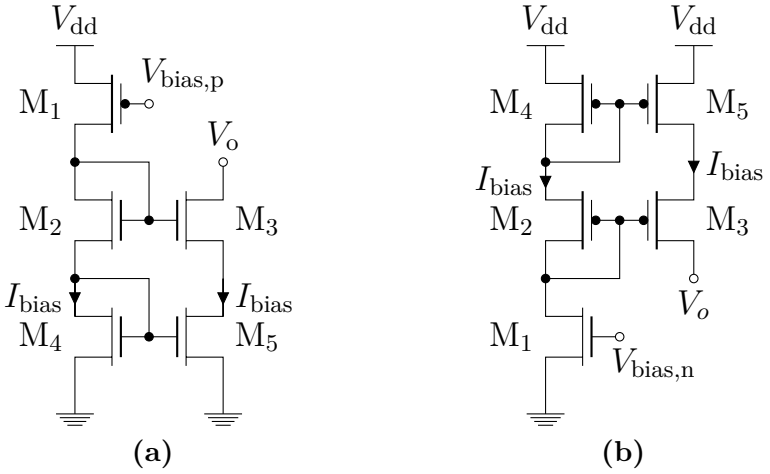


Figure 3.3: Schematic view of a cascode mirror structure. (a) Cascode structure as sink. (b) Cascode structure as source.

sink or source configuration: single transistors, cascode mirrors and wide swing mirrors. The cascode and wide swing configuration is discussed in the following two sections.

3.4.1 Cascode Mirror Structure

Both, the cascode and the wide swing structures, are designed to deliver a higher output resistance than the one provided by a single transistor stage. The cascode mirror circuit is depicted in Fig. 3.3a and Fig. 3.3b in sink and source configuration respectively. The output current is passed through the transistors M_3 and M_5. There exists a minimum for the node voltage V_o, which has to be exceeded in order to operate M_3 and M_5 in saturation. Transistor M_5 is biased by M_4, therefore the drain voltage of M_4 is mirrored to the drain at M_5, which results to

$$V_{D,5} = V_{DS,sat,4} + V_{th,4} = V_{DS,sat,5} + V_{th,5}, \tag{3.15}$$

where $V_{DS,sat,4}$ is the saturation voltage of M_4. Therefore, in order to keep M_3 in saturation, the output voltage V_o should fulfill

$$V_o \geq V_{DS,sat,5} + V_{th,4} + V_{DS,sat,5}. \tag{3.16}$$

If all dimensions are chosen identically, with $V_{DS,sat,5} = V_{DS,sat,3} = V_{DS,sat}$ and $V_{th,4} = V_{th,n}$, then the minimum voltage results to

$$V_{o,min} = 2 \cdot V_{DS,sat} + V_{th}. \tag{3.17}$$

This minimum can be reduced through the use of a wide swing configuration.

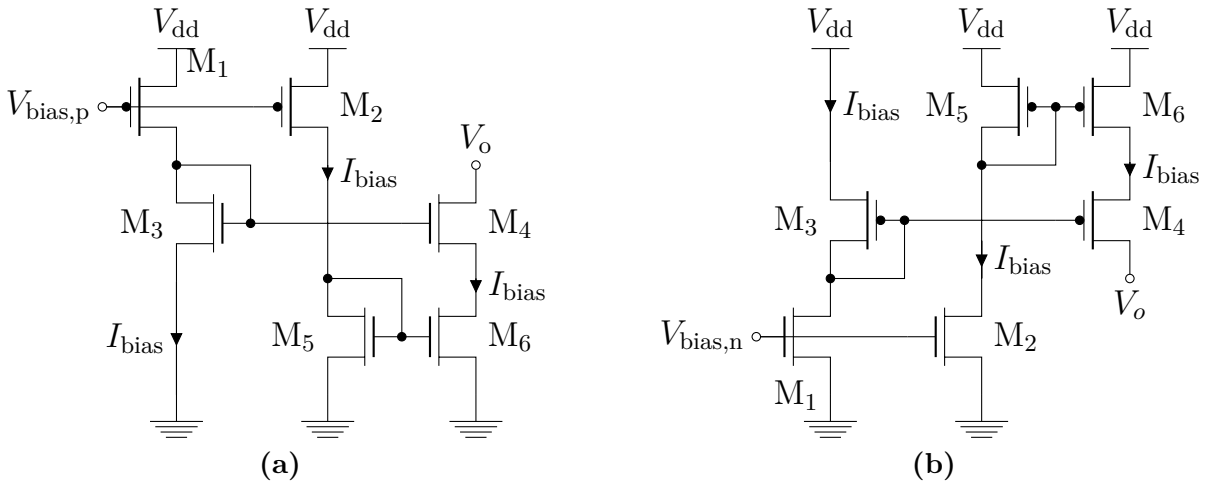


Figure 3.4: Schematic view of a wide swing mirror structure. (a) Wide swing structure as sink. (b) Wide swing structure as source.

3.4.2 Wide Swing Mirror Structure

The schematic view of a wide swing mirror structure is shown in Fig. 3.4 in sink and source configuration. This structure operates similar to a cascode mirror structure. Comparing Fig. 3.4a and 3.3a shows, that the cascode output transistors M_3 and M_5 in Fig. 3.4a are equivalent to M_4 and M_6 in Fig. 3.3a. The minimum output voltage for V_o compared to a cascode mirror is reduced by decoupling the gates of M_4 and M_6, as shown in sink configuration in Fig. 3.4a. Here, the source of M_4 is not bound to $V_{D,5}$ and therefore M_6 is saturated for $V_{D,6} > V_{DS,sat,6}$. The same constraint holds for M_4 and the minimum output voltage results to

$$V_{o,min} = 2 \cdot V_{DS,sat}. \tag{3.18}$$

Note that this is only true, if M_3 generates a suitable bias for M_4, which is achieved for

$$\left.\frac{W}{L}\right|_4 = \left.\frac{W}{L}\right|_3 \cdot 4. \tag{3.19}$$

3.5 Structural Degradation-Oriented Circuit Design

Within an analog integrated circuit, the discussed bias and mirror circuits from the previous section are used in various ways. In some cases, it is possible to interchange these circuits, without losing characteristics of the overlaying system, such as an amplifier. Nonetheless, equivalent current mirrors may provide identical functionality, but due to a difference in topology degrade differently. Within the following analysis, the degradation behavior of multiple combinations of bias and mirror circuits is investigated.

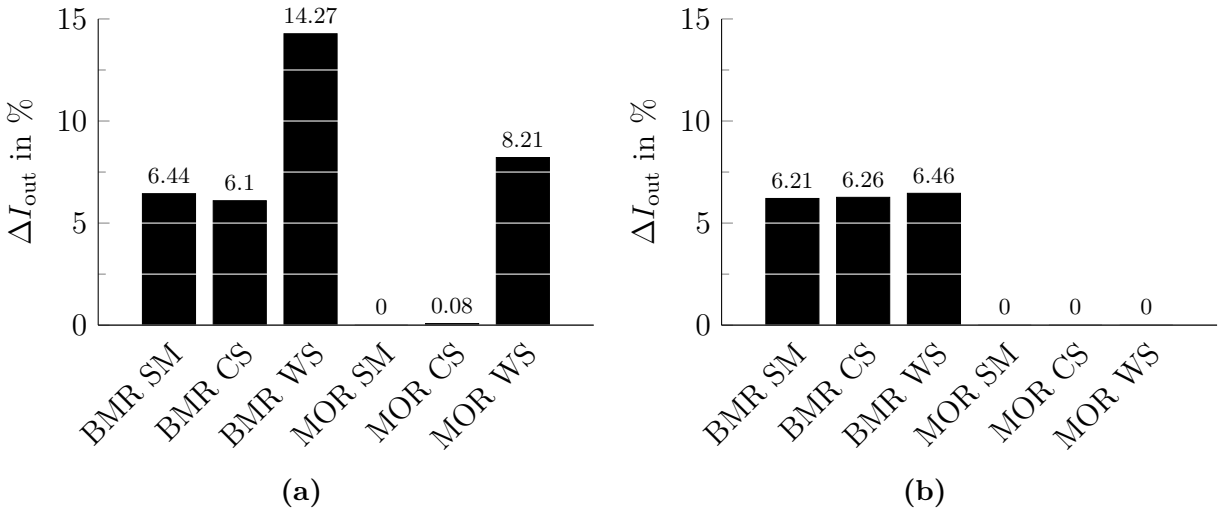


Figure 3.5: Deviation of sink currents (a) and source currents (b) for a stress time of 100 days. SM: Single MOS transistor; CS: Cascode structure; WS: Wide Swing structure.

The results are projected to the design of a folded cascode amplifier and reveal a gain in aging behavior for various circuit characteristics.

3.5.1 Degradation in Bias and Mirror Circuits

The degradation behavior of the following bias and mirror circuits is investigated: MOS only reference, beta matching reference, single transistor mirror (SM), cascode mirror (CS) and wide swing mirror (WS). The combination of the two bias circuits (MOR, BMR) and the three mirror circuits (SM, CS, WS) yields a total of 12 configurations, since each pair is available as current sink and source . The study involves degradation due to NBTI and HCD for a stress time of $t_{\text{stress}} = 100\,\text{days}$. The output voltage is set to $V_{\text{o}} = 1\,\text{V}$ for sinks and $V_{\text{o}} = V_{\text{dd}} - 1\,\text{V} = 2.3\,\text{V}$ for sources. Simulation of degradation is performed with the RelXpert reliability simulator within a 150 nm technology node using 3.3 V transistors. If not stated differently, all transistors share the dimensions as given in Table A.1 in the appendix. All circuits are designed to provide a bias current of $I_{\text{bias}} = 12\,\mu\text{A}$ at a supply voltage of $V_{\text{dd}} = 3.3\,\text{V}$.
A measure of degradation is provided using the deviation of the output current, which is given in percentage in Fig. 3.5 for a degradation time of $t_{\text{age}} = 100\,\text{days}$. Depending on the actual configuration, the maximum deviation reaches a value of almost 15 %. Current sinks, which are biased by an BMR, show a strong deviation. If the same structures are biased by a MOR, the output current seems to be more stable, especially in the case of a wide swing mirror. Both, the MOR SM as well as the MOR CS, show a deviation close to zero. Generally, sinks tend to degrade about twice as much as their source counterpart. This can be explained by the fact, that the output of sources

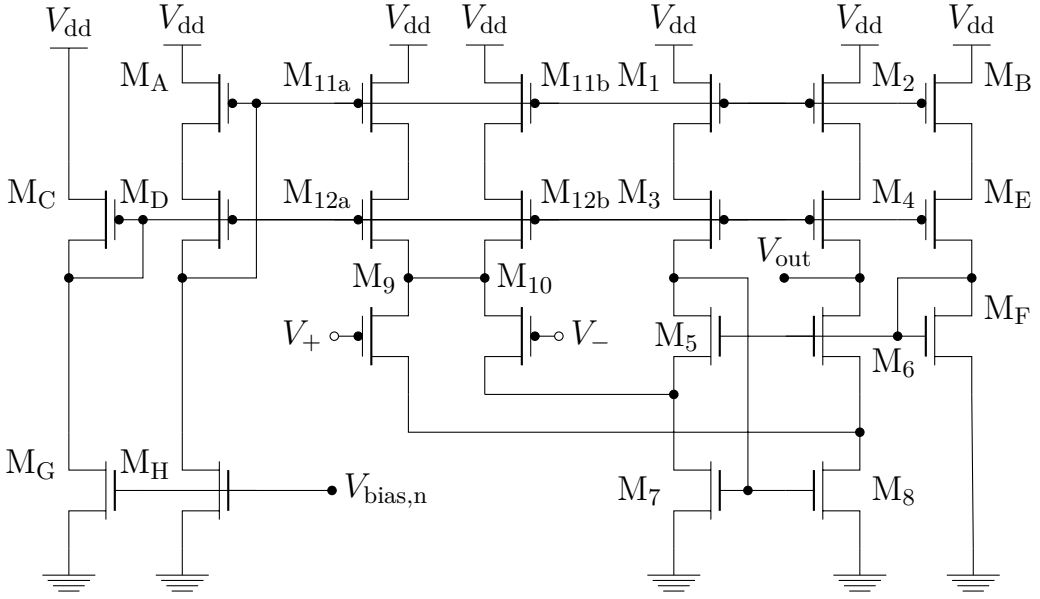


Figure 3.6: Schematic of Folded Cascode OTA.

mostly suffer from NBTI, whereas sinks experience HCD, which is susceptible to high drain voltages. The data from Fig. 3.1 implies, that the influence of V_{DS} and V_{GS} at the same value is worse for $V_{th,n}$ than it is for $V_{th,p}$ by a factor of two. The applied gate voltage for current sources are lower than the drain voltages applied for the current sinks, hence the degradation due to NBTI is less than the one induced by HCD.
The different response to degradation of biasing and mirror circuits can be used to methodically design reliable amplifier structures. By properly choosing mirror structures, the degradation of complex circuits can be greatly reduced.

3.5.2 Degradation in a Folded Cascode Amplifier

Figure 3.6 shows the schematic of a folded cascode amplifier, which has to be interpreted as two separate circuits: an inner and outer circuit. The inner circuit is constructed by transistors M_1 to M_{12}, which is the actual folded cascode amplifier. Devices M_{11} and M_{12} are shown as two sub-units ([M_{11a}, M_{11b}] and [M_{12a}, M_{12b}]), which shall emphasize that these units conduct twice the current of M_9 and M_{10}. This type of amplifier needs several bias voltages, in order keep all transistors in saturations. Namely, the group of [M_5, M_6], [M_{12a}, M_{12b}, M_3, M_4] and [M_{11a}, M_{11b}, M_1, M_2] share each a common bias voltage $V_{bias,1}$, $V_{bias,2}$ and $V_{bias,3}$ respectively. These voltages are generated by the outer circuit given by M_A to M_F. The bias network is activated through $V_{bias,n}$ connected to M_G and M_H. This voltage can be generated by reference circuits, such as a MOS only of beta matching reference, as discussed in the previous sections.
Observing Fig. 3.6 shows, that transistors M_A, M_C, M_D, M_G and M_H in combination with the cascode transistors M_{11}-M_{12}, M_1-M_3, M_2-M_4 and M_B-M_E are constructed as a wide swing current mirror (note that in comparison to Fig. 3.4b, transistor M_D is

Table 3.1: Fresh System Characteristics.

Specification	Unit	CS BMR	CS MOR	SI BMR	SI MOR
Phase Margin	°	88.6	88.7	88.6	88.6
Slew Rate Rise	V/µs	1.68	1.72	1.68	1.72
Slew Rate Fall	V/µs	2.32	2.38	2.46	2.53
DC Gain	dB	65.4	65.5	65.3	65.4
GBW	MHz	3.77	3.84	3.88	3.94
PSRR	dB	83.9	70.3	84.3	73.1
CMRR	dB	85.5	79.3	89.2	81.2
V_{dd}	V	3.3	3.3	3.3	3.3
Bias Current	µA	13.5	13.5	13.5	13.5

introduced to allow an enhanced bias voltage matching). This wide swing current mirror drives the differential input pair consisting of M_9 and M_{10}. This configuration can be modified by removing M_{12a} and M_{12b}, in which case the PMOS differential input pair (M_9, M_{10}) is driven by a single MOS current source instead of a wide swing current source.

The following analysis consists of 4 different configurations, where the bias voltage $V_{\mathrm{bias,n}}$ will be generated by a BMR and a MOR, the differential input pair is sourced by a single PMOS device and a wide swing structure. In each configuration, the whole structure is simulated in its fresh state and also for different stress times ranging from 200 days to 1000 days in a step size of 200 days. Fresh circuit characteristics are given in Tab. 3.1. The data from Tab. 3.1 shows almost no deviation in the specification values comparing the different configuration schemes. The chosen configurations may therefore be seen as identical within reasonable range.

The percentage deviation in different characteristics of the degraded folded cascode is given in Fig. 3.7 for a degradation time of $t_{\mathrm{age}} = 1000$ days. Each individual graph from Fig. 3.7a to Fig. 3.7f shows the degradation results for four configurations of bias and mirror circuit. Details of the results are presented in Tab. A.2 in the appendix.

Observing the results show that the shift due to degradation is highly dependent on the used current mirror and bias circuit. In particular, comparing the results for the DC Gain, the lowest deviation of 6.36 % is seen for the MOR SM configuration, whereas the highest value of 16.81 % is reported for BMR WS configuration. In order to allow an easy interpretation of the results, the lowest achieved deviation for each characteristic is highlighted in a gray tone. As it can clearly be observed, except for the PSRR characteristic, the single transistor current mirror biased by a MOS only reference leads to the lowest deviations from all configurations. By choosing this particular incarnation

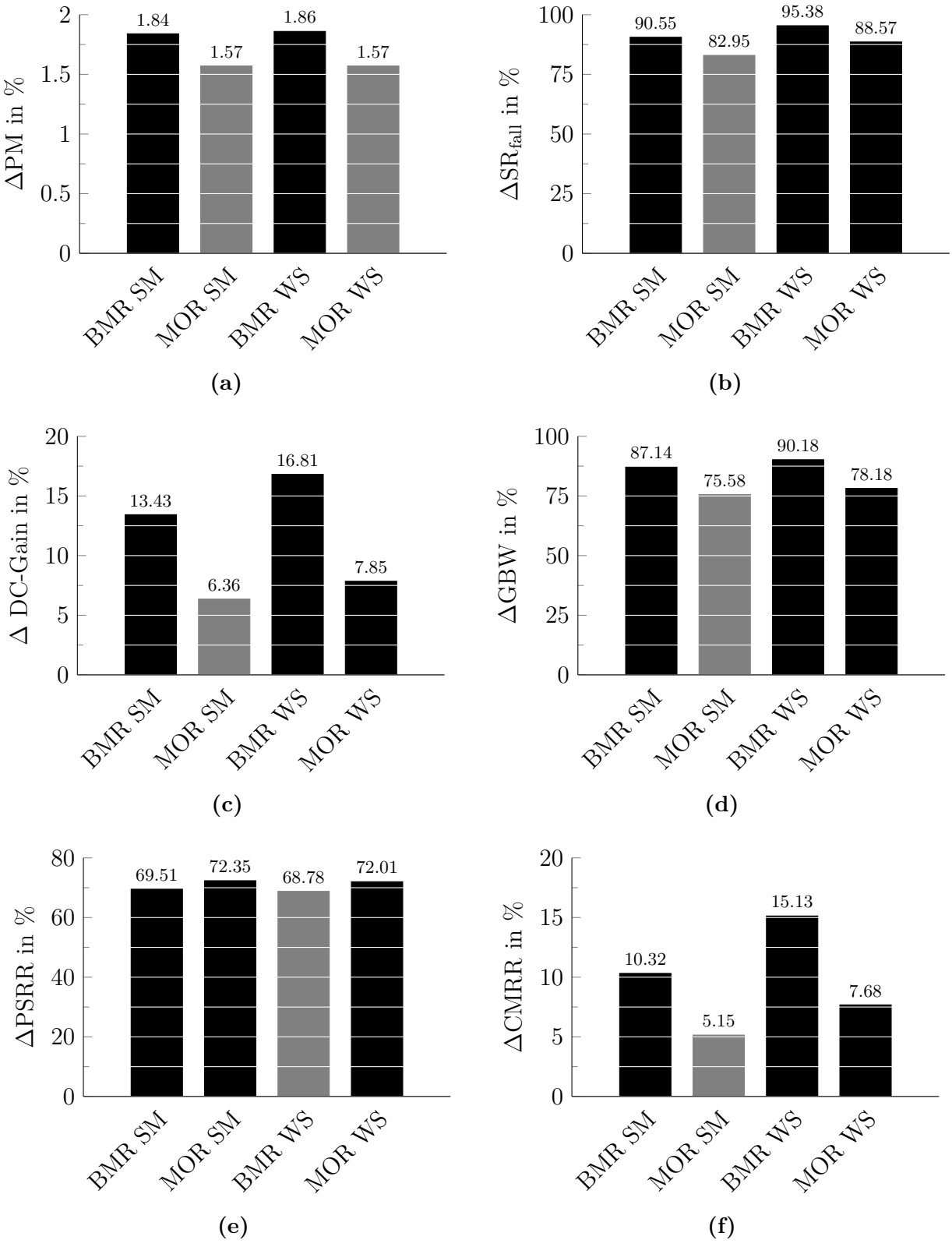


Figure 3.7: Degraded Circuit Characteristics of the folded cascode amplifier. MOR: MOS only reference; BMR: beta matching reference; SM: Single MOS transistor; WS: Wide Swing structure. (a) PM. (b) SR_{fall}. (c) DC-Gain. (d) Gain bandwidth. (e) PSRR. (f) CMRR.

of the folded cascode amplifier, a degradation-aware design is found, which can be used for safety-critical medical applications, as shown in the upcoming section 3.7.

3.6 Generation of Reliability Figures

Variability in CMOS technology nodes leads to a statistical shift in process parameters. In order to account for this stochastic distribution of technology parameters, each value is noted as a mean value and a standard deviation, e.g. for the threshold voltage V_{th} the expression is given as

$$V_{\mathrm{th}} \rightarrow \mu_{V_{\mathrm{th}}} \pm \sigma_{V_{\mathrm{th}}}, \tag{3.20}$$

where $V_{\mathrm{th}} = \mu_{V_{\mathrm{th}}}$ is referred to as the nominal value of the threshold voltage. In terms of variability, the nominal value represents the nominal process corner. Degradation modes, such as NBTI and HCD, cause a shift in transistor parameters, e.g. ΔV_{th} over time. The examination of these effects in chapter 2 did not account for the statistical behavior of the impacted parameters, but only refers to shifts of individual components in the nominal corner. The question remains, if these individual deviations only affect the mean value or also the standard deviation of a given parameter. There exist several studies in literature, which investigate the aging of parameter distribution due to NBTI and HCD [113] [44] [81]. Both groups in [113] and [44] show, that not only the mean but also the spread of a distribution is affected by degradation modes. However, both of these studies investigate rather small technology nodes, which are unsuitable for safety-critical applications, such as neural measurement systems. The author in [81] however shows, that this observation is not true for all technology nodes, especially for the 130 nm process discussed in the cited study. It can be assumed, that the degradation of the standard deviation of process parameters becomes more independent of degradation as the featured channel length increases.

3.6.1 Stochastic Circuit Simulation

The analysis, which maps the stochastic distribution of each parameter into a circuit simulation environment is known as the Monte-Carlo (MC) analysis. MC analysis generates a set of samples, each consisting of individual parameter sets derived from the statistical parameter distribution. In order to gain information on degraded circuit characteristics under the influence of variability, each derived sample has to be simulated against NBTI and HCD models. This additional degradation simulation increases the simulation time of each individual sample by at least a factor of 2, as it has been shown in [49], which results in an unreasonable simulation time. Preferably, the overall simulation time should be reduced without a decrease in accuracy of the results. This can be achieved in different ways, e.g. by the use of behavioral models [49] and also by minimizing the effort in aging simulations. Since [81] has shown, that a degradation

of the standard deviation can be neglected for nodes above 130 nm, the default MC algorithm can be rearranged to significantly reduce the number of aging simulations.

3.6.2 Variability-Aware Gradual Aging

RelXpert is used as the aging simulator, which outputs a degraded netlist description as the result of a degradation simulation. The term 'degraded netlist' is sometimes misleading, as the netlist itself does not degrade, the parameters within each component however do. Within this degraded netlist, RelXpert includes an aged model for each individually degraded transistor. Therefore, names of instances and models change compared to the original fresh netlist. This rearrangement within the netlist is not only done for single time points, but also for so called gradual aging simulation, where the degradation is subsequently simulated for given time points

$$\mathbf{t}_{\text{age}} = \{t_0, t_1, \ldots, t_n\}. \tag{3.21}$$

With this method RelXpert allows to gain access to parameter shifts over discrete time steps. Note that the degraded netlists generated by RelXpert are not compatible to statistical simulations and therefore only allow the degradation of within the nominal corner. In order to add the use of statistical simulation schemes, specifically the MC analysis, a reconfiguration of the resulting netlists has to be performed. The reconfigured aged netlists can be used in an innovative MC analysis, which efficiently calculates stochastic distributions of circuit characteristics under the influence of aging. Both aspects are described in the following.

Configuring Netlists

The reconfiguration of aged netlists is achieved using the aged netlist parser (ANP). The ANP consists of a Perl script, which is able to perform all necessary steps. A functional diagram of the ANP is given in Fig. 3.8. In a first step, the aged netlist provided by RelXpert is parsed and separated into three files containing the following:

1. Aged instances
2. Aged models
3. Symbol map

The Cadence design environment offers the tool SpiceIn, which is able to import a plain text netlist and convert it into a schematic view. This tool is used to import the netlist containing only the aged instances, which is provided by ANP. Once the instances are imported and converted into a schematic view, each instance has to be mapped to its specific aged transistor model. The symbol map associates each instance with its according model. This information is also provided by ANP. As stated before, models

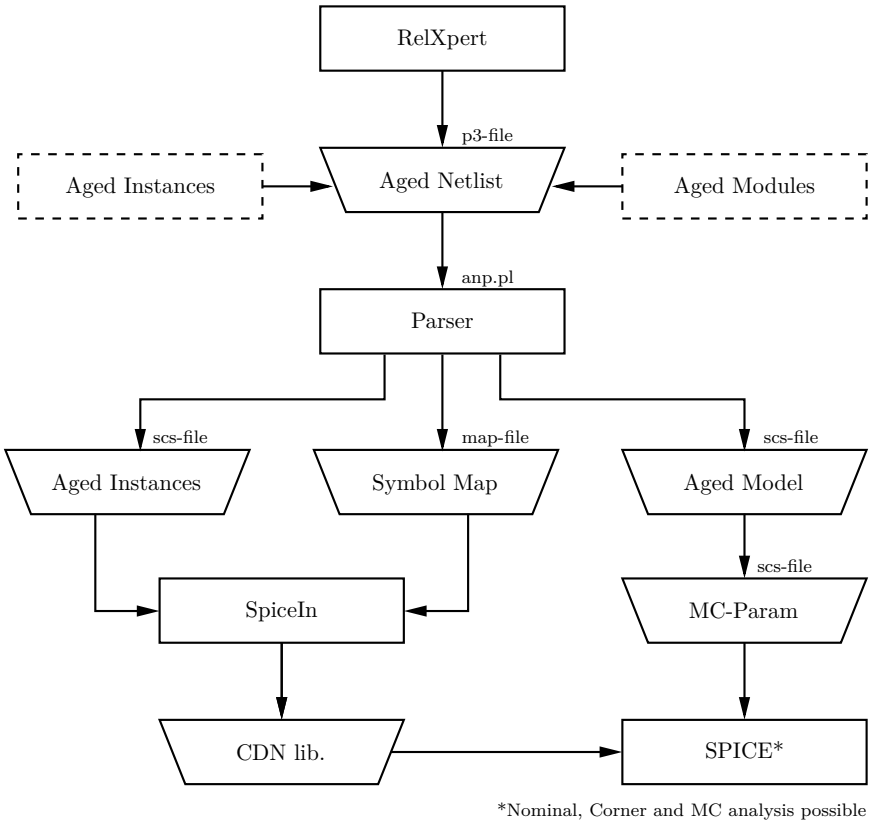


Figure 3.8: Functional diagram of the aged-netlist parser (ANP).

created by RelXpert only support nominal simulation. The degradation simulation rewrites the fresh model, such that needed MC parameters are not reported to the aged model. Given a small parameter set consisting of the threshold voltage vth0, the junction capacitance cj0 and the lateral nonuniform doping along channel nlx0, the description of the fresh netlist is expressed as

$$\begin{aligned} \mathrm{vth0} &= \mathrm{vth0_{const}} + \delta\mathrm{vth0_{std}} + \mathrm{vth0_{mc}} + \mathrm{vth0_{mis}} \\ \mathrm{cj0} &= \mathrm{cj_{const}} + \delta\mathrm{cj_{std}} \\ \mathrm{nlx0} &= \mathrm{nlx_{const}} + \delta\mathrm{nlx_{std}} + \mathrm{nlx_{mc}} \end{aligned}$$

where for each line X_{std} is controlled by corner analysis, X_{mc} is the variability parameter and X_{mis} is the mismatch parameter. The resulting degraded netlist loses the information about process variability, such that the description changes to

$$\begin{aligned} \mathrm{vth0} &= \mathrm{vth0_{aged}} \\ \mathrm{cj0} &= \mathrm{cj_{aged}} \\ \mathrm{nlx0} &= \mathrm{nlx_{aged}} \end{aligned}$$

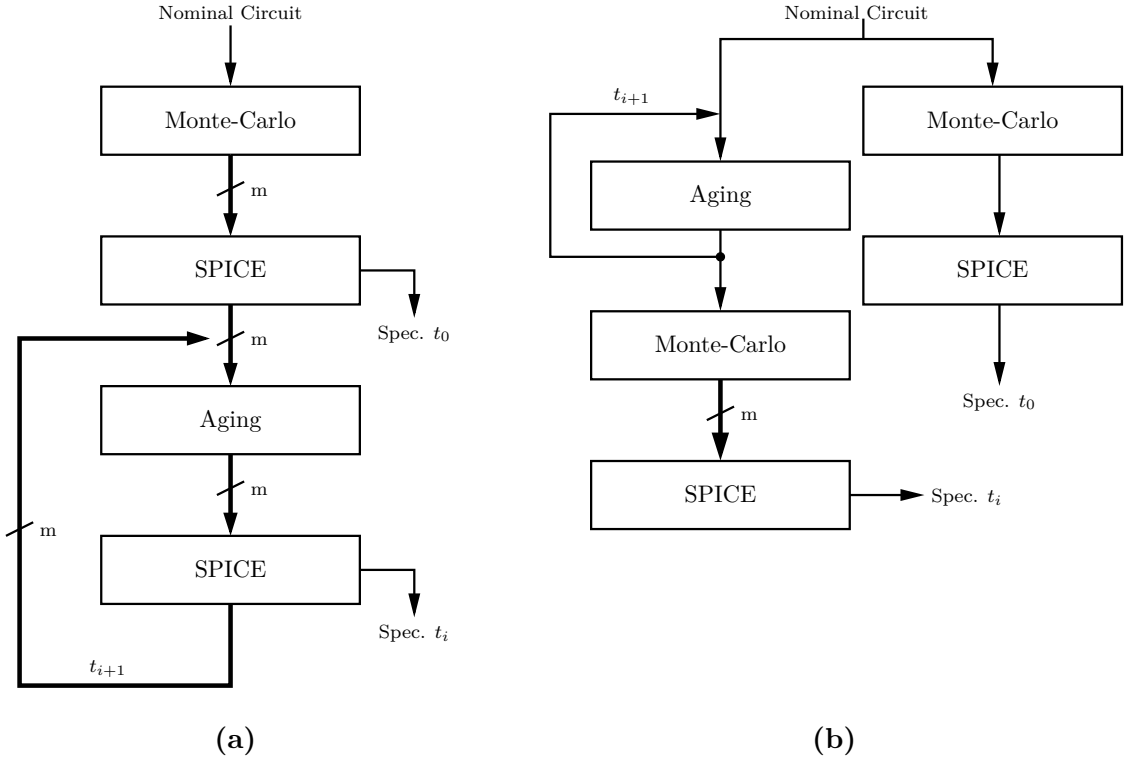


Figure 3.9: Algorithmic strategies for gradual aging under the influence of variability. (a) Direct aging. (b) Variability-aware gradual aging.

ANP detects the change of description for every involved parameter and recreates the according variability information, which is needed by the MC analysis. This method provides the basis for statistical simulation schemes on degraded circuits.

Aged Monte Carlo Analysis

In an aged Monte Carlo analysis simulation of degradation modes and process variability is performed within the same simulation run. This enables the investigation on distributed performance characteristics at a fresh and also at an aged state. The generation of an aged set of samples may be achieved using different approaches. In Fig. 3.9 two flows are proposed. With the basic MC analysis flow, a batch of samples is generated following statistical descriptions of associated process parameters. Each sample individually experiences the impact of degradation modes, which can be simulated through RelXpert. A flow diagram describing this direct aging approach is depicted in Fig. 3.9b. A graphical representation of this flow is given in 3.10 (top). Direct aging fulfills the requirement of providing subsequently aged distribution of performance values, but also requires a high number of RelXpert runs.
Since in this study no degradation of statistical spread for process parameters is assumed, the flow from Fig. 3.9b can be improved. Instead of simulating the degradation for all MC samples, only the sample from the nominal corner (nominal sample) is fed to

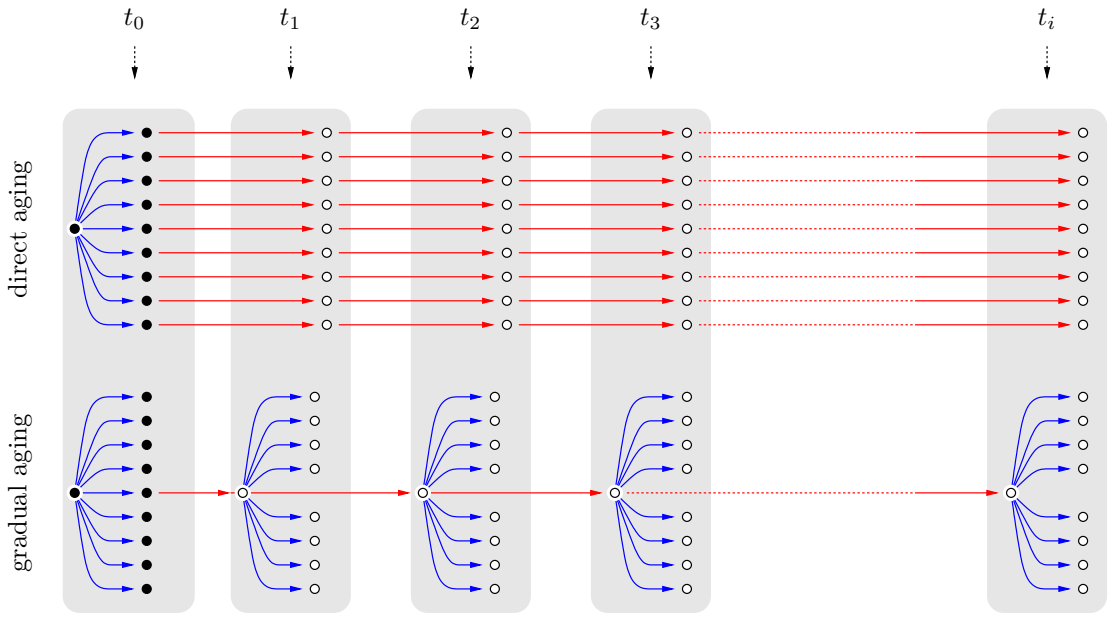


Figure 3.10: Graphical representation of the direct and gradual variability-aware aging simulation scheme.

RelXpert. ANP uses the netlist of this degraded nominal sample and adds support for statistical MC simulation. Therefore, in order to derive distributions of performance characteristics, only the nominal circuit undergoes degradation, whereas the distribution at the given stress time is provided by a simple MC analysis. A graphical representation of both strategies is shown in Fig. 3.10. A red arrow represents the simulation of degradation whereas a blue arrow is a SPICE equivalent simulation. If the number of time steps is given by n and the number of generated samples is m, then the count of RelXpert runs $R_{\#}$ is reduced from $R_{\#} = m \cdot n$ for the direct aging to $R_{\#} = n$ for the variability-aware gradual aging.

3.7 Reliability Analysis for a Neural Measurement System

Especially in safety-critical applications, such as automotive, avionic or medical, an insight of the expected reliability is mandatory for the development of such systems. Medical systems often require lifetimes in the range of decades or more, which needs to observable in the simulation environment. The variability-aware gradual aging scheme discussed in the previous section is able to provide such reliability figures, even for large and complex circuits. In the following, the generation of reliability figures is demonstrated for a neural measurement system.

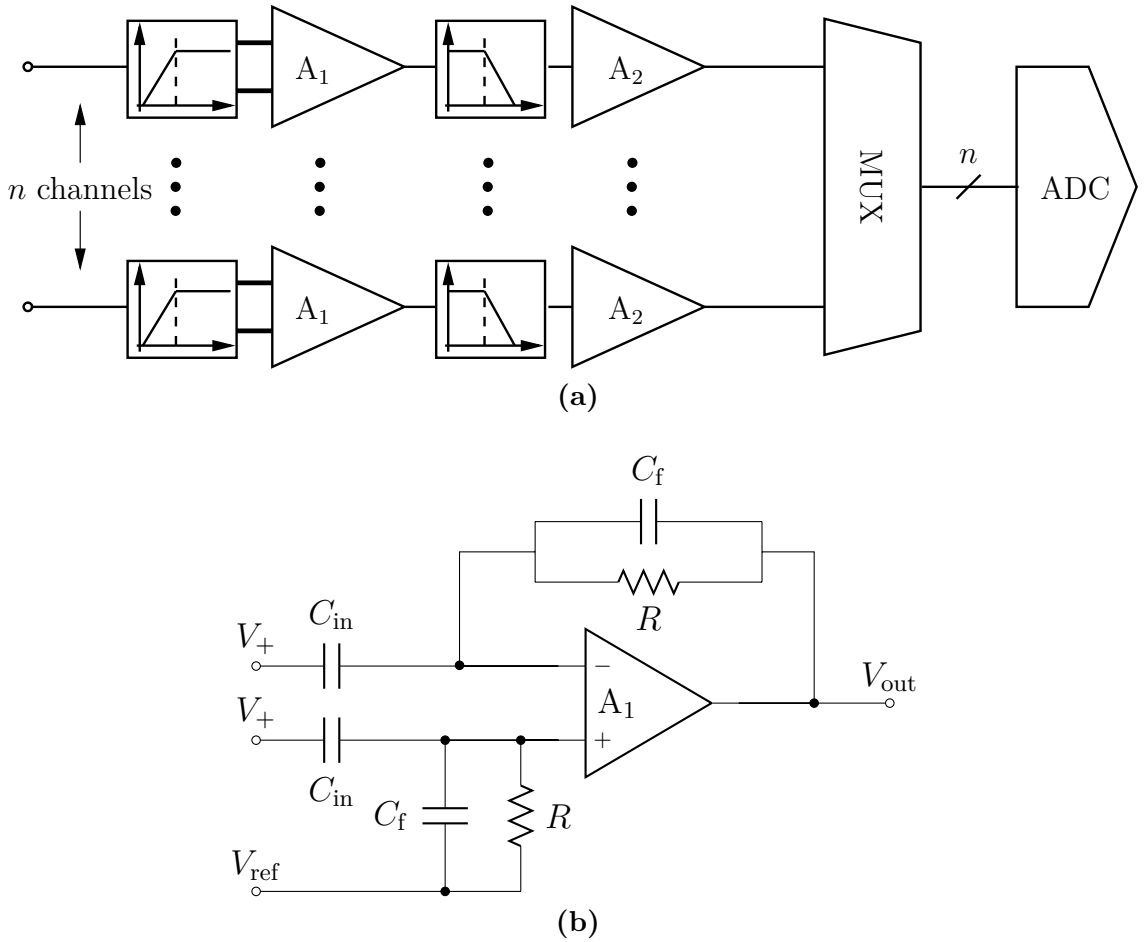


Figure 3.11: Analog sensor interface for a neural measurement system. (a) Multi-channel interface with multiplexer and analog to digital conversion. (b) High-pass filter with low noise amplifier in closed loop configuration.

3.7.1 Neural Measurement System

Neural measurement systems (NMS) are able to record neural activities inside the cortex. In general, neural signals can be distinguished into two different types: local field potentials (LFP) and neural action potentials (AP) [130] [83]. An AP signal is generated from a single neuron. The amplitude of the signal is in the range of 50 µV to 500 µV and consists of frequencies from 200 Hz to approximately 10 kHz. The actual measurement of AP is quite challenging, as in order to record an AP, the neuron has to be in direct contact with the electrode. This measurement techniques requires special electrodes, such as the Utah array [84]. These electrode arrays are inserted into the brain tissue, which leads to scarred tissues and reduced contact of the electrodes.
LFP measurement does not require direct contact to the activated neurons and is therefore preferred, considering the challenges involved with AP measurement. Local field potentials are the result of superimposed neural activities defined by multiple neurons and can therefore be mapped to a larger spatial area inside the cortex. There exist fully implantable measurement systems capable of recording LFP signals [111],

which use a flex-rigid electrode array, which lies on top of the cortex instead of being inserted into the tissue. Amplitudes of LFPs are larger than the ones of APs and can reach up to 5 mV. Due to the fact, that LFPs are a superposition of multiple APs, the resulting bandwidth is much lower and ranges from 1 Hz and 300 Hz.
Exact information regarding the amplitude and bandwidth of APs and LFPs vary in literature [56], [16], [73], but the demands on the measurement system seem to be very specific and distinct. A typical sensor interface for an NMS is shown in Fig. 3.11a. An NMS is mostly built as a multi-channel interface. Each channel records a single neural signal, which is to be digitized by the ADC. Each recording channel consists of an analog bandpass filter, which passes the desired frequencies of the LFPs or APs. Both types of signals suffer from a varying DC component below 1 Hz, therefore incoming signals are capacitively coupled and high-pass-filtered with a very low corner frequency. The high-pass is usually built using a low noise amplifier (LNA), e.g. as discussed in section 3.5.2, in a closed loop configuration as shown in Fig. 3.11b. The transfer function results to

$$H_{\mathrm{hp}} = \frac{sRC_{\mathrm{in}}}{1 + sRC_{\mathrm{f}}} = \begin{cases} \frac{C_{\mathrm{in}}}{C_{\mathrm{f}}} & \text{for } s \gg 1 \\ sRC_{\mathrm{in}} & \text{for } s \ll 1. \end{cases} \tag{3.22}$$

As seen in Eq. (3.22), the pole of the function is given by RC_{f}, which needs to be set to below 1 Hz in order to remove unwanted DC components. In order to achieve such large RC time constants, pseudoresistors are widely used [45], which reach impedances of up to 2 TΩ.
A precise recording of neural activities is essential for neural measurement systems, as these systems operate in a safety-critical environment. There exists no definite information on how accurate neural signals have to be recorded by an NMS, as researchers still work on different interpretation themes of the signals. If the data is transferred wireless outside the cortex, the available channel can limit the maximum data rate. Innovative signal processing algorithms, such as compressive sensing, are used to compress neural data without compromising accuracy [99] [98]. Nevertheless, in research applications the highest degree of available accuracy should be provided.
The NMS outputs a digitized bit stream, which is generated by the internal ADC. A way of describing the quality of the digitized signal is given by the effective number of bits (ENOB), which shall be used in the following as a main criteria. Different authors have shown that an ENOB of 8 to 10 Bit is sufficient for most neural applications. This ENOB value can be used to derive specific specifications for each sub-block of the NMS. The NMS shall be supplied by a voltage source of $V_{\mathrm{dd}} = 3.3\,\mathrm{V}$, which also acts as the reference voltage for the ADC. Considering that APs have a maximum amplitude of $V_{\mathrm{AP,max}} = 500\,\mu\mathrm{V}$, the amplification in the passband of the filter should be equal to

$$A_{\mathrm{filter}} = \frac{V_{\mathrm{dd}}}{V_{\mathrm{AP,max}}} = \frac{3.3\,\mathrm{V}}{500\,\mu\mathrm{V}} = 6600\,\mathrm{V/V} = 76.4\,\mathrm{dB}. \tag{3.23}$$

Table 3.2: Deviation of amplification ΔA_{LNA} dependent on resolution.

Resolution n	V_{LSB}	$V_{\mathrm{in,max}}$	ΔA_{LNA}	$\Delta A_{\mathrm{LNA}}\vert_{\%}$
8 Bit	12.89 mV	500 µV	0.258	0.390
9 Bit	6.45 mV	500 µV	0.129	0.195
10 Bit	3.22 mV	500 µV	0.065	0.097

The filter stage consists of two amplifiers: the LNA and a second stage operational amplifier. Both contribute to the overall amplification and are chosen to amplify at $A_{\mathrm{LNA}} = 66\,\mathrm{V/V} = 36.4\,\mathrm{dB}$ for the LNA and at $A_{\mathrm{OP}} = 100\,\mathrm{V/V} = 40\,\mathrm{dB}$ for the second stage amplifier. The bandwidth of the filter is set to satisfy the spectral range for APs of $200\,\mathrm{Hz} - 10\,\mathrm{kHz}$. The proposed method will be performed on the LNA described in section 3.5.2. All other components are considered to be free of degradation.
In order to keep the ENOB of the digitized signal within a specific specification limit, the above described parameters should not degrade and cross a certain limit. The influence of different system parameters on the ENOB varies and will be shown for the DC-Gain of the LNA. Considering an n-Bit resolution for the ADC and a reference voltage of $V_{\mathrm{ref}} = 3.3\,\mathrm{V}$, the voltage equivalent to the least significant bit (LSB) of the output signal is given by

$$V_{\mathrm{LSB}} = \frac{V_{\mathrm{ref}}}{2^n} = \frac{3.3\,\mathrm{V}}{2^n}. \tag{3.24}$$

If the signal is disturbed by an amount, which exceeds V_{LSB} prior to digitization, then the output word loses information and consequently decreases the vale of ENOB. This disturbance or influence on the signal may be the result of a degradation in amplification A_{LNA}. The shift in amplification ΔA_{LNA} due to degradation, which causes a change in the output voltage of LNA by V_{LSB} is calculated by

$$\Delta A_{\mathrm{LNA}} = \frac{V_{\mathrm{LSB}}}{V_{\mathrm{AP,max}} \cdot A_{\mathrm{OP}}} = \frac{3.3\,\mathrm{V}/2^n}{500\,\mu\mathrm{V} \cdot 40\,\mathrm{dB}} = 66\,\mathrm{V/V} \cdot 2^{-n}. \tag{3.25}$$

The results in Tab. 3.2 show the evaluated expression from Eq. (3.25) for different resolutions of n Bit, which will serve as a basis for the following reliability analysis. Inevitably, other performance characteristics, such as the phase margin PM, the slew rate SR and the gain bandwidth GBW also suffer from degradation and experience deviation. On the basis of the above described analysis for ΔA_{LNA}, equivalent expressions can be found for the remaining circuit performances, but are not considered for further analysis.
The initial design of the LNA shall provide a DC-Gain of slightly above $A_{\mathrm{LNA}} = 66\,\mathrm{V/V}$. Due to process variation, A_{LNA} shifts from sample to sample. As the specification limit is set to $A_{\mathrm{LNA,spec}} = 66$, the spread in process parameters would cause a decrease in the fresh yield. The other LNA parameters shall exceed a phase margin of $\mathrm{PM} \geq 75\,°$, a

slew rate of SR $\geq 100\,^{\mathrm{kV}}/_{\mathrm{sec}}$, a gain bandwidth of GBW $\geq 1\,\mathrm{MHz}$ and a power supply rejection ratio of PSRR $\geq 65\,\mathrm{dB}$.

3.7.2 Reliability Figures for Neural LNA

Using the proposed gradual aging flow, reliability figures for the neural measurement system are computed. As stated before, any component but the LNA will be considered free of degradation, therefore the reliability figures of the LNA is equivalent to the ones for the whole system. The investigated LNA is depicted in Fig. 3.6, but without the cascode transistor M_{12}.
Within the NMS, the LNA is used as a high-pass filter. The testbench for LNA is equivalent for the configuration shown in Fig. 3.11b. The high-pass filter is degraded for a total stress time of 10 yrs in a step size of 2 yrs. According to the gradual aging flow, within each degradation step, distributed performance characteristics are computed.
Not only the mean values, but also the standard deviations deviate over time, as observed in Tab. A.5 in the appendix. The DC-Gain shall be further analyzed using the derived specifications from Eq. (3.25). The sample distribution and cumulative distribution function of the DC-Gain are shown on Fig. 3.12 (a) and (b) respectively. Both figures are generated using $m = 3500$ samples. Vertical lines indicate specifications, which shall not be trespassed by a sample. According to Eq. (2.11), yield over time (reliability) is the relation of good samples to all samples, which is indicated through the specification lines in Fig. 3.12a. Each sample on the right of these lines is considered good, whereas each sample on left is considered faulty. The percentage of good samples is easily determined using Fig. 3.12b. The results for each specification limit are compiled in Fig. 3.12c. The initial yield value is slightly below 100 %, which is due to a trade-off between an over-designed circuit and matching the specific DC-Gain. In extreme cases, the reliability of the system is degraded to a value 9.2 % for a specification limit of 10 Bit. Note that the limits are chosen to demonstrate the generation of reliability figures and are most likely to be relaxed for in-field scenarios.

3.8 Summary

This chapter introduced a structural aging analysis for analog circuits. Different topologies of a current mirror are examined and investigated with respect to their response to NBTI and HCD. The analysis has been performed to bias and mirror circuits separately.Results are projected to a folded cascode amplifier. Through different configurations of the bias network of the amplifier, a difference in degradation is observed, which is foreseen by the degradation analysis of the current mirrors. In the second part of this chapter an approach for generating reliability figures is shown. Through an efficient reordering of Monte Carlo, SPICE and RelXpert simulations, the total simulation time has been greatly reduced.

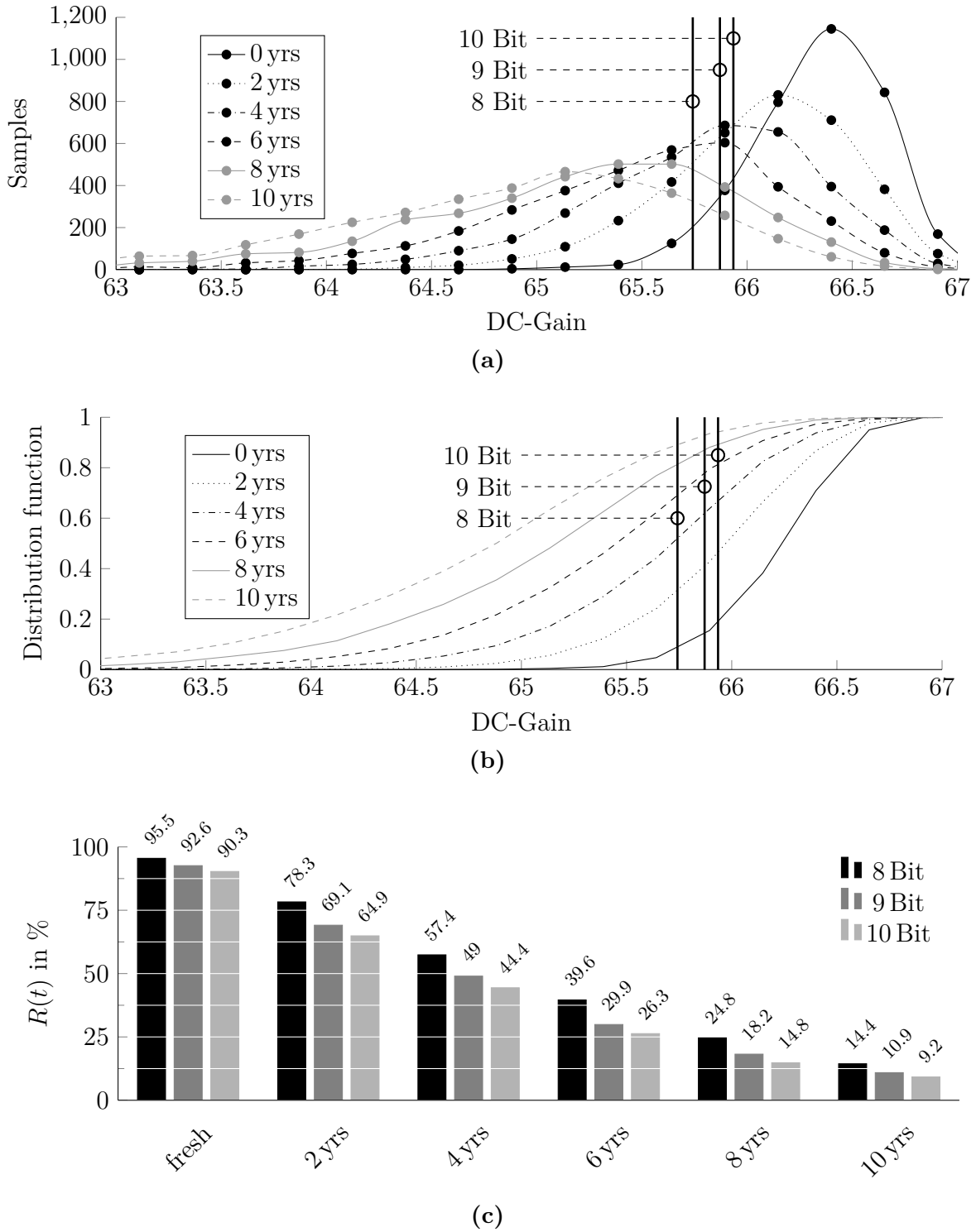


Figure 3.12: Stochastic degradation of the DC-Gain for the folded cascode amplifier. (a) Distribution of samples. (b) Cumulative distribution function. (c) Reliability function $R(t)$ for the DC-Gain of the folded cascode amplifier.

4 Operating Point Dependent-Degradation

The operating point of transistor is an important criteria in analog integrated circuit design and holds all essential small signal parameters of a device, which are evaluated for given constant terminal voltages. Especially circuits like amplifiers and comparators can almost be completely described using small signal models and do not rely on large signal models, as used in SPICE simulators. This circumstance enables a fast and straight-forward hand calculation of many design aspects, by the use of simple model equations valid in the small signal regime. Hand calculations of circuits have long been performed, but recent technology nodes provide innovative transistors with ever decreasing channel lengths, which are not compatible to outdated hand models, e.g. to the standard quadratic models. Therefore, a more sophisticated approach is presented in this section: the g_m/I_D sizing methodology.
This methodology does not only yield accurate results for fresh circuits, but can be extended to also provide predictions of degraded circuit characteristics. A method called operating point-dependent degradation is introduced in this chapter, which will combine the accuracy of the g_m/I_D sizing methodology for fresh states and projects it for an approximative degradation analysis. This chapter will first introduce different model approaches for MOS transistors and then explain the operating point-dependent degradation in depth followed by two circuit examples.

4.1 MOS Models

There exists a large number of model approaches for MOS transistors, which differ in complexity and physical accuracy. One of the most famous one is the analytical Shockley MOS model, which is still widely used in analog circuit design. Although being a very efficient hand analysis tool, it lacks of accuracy in short channel applications [95]. More precise models have been developed to close this gap, such as the SPICE LEVEL3 model [90] or the Berkeley short-channel IGFET model (BSIM). There also exist table based approaches [102] [94], which may even be applied to modern finFET technologies [110]. Within this work four basic categories of MOS models are defined:

physical, empirical, semi-physical and tabular models. This list extends and refines available categories, as found in [112] [8].

Physical models

These models are based on the specific device physics within the MOS component. All parameters use physical properties, such as the oxide thickness, substrate doping concentration and mobility of carriers. In the beginning of integrated circuit design, these models were used extensively. The most prominent representation might be Shockley's quadratic MOS model [103] and its refinement [25]. As the complexity of modern MOS devices increases, the physical modeling becomes extremely difficult and suffers from high computational effort, if used in circuit simulators. Today, simple models are mostly used for hand analysis, whereas the atomic scale models serve as a basis for empirical models.

Empirical models

Models of this type entirely rely on curve fitting and are not based on any physical parameter or constant. The exact curve is modeled from physical FEM simulations or taken from actual measurements. An empirical model uses any function or parameter, in order to most adequately fit the original curve. Model parameters have no physical meaning and are often tuples of parameters describing different orders of approximation.

Semi-physical models

The advantage of physical models lies in the use of actual, measurable parameters. Semi empirical models mainly use physical descriptions for the behavior of a MOS transistor, but add empirical methods, such as curve-fitting, in order to gain computational efficiency. These models are most used in circuit simulators, such as the BSIM model.

Tabular models

The fourth type of models is based on look-up tables. These multi-dimensional tables are built from measurements and contain the state of a MOS transistor in dependence on the terminal voltages or drain current. The accuracy within the specified data is extremely high and the computational runtime during simulation is very low. However, extrapolation of parameters outside the table data is unreliable and leads to false approximations. A table based approach is e.g. are supported by RelXpert for aging simulations.

4.2 Description of Operating Points

The operating point of a MOS transistor describes the steady state of all small signal parameters for fixed terminal voltages V_G, V_D, V_S and also V_B. Within an operating

point, the drain current I_D and also the width W and length L of the channel are fixed. The relation of process and transistor parameters causing a specific drain current I_D is given by MOS models, which have generally been categorized in section 4.1. As for the physical and semi-physical models, there exist two major descriptions to reference the operating point of a transistor. The first group (e.g. Shockley and BSIM) uses the overdrive voltage $V_{ov} = V_{GS} - V_{th}$, which is also known as the effective gate-source voltage, whereas the other group (e.g. the EKV model) combines the so called charge-based models.
The EKV model is able to continuously describe the drain current of a transistor from weak to strong inversion and also fully models the moderate inversion regime. The quadratic Shockley equations do not provide such a continuous model, which is its major drawback. Although both models derive their equations from a different perspective (charge based vs. overdrive voltage), the property of continuity during moderate inversion is not bound to either of these groups. The BSIM model version 3.3 uses fitting functions to properly model in between weak and strong inversion, but the resulting equations are not suitable for hand calculation.
In order to compare the traditional Shockley and the EKV models in regards of their accuracy for hand calculations, an intrinsic gain stage (IGS) is investigated and sized using both methods. An IGS consists of a transistors, a drain-connected constant current source I_D and a load capacitance C_L. This capacitance shall be large enough, to convert the IGS into a true first order system with only one pole. Given a fixed drain current I_D and a shape factor $S_f = {}^{W}/_{L}$, the IGS is designed to fulfill a specific GBWP

$$\omega_T = \frac{g_m}{C_L}. \tag{4.1}$$

Therefore, the transconductance g_m has to be connected to I_D and ${}^{W}/_{L}$ in order to find a solution. This connection is performed by MOS models. Using a simple quadratic approach, such as the Shockley model, the drain current in saturation is given by

$$I_D = \beta \frac{(V_{GS} - V_{th})^2}{2n}. \tag{4.2}$$

With

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = \sqrt{\frac{2\beta I_D}{n}} \tag{4.3}$$

it is possible to connect the transconductance g_m, the channel dimensions W and L and the drain current I_D to

$$\frac{W}{L} = \frac{n g_m^2}{2\mu C_{OX}} \cdot \frac{1}{I_D}. \tag{4.4}$$

The expression in Eq. (4.4) shows, that for a given g_m and I_D multiple solutions for W and L are possible. However, for very small values of I_D, or even $I_D = 0$ (when the transistor approaches weak inversion), this expression does not accurately model the

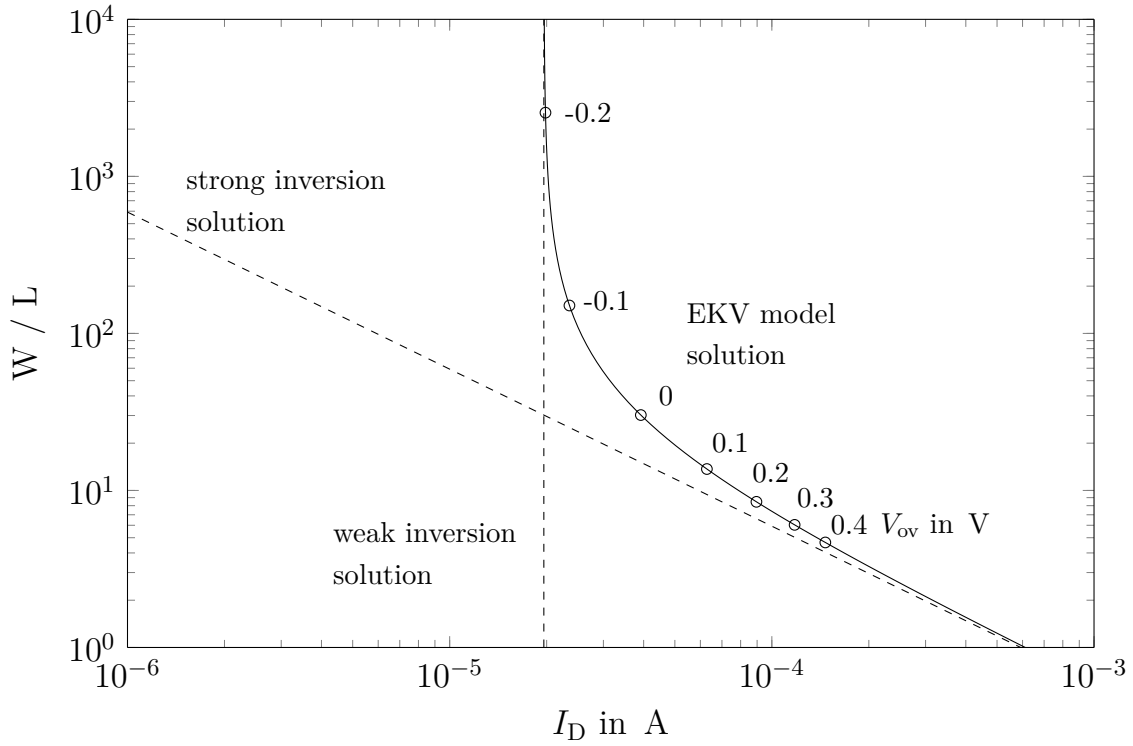


Figure 4.1: Different solutions for the shape factor $S_f = {}^{W}/_{L}$ for a constant g_m. Solutions based on weak and strong inversion model and also solution based on EKV model valid on all inversion regimes [63].

behavior of the MOS transistor. Therefore the standard quadratic approach is replaced by a different model in weak inversion. The drain current is given as

$$I_D = 2nC_{OX}v_T^2 \cdot \frac{W}{L} \exp\left(\frac{V_{GS} - V_{th}}{n \cdot v_T}\right) \tag{4.5}$$

and the relation with g_m is given by

$$I_{D,min} = nv_T g_m, \tag{4.6}$$

where $I_{D,min}$ is the minimal current to sustain a certain gain bandwidth product ω_T in weak inversion. This result is fairly different from Eq. (4.4). The transconductance alone is responsible for a fixed ω_{ugf} and does not depend on the $\frac{W}{L}$-ratio.
Sizing in moderate inversion requires an advanced model, which differs from the threshold voltage based models used before. The EKV charge-based model is able to continuously describe the drain current through all states of inversion. An expression valid from weak to strong inversion for a fixed ω_{ugf} is given by [63]

$$\frac{W}{L} = \frac{ng_m^2}{2\mu C_{OX}} \cdot \frac{1}{I_D - I_{D,min}}. \tag{4.7}$$

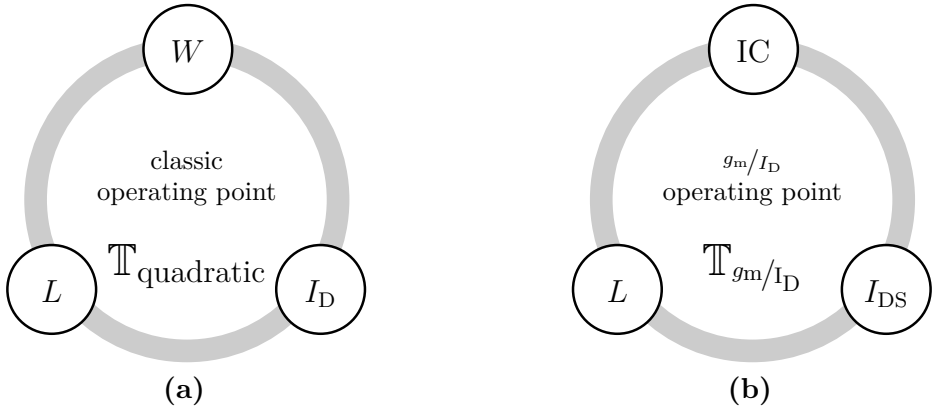


Figure 4.2: Operating point tuple for MOS transistors. (a) Classic approach. (b) g_m/I_D approach.

The solutions of the aspect ratio for a fixed GBWP (which corresponds to a fixed g_m) in weak and strong inversion compared to the solution valid in all regimes is shown in Fig. 4.1 for an example process as stated in [63]. It can be observed, that the weak and strong inversion approximations of the aspect ratio for an overdrive voltage of $-0.1 < V_{ov} < 0.4$ never fits the actual solution from Eq. (4.7). During design, this misfit leads to hand calculation results, which are far off the SPICE solution. In order to improve hand design in moderate inversion, a sizing method is needed, which is valid in all operating regions.

4.3 g_m/I_D Sizing Method

The EKV model approach (named after their inventors Enz-Krummenacher-Vittoz) provides unmatched accuracy in all operating regimes and only relies on rather simple relations and equations. Based on this continuous model, a modern sizing methodology exists, which is suitable to close the gap of needs for modern analog circuit design. The g_m/I_D sizing method uses a different description of the operating point, than used in the design with quadratic model approaches. Instead of using a tuple of $\mathbb{T}_{\text{quadratic}} = \{W, L, I_D\}$ in Fig. 4.2a, this method uses a tuple of $\mathbb{T}_{g_m/I_D} = \{\text{IC}, L, I_D\}$ as shown in Fig. 4.2b.
There are two main metrics within the g_m/I_D sizing scheme: the transconductance efficiency g_m/I_D and also the inversion coefficient IC, which is part of the new g_m/I_D operating point tuple. Both figures are discussed on the following two sections.

4.3.1 Transconductance Efficiency

The transconductance efficiency g_m/I_D combines two important metrics for analog circuit design: the drain current I_D and also the transconductance g_m. This metric proves to be highly relevant for the following reasons [105]:

1. The g_m/I_D value is strongly related to characteristic performances of analog circuits.
2. It is related to the operating region of a transistor.
3. It is suitable for calculating transistor dimensions in a tool-assisted manner.

The transconductance efficiency (or transconductance generation efficiency) is a measure on how much drain current I_D needs to be spent, in order to obtain a certain transconductance g_m. The definition of this relation is based on the logarithmic derivative of the sheet current $I_\square = I_D/S_f$ with respect to V_{GS}, which turns out to be equivalent to the logarithmic derivative of the drain current I_D with respect to V_{GS}

$$\frac{g_m}{I_D} = \frac{1}{I_D}\frac{\partial I_D}{\partial V_{GS}} = \frac{S_f}{I_D}\frac{\partial\,(I_D/S_f)}{\partial V_{GS}} = \frac{1}{I_\square}\frac{\partial I_\square}{\partial V_{GS}} = \frac{\partial(\ln I_\square)}{\partial V_{GS}}. \tag{4.8}$$

The definition in Eq. (4.8) also shows, that the value of g_m/I_D is independent of the shape factor S_f, which makes this relation universally applicable to transistors of all sizes.

4.3.2 Inversion Coefficient

The inversion coefficient IC is a fundamental metric within the g_m/I_D sizing methodology and is a numerical representation of the MOS inversion level. The term inversion coefficient has been introduced 1994 in [118] and is based on EKV MOS model, which surprisingly has been published later in 1996 [32], although first designs using this model have already been discussed in [118]. The idea of the coefficient consists of a normalization of the drain current I_D to a specific current I_{spec} and also to the shape factor $S_f = W/L$.
The specific current I_{spec} is defined as the drain current of a unit MOS device ($S_f = 1$) in moderate inversion ($I_D[\text{MI}]$). An approximation of $I_D[\text{MI}]$ can be derived from expressions for the transconductance in weak and strong inversion. The drain current $I_D[\text{WI}]$ and $I_D[\text{SI}]$ in weak and strong inversion shall be given according to [12]. With $g_m = \partial I_D/\partial V_{GS}$ it follows in weak inversion

$$I_D[\text{WI}] = 2nC_{OX}v_T^2 \cdot \frac{W}{L}e^{\left(\frac{V_{GS}-V_{th}}{n\cdot v_T}\right)}, \qquad g_m[\text{WI}] = \frac{I_D}{n\cdot v_T} \tag{4.9}$$

and in strong inversion

$$I_D[\text{SI}] = \frac{1}{2}\frac{\mu C_{OX}}{n}\frac{W}{L}(V_{GS} - V_{th})^2 \tag{4.10}$$

$$g_m[\text{SI}] = \sqrt{2I_D\left(\frac{\mu C_{OX}}{n}\right)\frac{W}{L}} = \left(\frac{\mu C_{OX}}{n}\right)\frac{W}{L}V_{ov} = \frac{2I_D}{V_{ov}}, \tag{4.11}$$

where n is the substrate factor, v_T is the thermal voltage and V_{ov} is the overdrive voltage. The current, which causes g_m in weak and strong inversion to be identical corresponds to I_D in moderate inversion and is given by

$$g_m[\text{WI}](I_D = I_D[\text{MI}]) = g_m[\text{SI}](I_D = I_D[\text{MI}]) \tag{4.12}$$

$$\frac{I_D[\text{MI}]}{n \cdot v_T} = \sqrt{2I_D[\text{MI}]\left(\frac{\mu C_{OX}}{n}\right)\frac{W}{L}} \tag{4.13}$$

$$\Rightarrow I_D[\text{MI}] = 2n\mu C_{OX}v_T^2 \cdot \left(\frac{W}{L}\right) \tag{4.14}$$

$$= I_{spec} \cdot \left(\frac{W}{L}\right). \tag{4.15}$$

The results from Eq. (4.15) corresponds to the drain current for a device operating in the center of moderate inversion where the predicted weak- and strong-inversion transconductances are equal. The expression in Eq. (4.15) is split up further into an S_f-independent part I_{spec} and the shape factor S_f itself. I_{spec} is called the specific current, which only depends on process specific parameters.
The drain current in moderate inversion from Eq. (4.15) can be used to normalize the drain current over all inversion regimes and leads to the definition of the traditional inversion coefficient

$$\text{IC}' = \frac{I_D}{2n\mu C_{OX}v_T^2\left(\frac{W}{L}\right)}. \tag{4.16}$$

The traditional inversion coefficient, as denoted in Eq. (4.16), is based on a varying specific current, since the substrate factor n as well as the mobility μ are dependent on the actual state of inversion. This definition is equivalent to normalized forward drain current i_f, which is used within the EKV model [33]. Noteworthy, the use of a normalized drain current has been used for a while in literature, but each of them differs. There exist different approaches on normalization, which allow slightly different interpretations of the g_m/I_D sizing method. In [24] the normalizing drain current is given as

$$I_{spec,[24]} = \frac{1}{2}n\mu C_{OX}v_T^2, \tag{4.17}$$

which is less by a factor of four compared to Eq. (4.16). In [105] the author completely waived the specific current and used the so called sheet current or current per square

$$I_{\square} = \frac{I_{\mathrm{D}}}{\frac{W}{L}} = \frac{I_{\mathrm{D}}}{S_{\mathrm{f}}}. \tag{4.18}$$

This normalization does not include process parameters n,μ and C_{OX}, which is still suitable for sizing, but does not allow a comparison over different processes.
The substrate factor n and the channel mobility μ are dependent on the overdrive voltage V_{ov} and therefore the specific technology current I_{spec} is not constant over different inversion regions. This change in I_{spec} is handled well in computer circuit simulation environments, such as SPICE, but complicates the hand calculation of inversion coefficients and also the associated expressions for transistor and circuit performance characteristics. In order to permit simple hand calculations during the design phase, the definition of a fixed-normalized inversion coefficient has been proven to be sufficient [12] [15] [11] [22]. Within the fixed-normalized IC the substrate factor is held at $n = n_0$, which corresponds to its average value in moderate inversion, and the mobility is set to its low-field value of $\mu = \mu_0$. This allows the normalization by a constant current I_0, which is independent of the inversion level. The final definition of IC results to

$$\mathrm{IC} = \frac{I_{\mathrm{D}}}{2n_0\mu_0 C_{\mathrm{OX}} v_{\mathrm{T}}^2 \left(\frac{W}{L}\right)} = \frac{I_{\mathrm{D}}}{I_0 \cdot \frac{W}{L}}. \tag{4.19}$$

Note that the introduced I_0 is fixed in value and therefore different from I_{spec}, which is only used for the traditional IC. Furthermore, this fixed-normalized inversion coefficient is highly related to the transconductance efficiency. Using the reasoning from Eq. (4.8) it can be seen, that

$$\frac{g_{\mathrm{m}}}{I_{\mathrm{D}}} = \frac{\partial \ln \mathrm{IC}}{\partial V_{\mathrm{GS}}}, \tag{4.20}$$

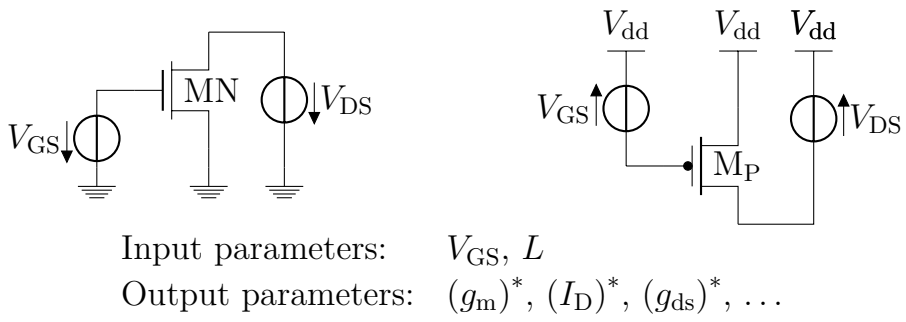
which nicely describes how these metrics interact.

4.3.3 Levels of Inversion and IC

The most important levels of inversion are weak inversion (WI) for $IC < 0.1$, moderate inversion (MI) for $0.1 < \mathrm{IC} < 10$ and strong inversion (SI) for $\mathrm{IC} > 10$. However, these regions may be split into further regimes to allow a more specific design language [12].

Deep weak inversion (IC < 0.01)

Operation in this area is not desired. The value of $g_{\mathrm{m}}/I_{\mathrm{D}}$ is at its maximum, but only differs little compared to the weak or weak-moderate inversion. These low values of IC demand very large ratios of W/L, as seen in Eq. (4.19). Operation is only suitable for ultra-low drain currents.



Input parameters: V_{GS}, L
Output parameters: $(g_m)^*$, $(I_D)^*$, $(g_{ds})^*$, ...

Figure 4.3: Testbench for generation of g_m/I_D parameters.

High side of weak inversion (IC = 0.1)
The boundary to moderate inversion. This region offers almost a maximum g_m/I_D value, low $V_{DS,sat}$ and bandwidth and high gain.

Weak-inversion side of moderate inversion (0.1 < IC < 1)
Operation here still provides high values of g_m/I_D , low V_{ov} and $V_{DS,sat}$, high gain, and an improved bandwidth compared to the high side of moderate inversion.

Center of moderate inversion (IC = 1)
Operation in the center of moderate inversion provides the best compromise of a high g_m/I_D, low V_{ov} and $V_{DS,sat}$, medium gain, and modest bandwidth.

Strong-inversion side of moderate inversion (1 < IC < 10)
This regime provides modest g_m/I_D, increasing V_{ov} and $V_{DS,sat}$, modest gain, and good bandwidth.

Onset of strong inversion (IC = 10)
Operation in this inversion level is coherent with a low g_m/I_D, high V_{ov} and $V_{DS,sat}$, low gain, and an improved bandwidth compared to the previous regimes.

Low side of strong inversion (10 < IC < 100)
Operation here provides low and declining g_m/I_D, high and increasing V_{ov} and $V_{DS,sat}$, low and declining gain, and excellent bandwidth. Due to high values of V_{ov} and $V_{DS,sat}$, this regime is not suitable for low-voltage operation.

Heavy or deep strong inversion (IC > 100)
This regime offers by far the smallest shape factors. Processing of these devises becomes difficult, as the minimum width given by the process forces large channel lengths, which in turn decrease the frequency behavior. Also, high values of V_{ov} and $V_{DS,sat}$ are not suitable for low-voltage operation.

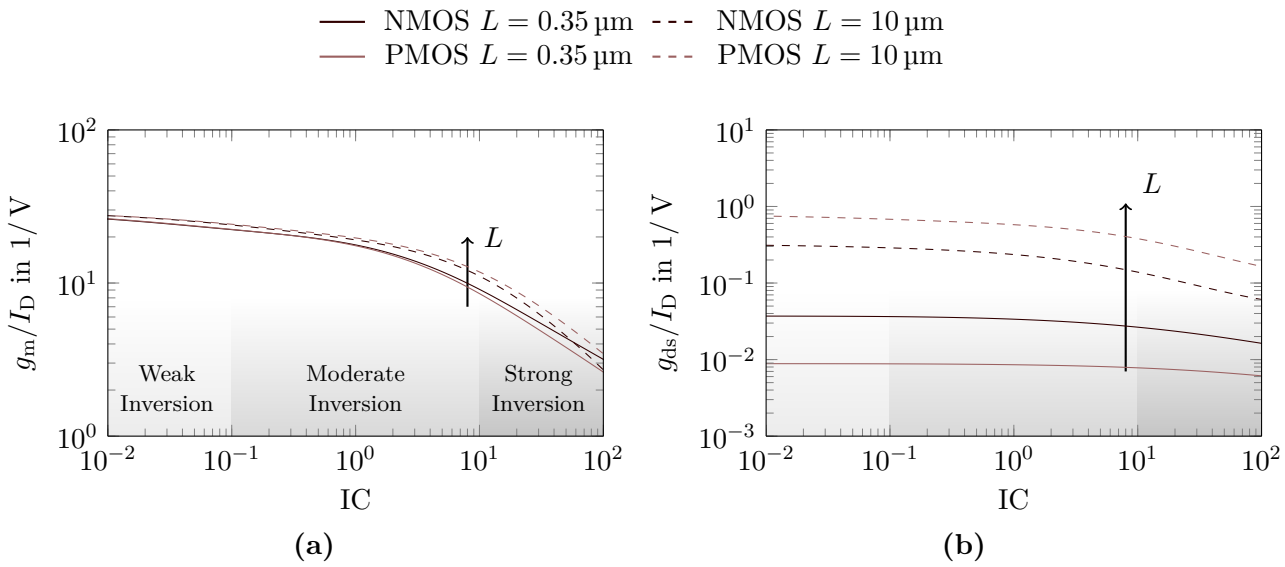


Figure 4.4: Simulated (a) g_m/I_D and (b) g_{ds}/I_D curve for a 3.3 V NMOS ($W = 10\,\mu m$) and PMOS ($W = 30\,\mu m$) transistor. Values for $L =$ are set by $\mathbb{T}_L = \{0.35\,\mu m, 10\,\mu m\}$.

4.3.4 Simulation Setup

In Fig. 4.3 the testbench for obtaining g_m/I_D, IC and related parameters is shown. The inversion coefficient is independent of V_{DS} (aside from short channel effects) as long as the transistor is kept in saturation with $V_{DS} \geq V_{DS,sat}$. Note that even in weak inversion for negative overdrive voltage, which is $V_{GS} < V_{th}$, the minimum $V_{DS,sat}$ is limited to [33]

$$V_{DS,sat}[\text{WI}] \geq 3v_T \text{ to } 5v_T, \tag{4.21}$$

showing that an operation in saturation is also possible in the sub-threshold regime. In this configuration, any value of V_{GS} for a specific L yields a different operating point. The use of a biased bulk is usually not considered within the g_m/I_D sizing scheme. It is also possible to use gate-drain-connected MOS diodes ($V_G \equiv V_D$) to generate the g_m/I_D data. However, a decoupling of V_{DS} from V_{GS} is necessary to introduce aging effects, as it will be discussed in section 4.4. The output of the testbench in Fig. 4.3 are all operating point parameters from the PMOS M_P and NMOS M_N for combinations of $\mathbb{T}_{g_m/I_D}$. In order to separate results from the testbench with ordinary parameters, output parameters are marked with $(\cdot)^*$.

4.3.5 Sizing Rules within the g_m/I_D Scheme

The inversion coefficient IC and the transconductance efficiency g_m/I_D have shown to be I_D and W/L independent metrics, which are suitable to describe all sizes of a specific transistor (e.g. NMOS or PMOS). Using the simulation results from the testbench in section 4.3.4, both metrics can be used in a single plot, as shown in Fig. 4.4a for an

NMOS and PMOS. Each device is depicted as a short and long channel device. The different regions of inversions are plotted as reference.
Each point on the this line represents a single operating point, which is defined by $\mathbb{T}_{g_\mathrm{m}/I_\mathrm{D}}$. The difference per device type is negligible for the weak and moderate inversion, however in strong inversion the PMOS preserves a larger $g_\mathrm{m}/I_\mathrm{D}$ for a long channel, but the opposite is true for the short channel device. The influence of the channel length is clearly observed, as a short channel degrades the achievable transconductance efficiency greatly. This plot can be used to determine a specific g_m from a given drain current I_D or vice versa, using

$$g_\mathrm{m} = \left(\frac{g_\mathrm{m}}{I_\mathrm{D}}\right)^* \cdot I_\mathrm{D}. \tag{4.22}$$

Note that $(g_\mathrm{m}/I_\mathrm{D})^*$ is known from simulation and almost independent of L. Therefore, Eq. (4.22) holds for any device size.
A similar approach is used for the output conductance g_ds, specifically $g_\mathrm{ds}/I_\mathrm{D}$, as shown in Fig. 4.4b for the same devices as before. Obviously, the independence of L is not provided anymore, resulting in a strong sensitivity with respect to the channel length. Also, the difference per device type is more clear than for $g_\mathrm{m}/I_\mathrm{D}$. A specific g_ds derived from a given I_D (or vice versa) is determined by

$$g_\mathrm{ds} = \left(\frac{g_\mathrm{ds}}{I_\mathrm{D}}\right)^* \cdot I_\mathrm{D} \text{ for a specific } L_0. \tag{4.23}$$

As noted in Eq. (4.23), this expression is only valid for the channel length L_0, which was used to generate $(g_\mathrm{ds}/I_\mathrm{D})^*$. Interpolating between two values of L is possible to reduce the simulation overhead, but this technique introduces some inaccuracy. Once g_m, g_ds and L (and ultimately I_D) have been chosen, $\mathbb{T}_{g_\mathrm{m}/I_\mathrm{D}}$ is satisfied. Therefore, the device width is then obtained using

$$W = (W)^* \frac{I_\mathrm{D}}{(I_\mathrm{D})^*} = \frac{I_\mathrm{D} \cdot (L)^*}{I_0 \mathrm{IC}}, \tag{4.24}$$

which is derived from the definition of the normalized inversion coefficient in Eq. (4.19).

4.4 Operating Point and Degradation

Degradation modes such as NBTI and HCD alter parameters of transistors, which ultimately leads to a shift in the operating point and therefore changes the output characteristics of whole circuits. This change in the operating point is dependent on the degradation time t_age. So far, the operating point within the $g_\mathrm{m}/I_\mathrm{D}$ sizing methodology

is described by the three dimensional tuple $\mathbb{T}_{\mathrm{g_m/I_D}}$ composed of IC, I_{DS} and L. In order to account for the age-dependent shift, this tuple needs to be expanded to

$$\mathbb{T}_{\mathrm{age}} = \{\mathrm{IC}, I_{\mathrm{DS}}, L, t_{\mathrm{age}}\}. \tag{4.25}$$

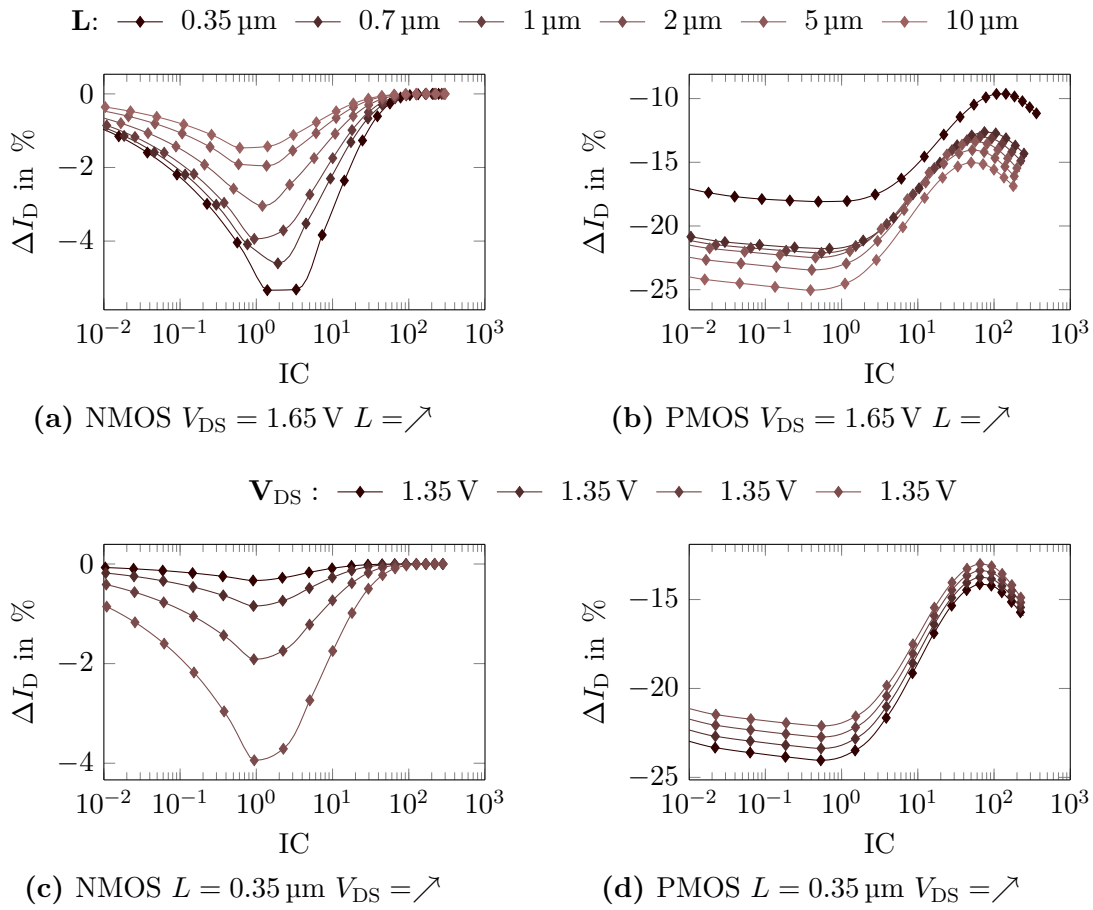
This change in $\mathbb{T}_{g_{\mathrm{m}}/\mathrm{I_D}}$ needs to be adapted by the testbench describe in Fig. 4.3. The input parameters are expanded by two additional entries V_{DS} and t_{age}. V_{DS} is introduced to account for hot carrier degradation, as this effect is highly dependent on the lateral field between the pinch-off and drain terminal in the channel. As the pinch-off changes in dependence on V_{DS}, this voltage is held variable to observe the effect of different drain source voltages. For each set of V_{GS}, V_{DS} and L the transistor is aged with an aging time of t_{age}. Degradation is simulated using the RelXpert simulator. Each output parameter is captured in its fresh and aged state.
The resulting data reveals information on operating point-dependent degradation. Each degradation of a transistor parameter may vary with a change in V_{GS}, V_{DS} or L. In general, all transistor parameters may be investigated, but the most influencing in analog design are the drain current I_{D}, the threshold voltage V_{th}, the transconductance efficiency $g_{\mathrm{m}}/I_{\mathrm{D}}$, and the output conductance efficiency $g_{\mathrm{ds}}/I_{\mathrm{D}}$, which are discussed in the next sections.
Observation of the results reveals basic characteristics of degraded parameters, which remain unchanged and independent of the exciting varying input parameter. Therefore, it is assumed that specifically L and V_{DS} only effect the magnitude of the shift due to degradation, but do not alter the general shape. The shift of each observed output parameter will be denoted as ΔP and is defined as

$$\Delta P = \frac{P|_{\mathrm{aged}} - P|_{\mathrm{fresh}}}{P|_{\mathrm{fresh}}} \cdot 100. \tag{4.26}$$

The degraded parameter set is obtained for a stress time of $t_{\mathrm{age}} = 10\,\mathrm{yrs}$, so ΔP is the difference in parameter value for the fresh and aged value. In the following a chosen set of transistor parameters shall be discussed in regards of their degradation behavior, specifically I_{D}, V_{th}, $g_{\mathrm{m}}/I_{\mathrm{D}}$ and $g_{\mathrm{ds}}/I_{\mathrm{D}}$.

4.4.1 Drain Current I_{D}

The shift in drain current ΔI_{D} due to NBTI and HCD in dependence on the inversion coefficient is depicted in Fig. 4.5. Both types, PMOS and NMOS share a general shape, which is independent on the sweeping parameter (L or V_{DS}). The NMOS has a maximum shift in moderate inversion at values around $0 < \mathrm{IC} < 5$. The shift increases for a decrease in L and an increase in V_{DS}, as shown in Fig. 4.5a and 4.5b. NMOS types are susceptible to HCD, which corresponds to the proportional growth of ΔI_{D} with a variation in V_{DS}. The PMOS type behaves differently. This is mostly due to the



(a) NMOS $V_{DS} = 1.65\,V$ $L = \nearrow$ **(b)** PMOS $V_{DS} = 1.65\,V$ $L = \nearrow$

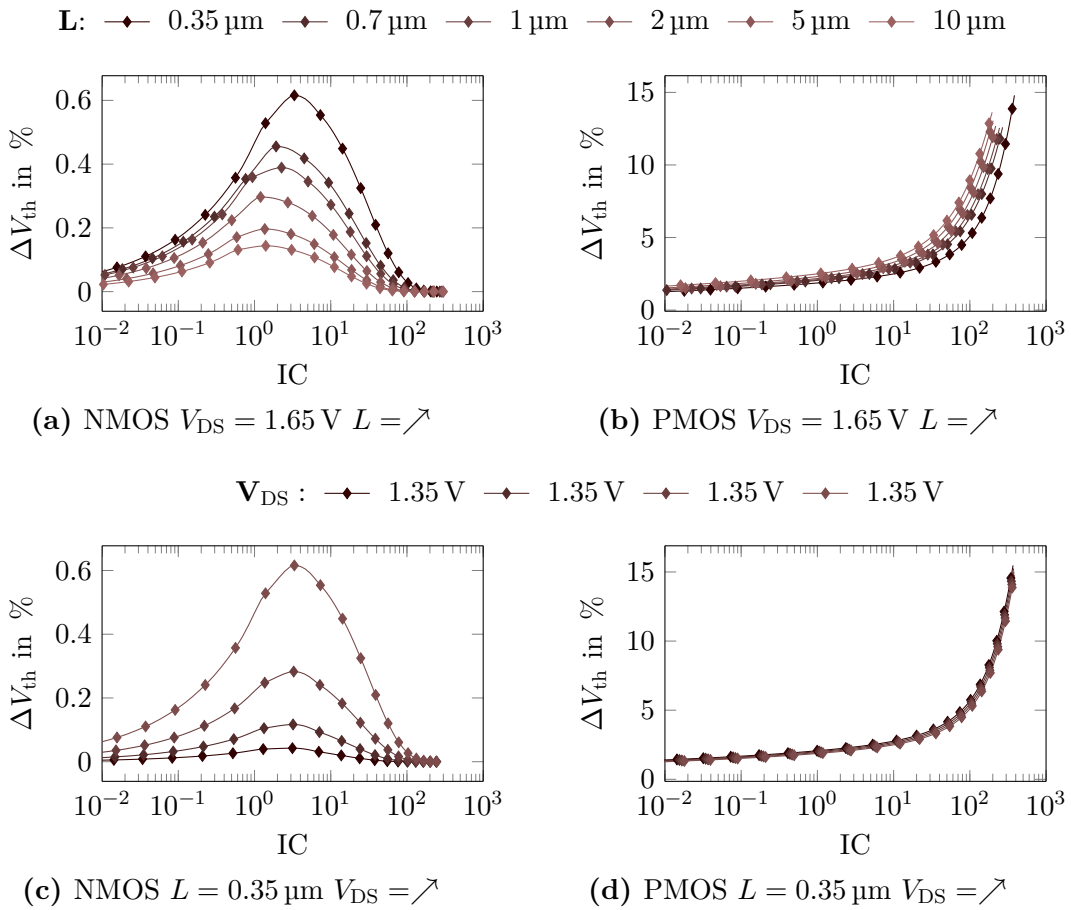
(c) NMOS $L = 0.35\,\mu m$ $V_{DS} = \nearrow$ **(d)** PMOS $L = 0.35\,\mu m$ $V_{DS} = \nearrow$

Figure 4.5: Shift in drain current ΔI_D for $150nm$ NMOS and PMOS devices over different lengths L and drain voltages V_{DS}. $L = \nearrow \hat{=} \{0.35, 0.7, 1, 2, 5, 10\}\,\mu m$, $V_{DS} = \nearrow \hat{=} \{1.35, 1.45, 1.55, 1.65\}\,V$.

different aging mechanism (NBTI). The shift ΔI_D is almost constant for IC < 1 and then decreases in strong inversion until a plateau is reached. This can be explained by the fact, that in strong inversion the pinch-off point reaches far in the active area within the channel, which decreases the area where NBTI can occur. Therefore, higher degrees of inversion provide an improved aging behavior over lower inversion regimes. The influence of L and V_{DS} is inverse to the observations of a constant V_{DS}, as higher channel lengths and lower drain-source voltages worsen the effect of NBTI.

4.4.2 Threshold Voltage V_{th}

As described in chapter 2, each sift in the transistor parameter is based on a shift in the threshold voltage ΔV_{th}, which is depicted in Fig. 4.6 for different variations in L and V_{DS}. The shift in $V_{th,n}$ for an NMOS device shows a distinct behavior, which is similar to ΔI_D as observed in Fig. 4.5a and 4.5c. The most degradation is seen on the strong-inversion side of moderate inversion for approximately ($1 <$ IC < 8), which decreases for smaller as well as larger values of IC. Depending on the used channel



(a) NMOS $V_{\mathrm{DS}} = 1.65\,\mathrm{V}$ $L = \nearrow$ **(b)** PMOS $V_{\mathrm{DS}} = 1.65\,\mathrm{V}$ $L = \nearrow$

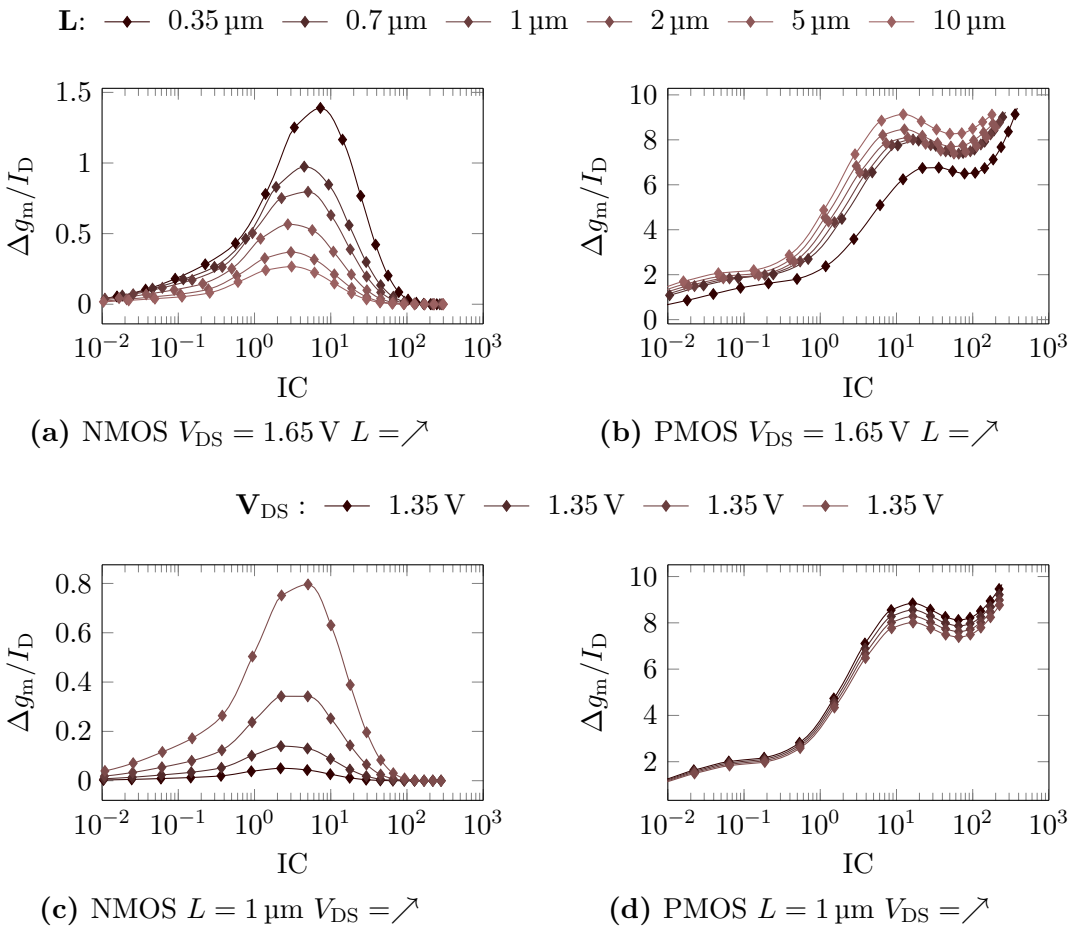
(c) NMOS $L = 0.35\,\mu\mathrm{m}$ $V_{\mathrm{DS}} = \nearrow$ **(d)** PMOS $L = 0.35\,\mu\mathrm{m}$ $V_{\mathrm{DS}} = \nearrow$

Figure 4.6: Shift in threshold voltage ΔV_{th} for 150nm NMOS and PMOS devices over different lengths L and drain voltages V_{DS}. $L = \nearrow \hat{=}\{0.35, 0.7, 1, 2, 5, 10\}\,\mu\mathrm{m}$, $V_{\mathrm{DS}} = \nearrow \hat{=}\{1.35, 1.45, 1.55, 1.65\}\,\mathrm{V}$.

length, the percentage deviation ranges from about 0.1 % to just above 0.6 % (Fig. 4.5a). This range is also given for a variation in V_{DS} in Fig. 4.5c.
The PMOS device only suffers from NBTI instead of HCD, which results in different observations. Both variations (L in Fig. 4.6b and V_{DS} in Fig. 4.6b result in similar deviation of the threshold voltage, which ranges from 1 % up to almost 9 %. This shift is worse by a factor of 10 compared to the influence on the NMOS device. Also, the degradation increases exponentially in dependence on IC, which is ultimately given by the fact, that high values of IC rely on large overdrive voltages and thus high values of V_{G}. A variation in L and V_{DS} has a rather low effect on ΔV_{th}.

4.4.3 Transconductance Efficiency $g_{\mathrm{m}}/I_{\mathrm{D}}$

The NMOS device in Fig. 4.7a and 4.7c shows expected behavior, which has already been discussed for ΔI_{D} and ΔV_{th} in section 4.4.1 and 4.4.2 respectively. However, the PMOS device degradation is unique for the transconductance efficiency. For both variations in L and V_{DS} in Fig. 4.7b and 4.7d the deviation in $\Delta g_{\mathrm{m}}/I_{\mathrm{D}}$ is rather constant



(a) NMOS $V_{DS} = 1.65\,V$ $L = \nearrow$ **(b)** PMOS $V_{DS} = 1.65\,V$ $L = \nearrow$

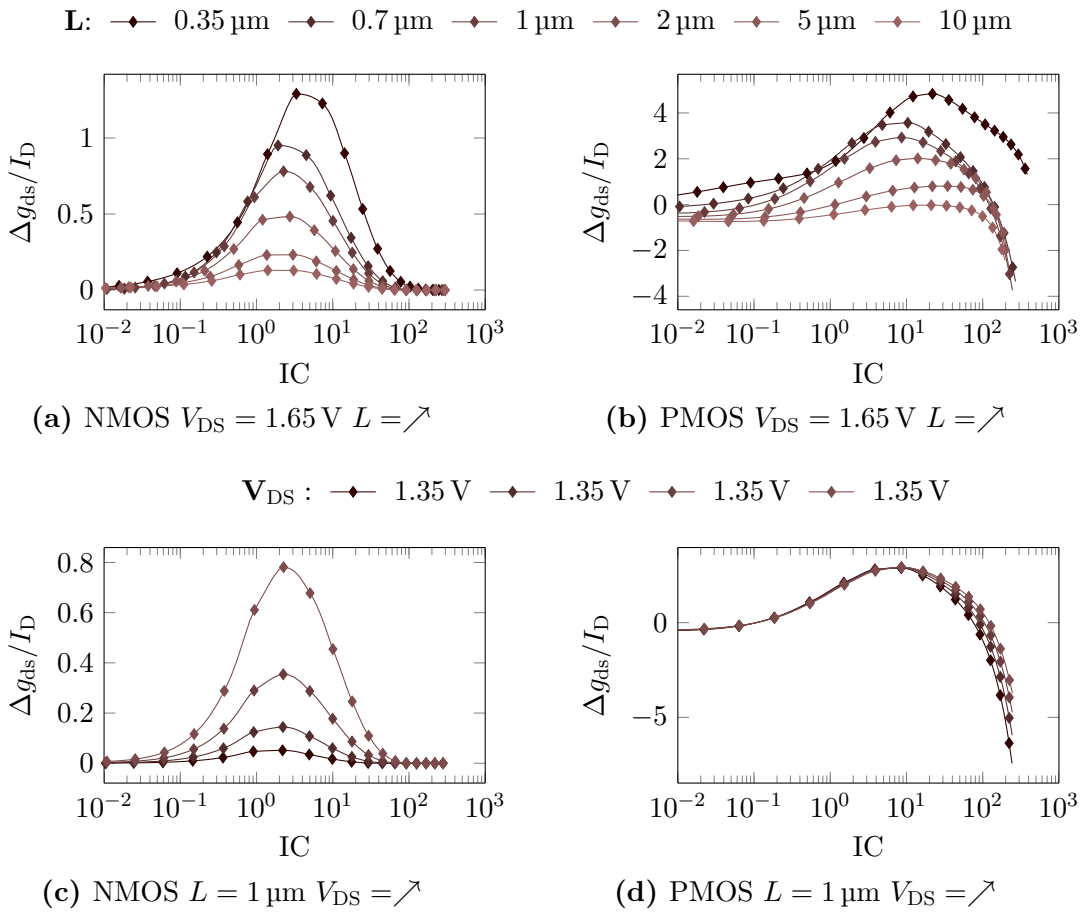
(c) NMOS $L = 1\,\mu m$ $V_{DS} = \nearrow$ **(d)** PMOS $L = 1\,\mu m$ $V_{DS} = \nearrow$

Figure 4.7: Shift in drain current $\Delta g_m/I_D$ for $150nm$ NMOS and PMOS devices over different lengths L and drain voltages V_{DS}. $L = \nearrow \hat{=} \{0.35, 0.7, 1, 2, 5, 10\}\,\mu m$, $V_{DS} = \nearrow \hat{=} \{1.35, 1.45, 1.55, 1.65\}\,V$.

and below 3 % until the weak-inversion-side of moderate inversion. From the center of moderate inversion at IC ≈ 1 until the onset of strong inversion the deviation steadily increases and is limited to about 8.5 %. For higher values of IC the degradation slightly decreases by about 2 % until IC ≈ 100 and from there rises with an increase in IC. This behavior is due to the individual degradation of g_m and I_D. A variation in L seems to have a larger effect on the amplitude of deviation than a variation in V_{DS}.

4.4.4 Output Conductance Efficiency g_{ds}/I_D

The output conductance efficiency is an important metric within the g_m/I_D sizing methodology and is therefore investigated in regards to its response to degradation. Again, the NMOS device in Fig. 4.8a and 4.8c shows a behavior, which has been observed through the other metrics as well. Observing the response to a variation in the channel length L in Fig. 4.8b shows a large change in $\Delta g_{ds}/I_D$. Despite the small channel device for $L = 0.35\,\mu m$, the general shape of the curve is similar, showing an increase in degradation within the moderate inversion and a continuous fall going into



(a) NMOS $V_{DS} = 1.65\,V$ $L = \nearrow$ **(b)** PMOS $V_{DS} = 1.65\,V$ $L = \nearrow$

(c) NMOS $L = 1\,\mu m$ $V_{DS} = \nearrow$ **(d)** PMOS $L = 1\,\mu m$ $V_{DS} = \nearrow$

Figure 4.8: Shift in drain current $\Delta g_{ds}/I_D$ for $150nm$ NMOS and PMOS devices over different lengths L and drain voltages V_{DS}. $L = \nearrow \hat{=} \{0.35, 0.7, 1, 2, 5, 10\}\,\mu m$, $V_{DS} = \nearrow \hat{=} \{1.35, 1.45, 1.55, 1.65\}\,V$.

deep strong inversion. Interestingly, the sign of degradation changes dependent on IC for certain channel lengths. The described shape is also seen for a variation in V_{DS} in Fig. 4.8d, but in this case $\Delta g_{ds}/I_D$ is almost independent of a change in V_{DS}.

4.5 Aged Performance Characteristics of a Common Source Amplifier

The results based on the current operating point in section 4.4 revealed interesting correlations between the inversion coefficient, channel length and drain potential with respect to device degradation of specific transistor parameters. These degraded parameter values can be used to approximate aged performance characteristics. As this analysis only covers the degradation of single transistors, interaction of multiple devices, such as seen in a circuit, is not simulated. Therefore, operating point-dependent degradation cannot provide the same results for circuits, as e.g. fully featured reliability simulators. However, as the following sections show, the approximated degradation results are still

suitable to be used within analog design and are a valid method for optimizing circuit characteristics. A second example, investigating a more complex circuit, is given in section 4.7.

In the following the operating point-dependent degradation is used optimize circuit characteristics towards higher resilience against aging. The method is performed for a common source amplifier design, consisting of an NMOS input transistor and a PMOS used as current source. The output node is open without any capacitive or resistive load connected to it. The circuit characteristic of interest is the DC-Gain A_{cs} with a target of 40 dB, which is composed of the g_{m} of the input transistor divided by the sum of output transconductances,

$$A_{\mathrm{cs}} = 20\log\left[\frac{g_{\mathrm{m,n}}}{g_{\mathrm{ds,n}} + g_{\mathrm{ds,p}}}\right] \stackrel{!}{=} 40\,\mathrm{dB}, \tag{4.27}$$

where the indices n and p denote parameters belonging to the NMOS and PMOS respectively. Within the $g_{\mathrm{m}}/I_{\mathrm{D}}$ sizing method, Eq. (4.27) is rewritten to

$$A_{\mathrm{cs}} = 20\log\left[\frac{\left(\frac{g_{\mathrm{m,n}}}{I_{\mathrm{D}}}\right)^{*}\cdot I_{\mathrm{D}}}{\left(\frac{g_{\mathrm{ds,n}}}{I_{\mathrm{D}}}\right)^{*}\cdot I_{\mathrm{D}} + \left(\frac{g_{\mathrm{ds,p}}}{I_{\mathrm{D}}}\right)^{*}\cdot I_{\mathrm{D}}}\right] = 20\log\left[\frac{\left(\frac{g_{\mathrm{m,n}}}{I_{\mathrm{D}}}\right)^{*}}{\left(\frac{g_{\mathrm{ds,n}}}{I_{\mathrm{D}}}\right)^{*} + \left(\frac{g_{\mathrm{ds,p}}}{I_{\mathrm{D}}}\right)^{*}}\right]. \tag{4.28}$$

As seen in Eq. (4.28), the amplification is only dependent on the normalized g_{m} and g_{ds} values, which are previously known from simulation results. Note that in order to obtain final dimensions of the transistors, the drain current of $\mathrm{M_n}$ and $\mathrm{M_p}$ has to be set to a specific value. The channel width can then be calculated through Eq. (4.24).

Two designs shall be compared regarding the expected aging and the use of the degraded $g_{\mathrm{m}}/I_{\mathrm{D}}$ and $g_{\mathrm{ds}}/I_{\mathrm{D}}$ tables. Both designs D_1 and D_1 should yield the identical gain of 40 dB. The dimension of each transistor is restricted in order to add some design constraints. Since the dimensions are constrained, the bias current I_{bias} as well as the inversion coefficients $\mathrm{IC_n}$ and $\mathrm{IC_p}$ need to be varied.

In an initial design D_1, as shown in Tab. 4.1, the desired gain is achieved with a degradation of $\Delta A_{\mathrm{cs}} = -0.31\,\mathrm{dB}$ for a stress time of $t_{\mathrm{age}} = 10\,\mathrm{yrs}$. The shift ΔA_{cs} shall now be reduced. Simulation data show, that the shift in the output transconductances $g_{\mathrm{ds,n}}/I_{\mathrm{D}}$ and $g_{\mathrm{ds,p}}/I_{\mathrm{D}}$ have the most influence on ΔA_{cs} (compared to $g_{\mathrm{m,n}}/I_{\mathrm{D}}$). Also, the shift in $g_{\mathrm{ds,p}}/I_{\mathrm{D}}$ seems to be dominant over the shift in $g_{\mathrm{ds,n}}/I_{\mathrm{D}}$. In order to optimize A_{cs} towards less susceptibility of aging, $\mathrm{IC_p}$ is adjusted to minimize the effect of $g_{\mathrm{ds,p}}/I_{\mathrm{D}}$. In design D_2 from Tab. 4.1 $\mathrm{IC_p}$ is changed from 366 to 281. $\mathrm{IC_n}$ and I_{bias} are also adjusted to maintain the dimensions of the transistors and keep the gain at 40 dB. As a results, the degradation of the gain has been reduced to $\Delta A_{\mathrm{cs}} = -0.06$, which is less by factor of more than five compared to design D_1.

Both designs share the same fresh gain A_{cs} by design. A change in I_{bias}, $\mathrm{IC_n}$ and $\mathrm{IC_p}$ has reduced the influence of degradation modes while maintaining the desired specification of $A_{\mathrm{cs}} = 40\,\mathrm{dB}$. However, a change of I_{bias} ultimately alters other characteristics,

Table 4.1: Initial design D_1 and aging-aware design D_2.

	IC_n –	IC_p –	I_{bias} (µA)	V_{bias} (V)	$S_f(M_n)$ µm/µm	$S_f(M_p)$ µm/µm	A_{cs} dB	ΔA_{cs}|10 yrs dB
D_1	50	366	103	0.96	10/1	30/10	39.72	−0.31
D_2	40	281	81	0.91	10/1	30/10	40.45	−0.06

such as the slew rate. A method which is capable of incorporating multiple output characteristics is presented in chapter 5.

4.6 Tool-Assisted Design within the g_m/I_D Scheme

Since the g_m/I_D sizing methodology is based on previously generated simulation data, many authors use electronic design automation to simplify and automate designs of integrated circuits. There exist specifically designed tools, which exploit the fundamental g_m/I_D sizing approach [31] [12]. One approach is shown in Fig. 4.9, which is called the *analog designer* from C. Enz. This tool shows and sets basic transistor parameters, which in this case are based on the EKV model, but these are also valid for the g_m/I_D sizing methodology. Within the analog designer six different groups of parameters can be adjusted, where within each group parameters are proportionally linked to each other. For example, the first group holds the normalized forward current i_f (equivalent to the inversion coefficient), the overdrive voltage $V_{ov} = V_{GS} - V_{th}$ and pinch-off voltage in reference to source $V_p - V_S$. If one of these parameters is set through the slider, all others within this group are automatically set as well. The graphical user interface shown in Fig. 4.9 is able to set the operating point of a single transistor, in this specific example the interface is shown for an n-channel type. Although the analog designer is able to intuitively design a single transistor, the interaction of multiple transistors and their impact on circuit characteristics is not covered.
This work proposes a new an innovative approach for the combination of EDA and the g_m/I_D sizing method: the GMID-Tool [55]. The main purpose is to provide a design environment, which enables direct access to simulation results generated within the g_m/I_D scheme in its fresh and also in aged states. The GMID-Tool includes handling of simulation results, setting of operating points, retrieving final dimensions of transistors and also constructing performance characteristics, which depend on multiple transistors. The GMID-Tool consists of two individual parts, one responsible for the design of a single device (device designer) and one for characteristics of a whole circuit (circuit designer). Within the device designer each single transistor can individually be adjusted in its operating point. Construction of output performance values for whole circuits, such as gain or bandwidth, is performed in the circuit designer. Additionally, this tool is able to include information on operating point-dependent degradation, which is

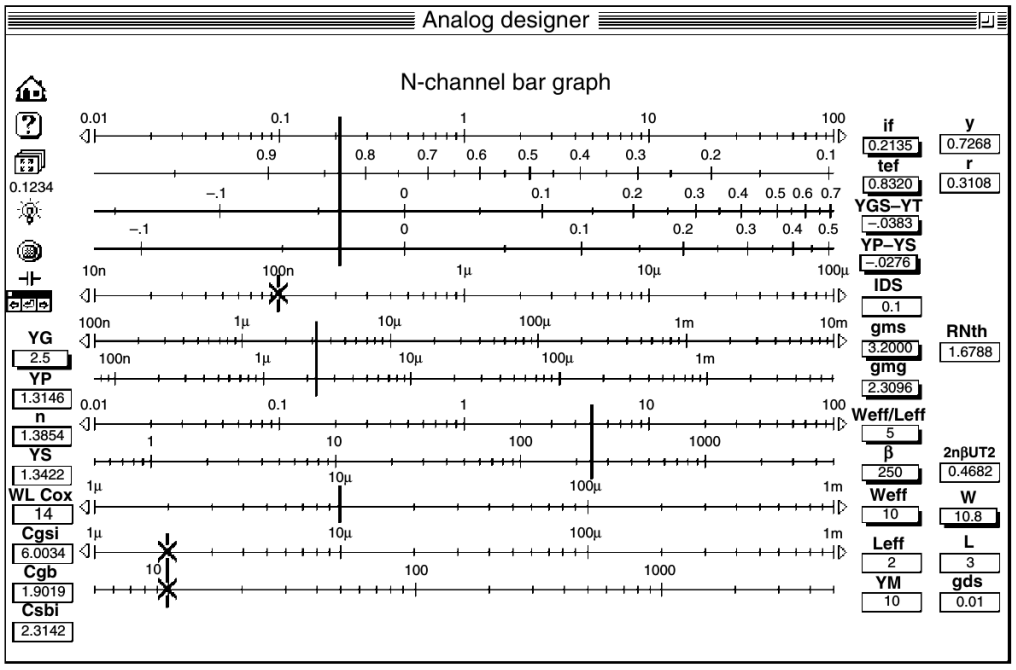


Figure 4.9: Tool called analog designer for setting operating points of transistors as shown in [31].

directly available to the designer. By this inclusion, the traditional g_m/I_D sizing method is expanded to an aging-aware design flow, where the impact of degradation on single transistors as well as system performance values can be approximated during the design phase.

4.6.1 The GMID-Tool

The GMID-Tool is specifically designed to account for multiple transistors, each contributing to various circuit characteristics. The inside of the GMID-Tool consists of three blocks, which each contain different information classes. There are *Circuit*, *MOS* and *Process* blocks. The *Circuit* block is the main object and contains a list of circuits elements. Each element is differentiated in transistor and non-transistor elements. These non-transistor elements may also contain different elements (transistor and non-transistor objects). This allows a cascaded structure of nested circuits, as it is also possible e.g. in SPICE netlists. The *MOS* block holds the operating point data of a transistor element, which is the tuple $\mathbb{T}_{\text{age}}$ and also the type of the transistor (NMOS, PMOS). Process-dependent information, such as the values for mobility μ, substrate factor n etc. are stored within the *Process* block. The operating point tuple $\mathbb{T}_{\text{age}}$ from the *MOS* block can interact with this information to calculate individual transistor parameters (g_m, g_m/I_D, W) using the sizing rules from section 4.3.5.
Simulation data is obtained for every iteration from the testbench in Fig. 4.3. The resulting operating point parameters $\mathbf{d}$ need to be saved in a structure, which allows a convenient access to all relevant data points. Also, this structure should allow an

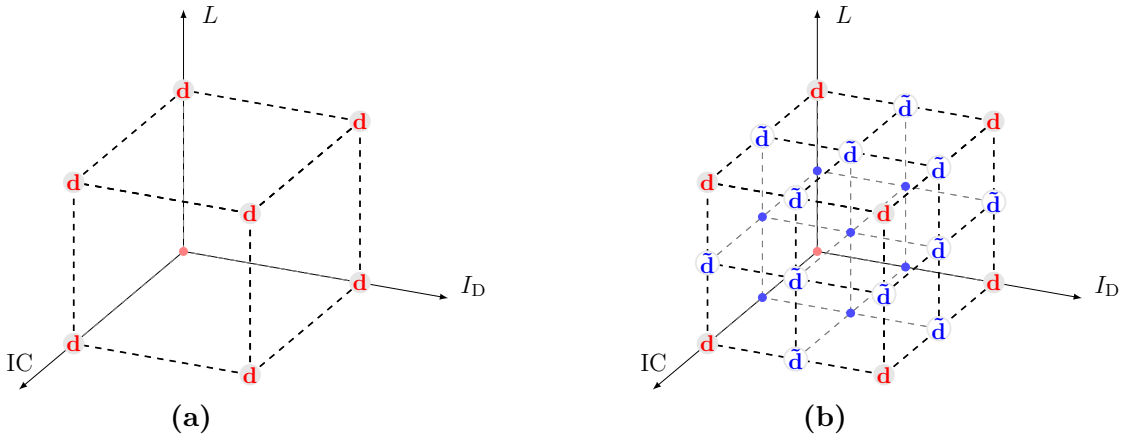


Figure 4.10: Interpolation of simulation results using predefined interpolation methods. (a) Location of **d**-values in the grid. (b) Interpolation of simulated data.

interpolation of the simulation data, if additional data is required by the designer. The GMID-Tool uses a multidimensional matrix for this purpose: the **d**-grid. This grid stores **d** data in dependence on $\mathbb{T}_{\text{age}}$. The testbench for generating g_m/I_D metrics in Fig. 4.3 two types of parameters are listed: input and output parameters. Both types need to be processed in order to generate the full **d**-grid. A graphical interpretation of the grid is shown in Fig 4.10, where Fig. 4.10a shows the grid without interpolation and Fig. 4.10b depicts an interpolation index of 1. Both share an outer dimension of three (L, IC, I_D), which represent the members of the operating point tuple $\mathbb{T}_{\text{age}}$. The fourth dimension t_{age} is not shown. Red marked entries represent data generated from Fig. 4.3, whereas blue fields are interpolated values. Markers inside the cube are only shown as colored dots to improve visual appearance of the plot. The generation of the parameter set may be very time consuming, depending on the step size and number of iterations defined for the initial simulation run. The sweep of V_{GS} should be performed logarithmically, as this helps to concentrate the density of data points in V_{GS} where needed.

There are parameters, such as the inversion coefficient IC, which are not natively available through simulation results. In order to calculate these values, additional information needs to be provided. As stated before, this information is held within the *Process* block. The generation of such second order output parameters is handled through a configuration file. An example of a configuration file for the *Process* block is shown in Fig. 4.11. The configuration file uses two sections, where in the first process-dependent constants are defined and in the second one mathematical descriptions of operating point data is found. The `EQUATIONS` section holds data from the simulation itself (prefixed with `sim_`) and constants, which have been defined in the upper section. For example, in order to calculate the value of IC, three parts need to be known: the previously simulated drain current `sim_ids`, the specific current `I_0` and also the device dimensions. All of these parameters are known from the `CONST` section. It is also

```
[CONST]
cox = 4.54e-3
Ut = 0.026
nW = 10e-6
nu0 = 370e-4
pW = 30e-6
pu0 = 126e-4

[EQUATIONS]
%Ids = 1e-6
IC = abs( sim_ids ) ./ ( I0 .* ( W ./ L ) )
I0 = 2 * n * u0 * cox * Ut * Ut
gm = abs(sim_gmoverid .* %Ids)
gds = abs((sim_gds./sim_ids) .* %Ids)
vov = abs( sim_vgs ) - abs( sim_vth )
```

Figure 4.11: Example of a configuration file for GMID-Tool [55].

possible to source results from other equations, such as for `I_0`. The calculation of g_m relies on the simulated value of g_m/I_D (`sim_gmoverid`) and also a chosen drain `%Ids`. This current is neither given through the simulation nor defined through a constant. It needs to be set by the designer and therefore, any equation holding this current, needs to be evaluated at runtime. In order to separate the simulated from the runtime drain current the percentage sign `%` is introduced.

4.6.2 Aging-Aware Design Flow

The aging-aware flow, which is adapted by the GMID-Tool, is depicted in Fig. 4.13. The flow graph compares the traditional and the aging-aware approach for generating degraded performance value of system characteristics. Within the traditional flow, which is based on the reliability simulator RelXpert, the circuit is given as a SPICE netlist. In order to determine fresh circuit performances, an ordinary SPICE solver is used to evaluate the netlist. The degraded representation of these values is gained from a reliability simulation in RelXpert, as discussed in chapter 2.4.4. A designer has to decide, whether the degraded performance values are within certain specification bounds. If the specifications are not met, the designer needs to change the fresh circuit until the desired aged performance characteristic is obtained. This flow is rather inefficient, as multiple simulation tools need to be incorporated, which is tedious and unhandy, and also the simulation time rises significantly, as degradation simulations in RelXpert prove to be very time consuming.
The proposed aging-aware flow does not build upon a netlist description. Within this flow, the output descriptions need to be described by analytical equations. The

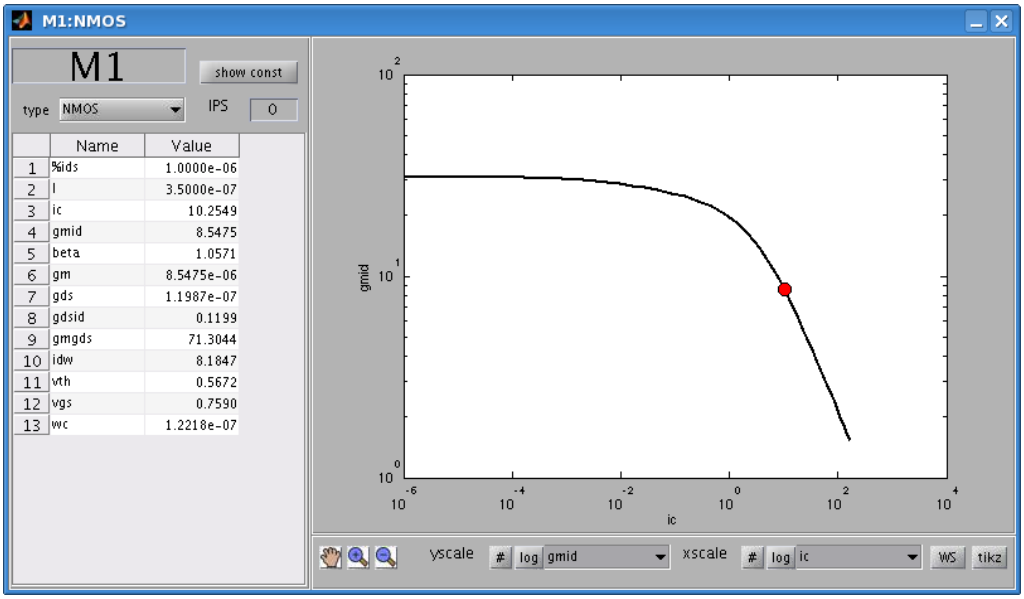


Figure 4.12: Transistor sizing within the GMID-Tool, showing operating point parameters on the left and a visualization of the g_m/I_D curve on the right.

evaluation of these equations is performed by using either the fresh or degraded g_m/I_D parameters, which directly yields fresh or degraded circuit characteristics. Note that the evaluation at an aged state is performed without any RelXpert simulation at design time. As explained before, the aged operating point data needs only to be derived once within a technology process node. The traditional as well as the aging-aware flow both provide access to fresh and aged performance values, but the latter one reaches by far a superior computational efficiency and can speed up the design process incomparably.

4.7 Aging-Aware Design of a Miller Amplifier

Section 4.5 already demonstrated the usage of the operating point-dependent degradation information in order to optimize the gain of a simple common source amplifier. In this section we use the same information for a Miller operational amplifier design within a 350 nm technology node and the use of the GMID-Tool. The schematic of the amplifier is shown in Fig. 4.14, the according dimensions are given in Tab. A.6 in the appendix. The Miller operational amplifier consists of a differential input stage and a an output stage, which is implemented as a common source amplifier. The input stage is given by M_1 to M_5 and the output stage is constructed by M_6 and M_7. Transistor M_8 is used for biasing only. Both stages contribute to the overall gain, which is denoted as

$$A_{\text{Miller}} = A_{\text{I}} \cdot A_{\text{II}}, \tag{4.29}$$

where A_{I} and A_{II} denote the DC-Gain of the input and output stage respectively. The gain for each stage is calculated using the individual transconductances of the composing

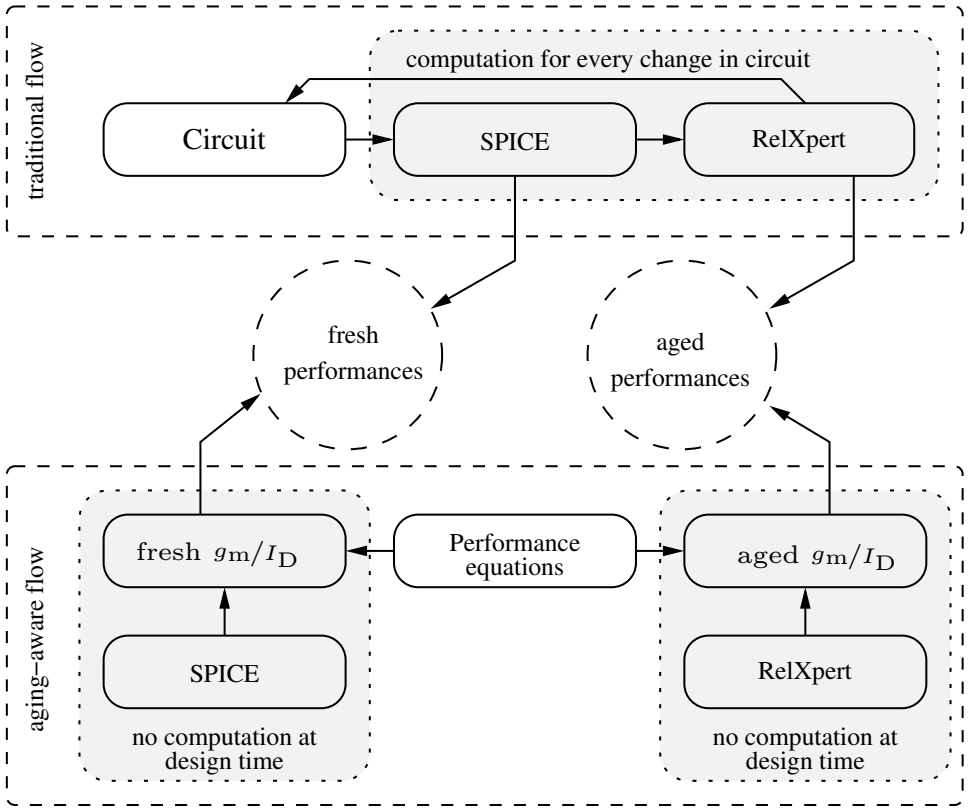


Figure 4.13: Aging-Aware design flow using the GMID-Tool. Performance values are available for fresh and aged circuit state [55].

transistors. In order to reduce the number of influencing parameters, M_5 is excluded in the description of the first stage gain A_I. The overall unsigned gain results to

$$|A_{\text{Miller}}| = \underbrace{\left(\frac{gm,1}{gds,1+gds,3}\right)}_{|A_I|} \cdot \underbrace{\left(\frac{gm,7}{gds,7+gds,6}\right)}_{|A_{II}|}. \tag{4.30}$$

Each individual parameter is replaced by its normalized equivalent within the g_m/I_D sizing methodology. Both stages may be biased by different currents, which will be denoted as I_I and I_{II} for the input and output stage respectively. The expression from Eq. (4.30) is reorganized to

$$|A_{\text{Miller}}| = \left(\frac{\left(\frac{gm,1}{I_{DS,1}}\right)^* \cdot I_I}{\left(\frac{gds,1}{I_{D,1}}\right)^* \cdot I_I + \left(\frac{gds,3}{ID,3}\right)^* \cdot I_I}\right) \cdot \left(\frac{\left(\frac{gm,7}{ID,7}\right)^* \cdot I_{II}}{\left(\frac{gds,7}{ID,7}\right)^* \cdot I_{II} + \left(\frac{gds,6}{ID,6}\right)^* \cdot I_{II}}\right) \tag{4.31}$$

$$|A_{\text{diff}}| = \left(\frac{\left(\frac{gm,1}{ID,1}\right)^*}{\left(\frac{gds,1}{ID,1}\right)^* + \left(\frac{gds,3}{ID,3}\right)^*}\right) \cdot \left(\frac{\left(\frac{gm,7}{ID,7}\right)^*}{\left(\frac{gds,7}{ID,7}\right)^* + \left(\frac{gds,6}{ID,6}\right)^*}\right). \tag{4.32}$$

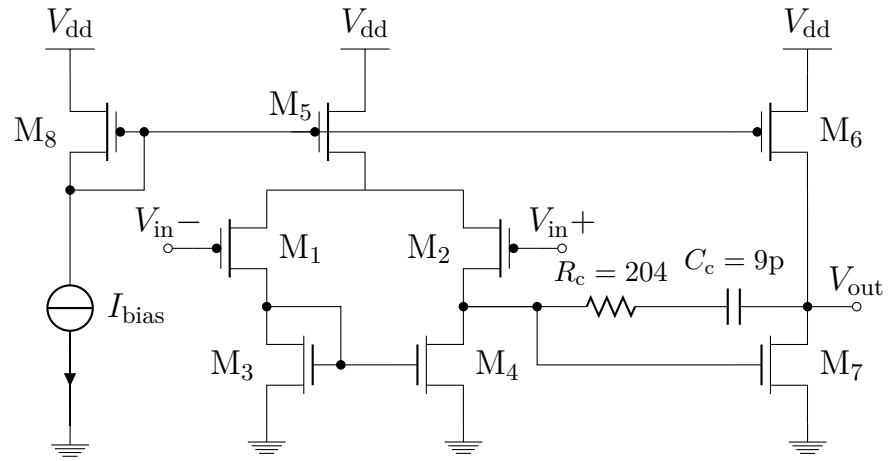


Figure 4.14: Two-stage differential Miller amplifier with compensation network.

Interestingly, the overall gain can be designed independently from the individual drain currents, which corresponds to the results obtained from the discussion in section 4.5. Note that each parameter marked by $(\cdot)^*$ is dependent on IC, I_D, L and t_{age} and has been simulated before.

The design should achieve the following target specifications: a total DC-Gain $|A_{Miller}|$ exceeding 100 dB, where $|A_I|$ is set to about 60 dB and $|A_{II}|$ to approximately 40 dB. The currents I_I and I_{II} may be set arbitrarily, if Eq. (4.32) is the only specification to follow. In order to have constraints of these currents, other specifications, such as the slew rate, need to be taken into account. In this study the current in the first stage is set to $I_I = I_{DS,1/2} = I_{DS,3/4} = 11\,\mu A$. The current mirror consisting of M_8 and M_5 is designed to supply a sufficient current of $I_{DS,5} = 22\,\mu A$. In order to ensure a high output impedance, the channel length of M_8 and M_5 is chosen to $L_{5/8} = 10.05\,\mu m$.

The input pair consisting of M_1 and M_2 defines the g_m of the first stage. High values of g_m/I_D provide a high power efficiency, which is achieved for inversion coefficients within the moderate inversion. The current source load constructed by M_3 and M_4 should provide a high output impedance, which is achieved with low valued of g_{ds}. Therefore, M_3 and M_4 are set in strong inversion and use rather long channel lengths of $L_{3/4} = 14.05\,\mu m$. A similar approach is chosen to design the second stage common source amplifier consisting of M_6 and M_7. Here, the current is set to $I_{II} = I_{ds,6/7} = 260\,\mu A$, which guarantees a good driving strength for capacitive and resistive output loads. The resulting device dimensions and bias currents are listed in Tab. A.6.

The GMID-Tool holds all information on the operating point-dependent degradation for each device, allowing native access to both fresh and aged transistor parameters. The simulation base for the degradation includes BTI and HCD for a simulated aging time of $t_{age} = 10\,yrs$. A comparison of GMID-Tool-approximated and SPICE-evaluated parameters of the Miller amplifier is given in Fig. 4.15 (a detailed list containing all degradation results is given in the appendix in Tab. A.4). Results for the g_m/I_D of the input transistors M_1 and M_2 is shown in Fig. 4.15a. Aside from the absolute values

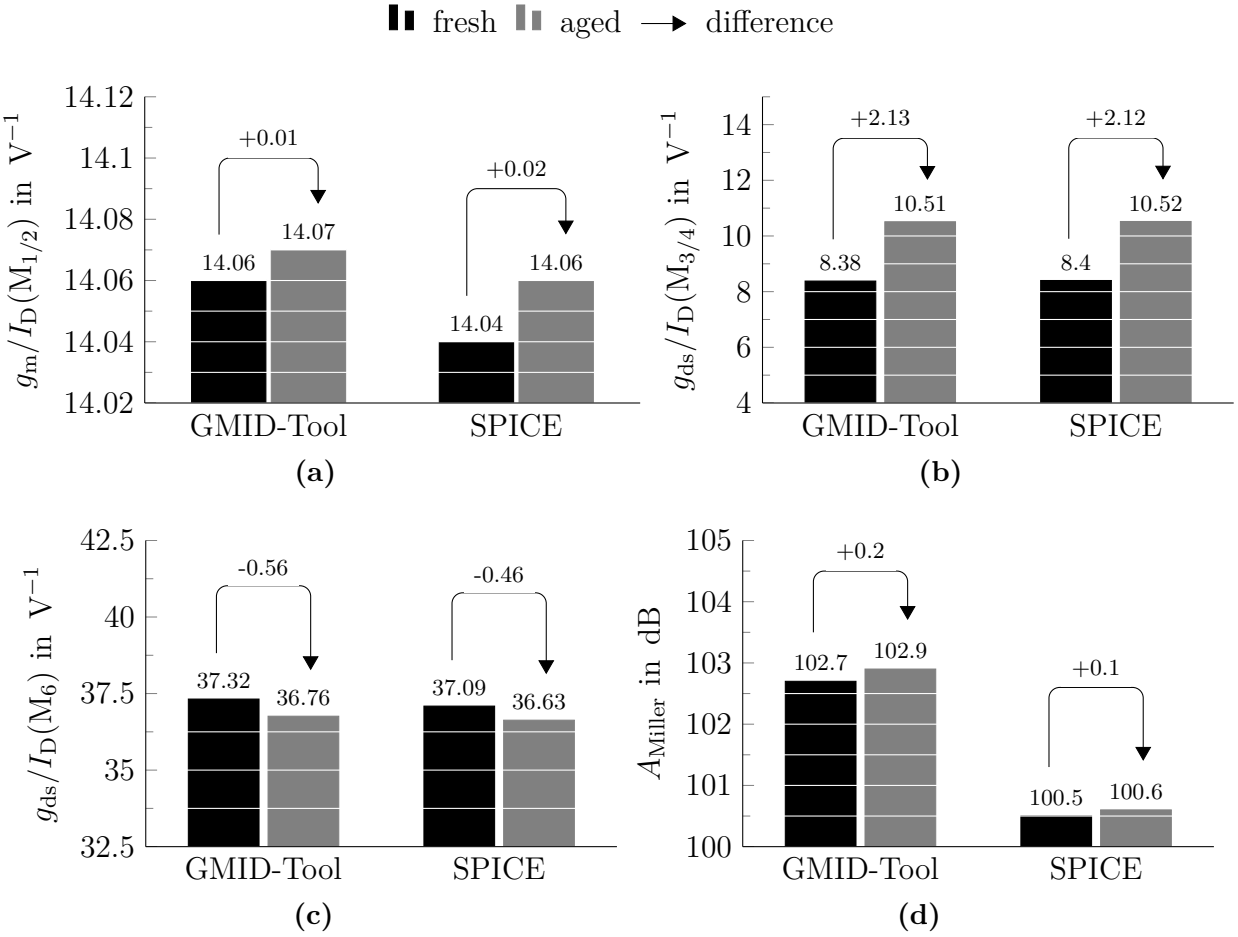


Figure 4.15: Comparison of results for the degradation of the Miller amplifier obtained through the GMID-Tool and SPICE with RelXpert. (a) g_m/I_D of $M_{1/2}$. (b) g_{ds}/I_D of $M_{3/4}$. (c) g_{ds}/I_D of M_6. (d) Total DC-Gain A_{Miller}.

approximated by the GMID-Tool and the ones evaluated with SPICE and RelXpert, also the difference in the fresh and aged state is depicted, which allows to compare the tendency of degradation as well. It can be observed, that the estimated g_m/I_D values from the GMID-Tool are well confirmed by the SPICE analysis. The same is observed for the output conductance efficiency g_{ds}/I_D of M_3 and M_4 in Fig. 4.15b, which the GMID-Tool almost perfectly predicts in regards of absolute value and tendency. Transistor M_6 represents the part of the output stage and contributes to A_{II} by its g_{ds}/I_D, which is depicted in Fig. 4.15c. The estimate of the difference in g_{ds}/I_D due to degradation ($-0.56\,V^{-1}$) is within reasonable range of the SPICE evaluated value ($-0.46\,V^{-1}$).
So far, only single transistor parameters have been compared. The total DC-Gain A_{Miller} consists of multiple transistor parameter and is therefore more exposed to changes in aging behavior of single transistors interacting with each other. The results of the degradation analysis for A_{Miller} are depicted 4.15d. In the fresh state the GMID-Tool over-estimated the DC-Gain by 2.2 dB. This difference is little, considering that only small signal parameters obtained from an operating point analysis of a single transistor

were used to calculate that value, whereas the SPICE solution relies on a full schematic representation of the circuit. Omitting this offset, the predicted difference in degradation is very close to the RelXpert solution.
The data obtained from the GMID-Tool shows a very good agreement with the SPICE and RelXpert solution, proving that

1. the g_m/I_D sizing method serves well for designing circuits in a fresh state.
2. the GMID-Tool and its inherent method of operating point-dependent degradation is suitable to provide well estimated values at a degraded state for single transistors, as well as composed characteristics of complex circuits.

4.8 Summary

In this chapter the operating point-dependent degradation is introduced and its application on transistors and circuits has been analyzed. Different MOS model types have been discussed and the differences in each type is investigated. From there the g_m/I_D sizing method is presented and the most important parameters, specifically the transconductance efficiency g_m/I_D and the inversion coefficient IC, are explained. Based on the operating point a degradation analysis using the RelXpert is performed. The impact on different transistor parameters is investigated, which included the drain current I_D, the threshold voltage V_{th} and the normalized conductances g_m/I_D and g_{ds}/I_D. A innovative EDA program, the GMID-Tool, is used to demonstrate the usage of an automated design approach for a common source and also a Miller amplifier. The analysis of the results show a well estimated degradation for single transistor parameters as well as for whole circuit characteristics. An advanced approach based on these results is presented in the following chapter.

5 Aging-Aware Sensitivity for Analog Circuit Design

In the previous two chapters methods and techniques have been presented, in order to determine the impact of degradation on analog integrated circuits. Already, these methods combine an accurate prediction of the aged behavior, an intuitive design flow and also efficient simulation schemes. In this chapter, the discussed work is extended to be used as an optimization tool.

Circuit optimization is widely used for integrated circuits. There exist several approaches on how to provide the optimum parameter set for a given design. The author in [38] uses an iterative optimization design flow approach, where degraded circuit performances values are optimized by modification of transistor dimensions based on an objective function. This function uses the so called worst case distance (WCD) as an indicator for how well the design is centered considering process variability. Instead of evaluating the WCD at a fresh state, a RelXpert simulation is performed to provide an aged representation of the design, which is used for WCD analysis. The WCD is minimized for the aged state by adjusting the transistor dimensions in the fresh circuit. This technique can be considered as an aged yield optimization and provides a functional methods for a reliability optimization. However, within this concept it is rather difficult to retrace the decisions on why dimensions of transistor have changed, leaving the designer out of control of the design. Another approach on optimization is proposed in [36] and [76]. Here, the authors suggest the use of knob and monitor structures, which tune the circuit in case of detected deviations. These concepts always involve additional circuitry and most likely increase the digital as well as the analog overhead.

The above discussed examples cover the impact of degradation on analog circuits, but lack of intuitive insight and identification of critical transistors. In this chapter an innovative method called aging-aware sensitivity (AAS) is proposed. As it will be shown, it is capable of optimizing circuits towards higher resilience against aging, while only relying on the concept of operating point-dependent degradation.

5.1 Sensitivity in Circuit Design

The term sensitivity is mathematically not strictly defined and is referred to as a measure, which indicates the influence of a change in one parameter on another. The use of a sensitivity analysis in circuit design has first been extensively reported in [119] and [88], where basic computation methods and concepts are described. There exist different methods on how to describe a sensitivity: the perturbation method [26], the adjoint-based sensitivity [35] and a direct approach [119]. The first two rely on the description of a topology τ. Through perturbation the parameters of the components of the topology deviate, which yields into to two different incarnations of the same topology. Both networks are then used to determine a sensitivity matrix. The sensitivity used in this work does not rely on a specific circuit topology, therefore the direct approach is used in the following.
Let $\mathcal{P}$ be a parameter which is influenced by a change in a variable h. The most elementary definition for the sensitivity is the derivative of the impacted parameter $\mathcal{P}$ with respect to h:

$$D_h^{\mathcal{P}} = \frac{\partial \mathcal{P}}{\partial h}. \tag{5.1}$$

Generally, the definition from Eq. (5.1) is suitable to describe the sensitivity of $\mathcal{P}$, however, there exists a major drawback of this description: the resulting figure is not unit-free. This property is important when different parameters need to be compared, especially in an electronic design automation environment.
A unit-free definition of a sensitivity is described in [119]. Here, the parameter $\mathcal{P}$ as well as the variable h are replaced by their logarithmic representations, which leads to the normalized sensitivity S_P^h

$$\mathsf{S}_{\mathcal{P}}^h = \frac{\partial \ln \mathcal{P}}{\partial \ln h} = \frac{h}{\mathcal{P}} \cdot \frac{\partial \mathcal{P}}{\partial h} = \frac{h}{\mathcal{P}} \cdot D_h^{\mathcal{P}}. \tag{5.2}$$

As $\mathsf{S}_{\mathcal{P}}^h$ in Eq. (5.2) is unit-free, these values may also be used and compared with sensitivities from different parameters, e.g. in an optimization algorithm.
Although the disadvantage of a unit-based definition from Eq. (5.1) has been overcome, the description from Eq. (5.2) is still not usable under all circumstances. If used in circuit design, $\mathcal{P}$ and h are related to circuit metrics, e.g. $\mathcal{P}$ being a resistance and h being the temperature. Then, $\mathsf{S}_{\mathcal{P}}^h$ is the sensitivity of a resistance with respect to temperature. It can be shown, that in some cases the definition in Eq. (5.2) does not provide a useful value. In a specific case, $\mathcal{P}$ may be a parasitic resistance, which its initial value is $\mathcal{P} = 0\,\Omega$ and is only activated through a change in temperature. In another case, a sensitivity might be evaluated at temperature value of $h = 0\,°C$. Both scenarios results in either $h = 0$ or $\mathcal{P} = 0$. In this case, the normalized sensitivity S_P^h fails and a different approach has to be applied. Therefore, also semi-normalized

approaches exist, which exchanges only one character by its logarithmic representation. The semi-normalized sensitivities for $\mathcal{P} = 0$ results to

$$\zeta_h^{\mathcal{P}} = \frac{\partial \mathcal{P}}{\partial \ln h} = h \cdot \frac{\partial \mathcal{P}}{\partial h} = h \cdot D_h^{\mathcal{P}} \tag{5.3}$$

and for $h = 0$ to

$$\xi_h^{\mathcal{P}} = \frac{\partial \ln \mathcal{P}}{\partial h} = \frac{1}{\mathcal{P}} \cdot \frac{\partial \mathcal{P}}{\partial h} = \frac{1}{\mathcal{P}} \cdot D_h^{\mathcal{P}}. \tag{5.4}$$

Given the example for $\mathcal{P}$ being a resistance and h being a temperature, it is also possible that a sensitivity value has to be evaluated for the case of both parameters being equal to zero, such that $h = 0$ and $\mathcal{P} = 0$. Here, the derivative from Eq. (5.1) is to be used.

5.2 Aging-Aware Sensitivity (AAS)

As stated in [53], sensitivities used for integrated circuit design reveal the interactions between parameter variations and their influence on system performance values and can also provide response gradients in optimization applications. Recent studies show, that sensitivities are productively used in combination with the $g_\mathrm{m}/I_\mathrm{D}$ sizing method. In [20] a sensitivity analysis is used to detect device size sensitivities for different circuit performance characteristics. This approach provides a semi-automatic sizing tool based on the EKV model. However, although some authors have updated the $g_\mathrm{m}/I_\mathrm{D}$ sizing method for recent technology nodes, e.g. in [89], the influence of aging on the sizing method is only little investigated. As it will be shown, the analysis of the operating point-dependent degradation of transistor parameters, as discussed in chapter 4, provides a basis for an aging-aware sensitivity analysis.

5.2.1 Definition of AAS

Using the results from section 5.1 it is possible to define a sensitivity value, which is a measure of degradation for a single transistor parameter. This sensitivity value uses the degradation time t_age as the varying parameter h, such that $h = t_\mathrm{age}$. However, in case of a fresh circuit, the time of degradation $t_\mathrm{age} = 0$. Therefore, the aging-aware sensitivity (AAS) is based on the semi-normalized sensitivity as defined in Eq. (5.4). The AAS of a transistor parameter $\mathcal{P} = \mathrm{d}_i$ is defined as

$$\xi_\mathrm{age}^{\mathrm{d}_i}\left(\mathbb{T}_\mathrm{age}\right) = \frac{\partial \ln \mathrm{d}_{i,\mathbb{T}_\mathrm{age}}}{\partial t_\mathrm{age}} = \frac{1}{\mathrm{d}_{i,\mathbb{T}_\mathrm{age}}} \cdot \frac{\partial \mathrm{d}_{i,\mathbb{T}_\mathrm{age}}}{\partial t_\mathrm{age}} \tag{5.5}$$

$$\left[\xi_\mathrm{age}^{\mathrm{d}_i}\left(\mathbb{T}_\mathrm{age}\right)\right] = \mathrm{yrs}^{-1}, \tag{5.6}$$

The term d_i is an operating point-dependent parameter and is therefore dependent on $\mathbb{T}_{\mathrm{age}} = \{\mathrm{IC}, L, I_{\mathrm{DS}}, t_{\mathrm{age}}\}$. Since this dependence is given for all instances of d_i, the declaration of $\mathbb{T}_{\mathrm{age}}$ is omitted and the AAS is only noted as $\xi_{\mathrm{age}}^{\mathrm{d}_i}$.

5.2.2 AAS of Composed Performance Values

In circuit design, not only single transistor parameters are of interest, but also circuit characteristics (such as gain), which are composed of different parameters of multiple devices. The AAS is suitable to describe the sensitivity of such a circuit characteristic, which relies on multiple transistor parameters. It can be shown, that the sensitivity of a circuit characteristic is only dependent on the AAS of single transistor parameters. Let $\mathcal{C}$ be an arbitrary circuit characteristic, which is constructed by multiplication and division of different transistor parameters, such that

$$\mathcal{C} = \frac{\prod_{n=1}^{N} \mathrm{d}_n}{\prod_{m=1}^{M} \mathrm{d}_m}, \tag{5.7}$$

with N and M being the number of considered operating point parameters. The associated aging-aware sensitivity is constructed by using the definition from Eq. (5.5)

$$\xi_{\mathrm{age}}^{\mathcal{C}} = \frac{\partial \ln \mathcal{C}}{\partial t_{\mathrm{age}}} = \frac{\partial \ln \left(\frac{\prod_{n=1}^{N} \mathrm{d}_n}{\prod_{m=1}^{M} \mathrm{d}_m} \right)}{\partial t_{\mathrm{age}}}. \tag{5.8}$$

The logarithmic part can be further split up. Using the relation $\ln(a/b) = \ln a - \ln b$ the equation above results to

$$\xi_{\mathrm{age}}^{\mathcal{C}} = \frac{\sum_{n=1}^{N} \ln \mathrm{d}_n - \sum_{n=1}^{M} \ln \mathrm{d}_m}{\partial t_{\mathrm{age}}}. \tag{5.9}$$

The ∂-operator can be drawn into each summation part, which yields

$$\xi_{\mathrm{age}}^{\mathcal{C}} = \sum_{n=1}^{N} \xi^{\mathrm{d}_n} - \sum_{n=1}^{M} \xi^{\mathrm{d}_m}. \tag{5.10}$$

The result from Eq. (5.10) shows that the aging-aware sensitivity $\xi_{\mathrm{age}}^{\mathcal{C}}$ can be represented by only using the sensitivities of the individual operating point-dependent transistor parameters $\xi_{\mathrm{age}}^{\mathrm{d}_i}$.
This is only true for composed characteristics that follow the construction rules set in Eq. (5.7) (multiplication and division only). In case of a characteristic which uses addition or subtraction, e.g.

$$\mathcal{D} = \sum_{i=1}^{I} \mathrm{d}_i, \tag{5.11}$$

a different result is observed. The aging-aware sensitivity of the characteristic $\mathcal{D}$ is then given by

$$\xi_{\text{age}}^{\mathcal{D}} = \frac{\partial \ln \sum_{i=1}^{I} \mathrm{d}_i}{\partial t_{\text{age}}}. \tag{5.12}$$

In contrast to a multiplication or division, differences and sums are not easily split within an logarithmic argument. Performing the derivation yields

$$\xi_{\text{age}}^{\mathcal{D}} = \frac{1}{\sum_{i=1}^{I} \mathrm{d}_i} \cdot \frac{\partial \sum_{i=1}^{I} \mathrm{d}_i}{\partial t_{\text{age}}} \tag{5.13}$$

$$= \sum_{i=1}^{I} \frac{1}{\frac{1}{\mathrm{d}_i} \sum_{u=1}^{I} \mathrm{d}_u} \cdot \underbrace{\frac{1}{\mathrm{d}_i} \frac{\partial \mathrm{d}_i}{\partial t_{\text{age}}}}_{\xi_{\text{age}}^{\mathrm{d}_i}}. \tag{5.14}$$

The last part of Eq. (5.14) already yields the AAS of each contributing parameter d_i. However, each $\xi_{\text{age}}^{\mathrm{d}_i}$ is accompanied by an additional correction term. The final result can be written as

$$\xi_{\text{age}}^{\mathcal{D}} = \sum_{i=1}^{I} \epsilon_{\mathrm{d}_i} \cdot \xi_{\text{age}}^{\mathrm{d}_i} \tag{5.15}$$

with

$$\epsilon_{\mathrm{d}_i} = \frac{1}{\frac{1}{\mathrm{d}_i} \sum_{u=1}^{I} \mathrm{d}_u}. \tag{5.16}$$

Note that the sum of all correction factors yields

$$\sum_{i}^{I} \epsilon_{\mathrm{d}_i} = 1. \tag{5.17}$$

The result from Eq. (5.15) is similar to the one obtained in Eq. (5.10), except that each individual AAS is weighted with a factor ϵ_{d_i}. The impact of ϵ_{d_i} strongly depends on the relation of each individual d parameter to each other. For example, given a characteristic constructed by $\mathrm{d}_1 + \mathrm{d}_2$ the according AAS results to

$$\xi_{\text{age}}^{\mathrm{d}_1+\mathrm{d}_2} = \underbrace{\frac{1}{1 + \frac{\mathrm{d}_2}{\mathrm{d}_1}}}_{\epsilon_{\mathrm{d}_1}} \cdot \xi_{\text{age}}^{\mathrm{d}_1} + \underbrace{\frac{1}{1 + \frac{\mathrm{d}_1}{\mathrm{d}_2}}}_{\epsilon_{\mathrm{d}_2}} \cdot \xi_{\text{age}}^{\mathrm{d}_2}. \tag{5.18}$$

The value of ϵ_{d_1} and ϵ_{d_1} in dependence on the ratio $\mathrm{d}_1/\mathrm{d}_2$ is shown in Fig. 5.1. In case of $\mathrm{d}_1 = \mathrm{d}_2$ both correction factors share a value of 0.5. Ratios of $\mathrm{d}_1/\mathrm{d}_2 \gg 1$ or $\mathrm{d}_1/\mathrm{d}_2 \ll 1$

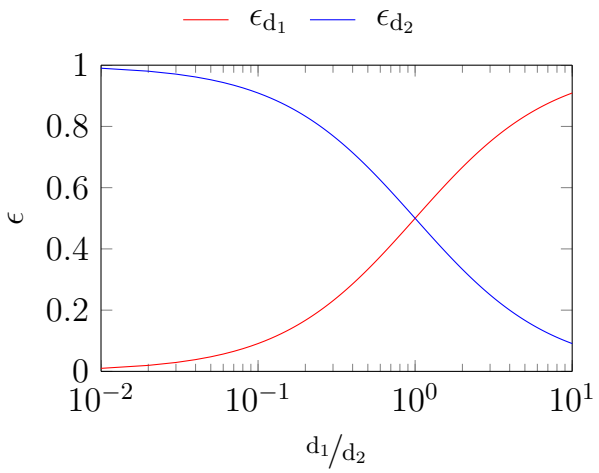


Figure 5.1: Value of ϵ_{d_1} and ϵ_{d_2} in dependence on the ratio of d_1/d_2.

lead to either $\epsilon_{d_1} = 0$ or $\epsilon_{d_2} = 0$ and therefore to a domination of one factor. This can be used to approximate $\xi_{age}^{d_1+d_2}$ as

$$\xi_{age}^{d_1+d_2} = \begin{cases} \xi_{age}^{d_1} & \text{for } d_1 \gg d_2 \\ \xi_{age}^{d_2} & \text{for } d_1 \ll d_2. \end{cases} \tag{5.19}$$

5.3 AAS for a Common Source Amplifier

In order to demonstrate the use of the aging-aware sensitivity, a common source amplifier is investigated with respect to certain output characteristics. The circuit consists of an NMOS input transistor and PMOS current source and no output load connected. The following study uses the AAS on two distinct characteristics: the DC-Gain and the unity gain frequency. The obtained AAS values are compared to actual simulations using the RelXpert simulator for several different designs. All simulations have been performed using a 150 nm technology node. If not stated differently, all transistors share a channel length of $L = 1\,\mu m$.

5.3.1 DC-Gain

A major performance value is the DC-Gain A_{cs}, which can be expressed through small signal parameters by

$$|A_{cs}| = \frac{gm, n}{gds, n + gds, p}, \tag{5.20}$$

where the index n and p denote a parameter associated with NMOS and PMOS respectively. This description can be transferred into the $g_\mathrm{m}/I_\mathrm{D}$ sizing methodology, which results to

$$|A_\mathrm{cs}| = \frac{\left(\frac{gm,n}{I_\mathrm{D}}\right)^* \cdot I_\mathrm{D}}{\left(\frac{gds,n}{I_\mathrm{D}}\right)^* \cdot I_\mathrm{D} + \left(\frac{gds,p}{I_\mathrm{D}}\right)^* \cdot I_\mathrm{D}} = \frac{\left(\frac{gm,n}{I_\mathrm{D}}\right)^*}{\left(\frac{gds,n}{I_\mathrm{D}}\right)^* + \left(\frac{gds,p}{I_\mathrm{D}}\right)^*}, \tag{5.21}$$

where I_D is the common drain current for both transistors. Using AAS, the tendency of degradation of the DC-Gain is obtained. Given the rules derived from Eq. (5.10) and Eq. (5.15), $\xi_\mathrm{age}^{|A_\mathrm{cs}|}$ results so

$$\xi_\mathrm{age}^{|A_\mathrm{cs}|} = \xi_\mathrm{age}^{\left(\frac{gm,n}{I_\mathrm{D}}\right)^*} - \left[\epsilon_{\left(\frac{\mathrm{gds,n}}{\mathrm{I_D}}\right)^*} \cdot \xi_\mathrm{age}^{\left(\frac{gds,n}{I_\mathrm{D}}\right)^*} + \epsilon_{\left(\frac{\mathrm{gds,p}}{\mathrm{I_D}}\right)^*} \cdot \xi_\mathrm{age}^{\left(\frac{gds,p}{I_\mathrm{D}}\right)^*}\right], \tag{5.22}$$

where the correction factors are given by

$$\epsilon_{\left(\frac{\mathrm{gds,n}}{\mathrm{I_D}}\right)^*} = \left[\frac{\left(\frac{gds,n}{I_\mathrm{D}}\right)^* + \left(\frac{gds,p}{I_\mathrm{D}}\right)^*}{\left(\frac{gds,n}{I_\mathrm{D}}\right)^*}\right]^{-1} = \left[1 + \frac{\left(\frac{gds,p}{I_\mathrm{D}}\right)^*}{\left(\frac{gds,n}{I_\mathrm{D}}\right)^*}\right]^{-1} \tag{5.23}$$

$$\epsilon_{\left(\frac{\mathrm{gds,p}}{\mathrm{I_D}}\right)^*} = \left[\frac{\left(\frac{gds,n}{I_\mathrm{D}}\right)^* + \left(\frac{gds,p}{I_\mathrm{D}}\right)^*}{\left(\frac{gds,p}{I_\mathrm{D}}\right)^*}\right]^{-1} = \left[1 + \frac{\left(\frac{gds,n}{I_\mathrm{D}}\right)^*}{\left(\frac{gds,p}{I_\mathrm{D}}\right)^*}\right]^{-1}. \tag{5.24}$$

Note that a further reduction of $\xi_\mathrm{age}^{\left(\frac{gm,n}{I_\mathrm{D}}\right)^*}$, $\xi_\mathrm{age}^{\left(\frac{gds,n}{I_\mathrm{D}}\right)^*}$ and $\xi_\mathrm{age}^{\left(\frac{gds,p}{I_\mathrm{D}}\right)^*}$ is not necessary, since the values marked by $(\cdot)^*$ are already available from the initial generation of $g_\mathrm{m}/I_\mathrm{D}$ metrics and can be regarded as single values. As shown in Eq. (5.19), the AAS can be approximated, if the correction factors are either close to 1 or 0. The value of $\epsilon_{(gds,n/\mathrm{I_D})^*}$ in dependence on the inversion coefficients $\mathrm{IC_n}$ and $\mathrm{IC_p}$ is shown in Fig. 5.2. It can be observed that $\epsilon_{(gds,n/\mathrm{I_D})^*}$ is not constant over the plane and also only sparsely approaches the limits of 1. The limit of $\epsilon_{(gds,n/\mathrm{I_D})^*} \to 0$ is not reached, as the overall minimum is only close to 0.3. Since from Eq. (5.17) it can be followed that $\epsilon_{(gds,p/\mathrm{I_D})^*} = 1 - \epsilon_{(gds,n/\mathrm{I_D})^*}$, the same observation holds for $\epsilon_{(gds,p/\mathrm{I_D})^*}$. Therefore no approximated computation of $\xi_\mathrm{age}^{|A_\mathrm{cs}|}$ is performed.

Observing Eq. (5.21) reveals, that the DC-Gain is only dependent on priorly simulated normalized transconductance values. These values are dependent on $\mathbb{T}_\mathrm{age}$, which consists of IC, L, I_DS and t_age. It has already been shown, that A_cs is independent of I_DS. In the case of no degradation at $t_\mathrm{age} = 0$ and a fixed L, these normalized values can be configured solely by the inversion coefficients $\mathrm{IC_n}$ and $\mathrm{IC_p}$ of the NMOS and PMOS. Therefore, the DC-Gain of the common source amplifier can be expressed as $A_\mathrm{cs}(\mathrm{IC_n}, \mathrm{IC_p})$. A visualization of A_cs in dependence on $\mathrm{IC_n}$ and $\mathrm{IC_p}$ is shown in Fig. 5.3a. In the same way, the aging-aware sensitivity can be constructed using only the inversion coefficients as input. The according AAS-map is shown in Fig. 5.3b.

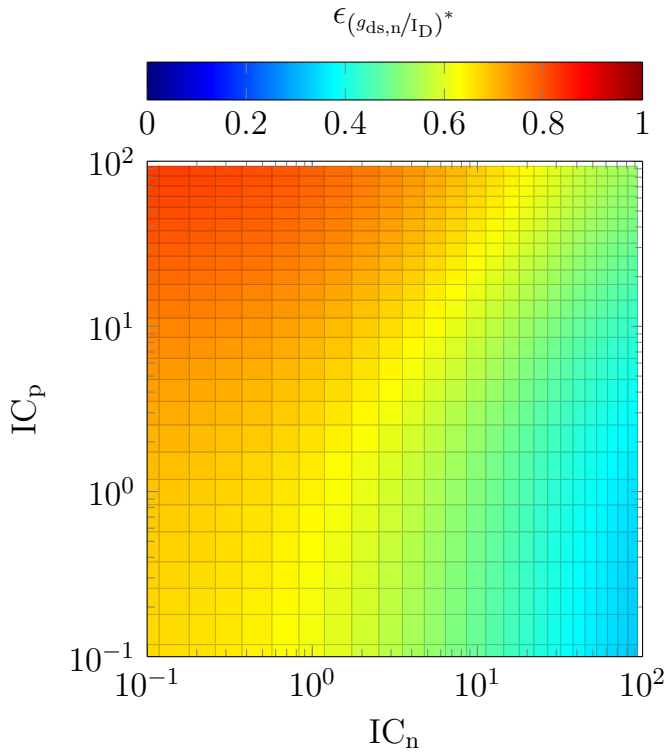


Figure 5.2: $\epsilon_{(gds,n/I_D)^*}$ in dependence on the inversion coefficients IC_n and IC_p.

The $|A_{cs}|$-map from Fig. 5.3a shows the DC-Gain in dB, which is targeted by a designer. If the PMOS current source transistor provides a high output impedance, then the amplifier acts similar to an intrinsic gain stage and therefore the DC-Gain is mostly influenced by a change in IC_n, which fits nicely with the results from the plot. Comparing this to the related AAS-Map in Fig. 5.3b reveals a different behavior. The computed sensitivity is not fully correlated to the DC-Gain. Whereas the DC-Gain is mostly independent of IC_p, $\xi_{age}^{|A_{cs}|}$ shows a dependence on both inversion coefficients IC_n and IC_p. This leads to the conclusion that designs with identically designed DC-Gain do not necessarily share the same sensitivity with respect to aging. Note that the results within the AAS-Map strongly depend on the used models within RelXpert. Different aging models will lead to dissimilar results, although the described method does not depend on the actual degradation model.

5.3.2 Unity-Gain-Frequency

The DC-Gain is important for static and low frequency analysis of analog circuits. Another important characteristic is the unity-gain frequency (UGF), which is equivalent to the gain bandwidth product (GBWP). The UGF is given by

$$f_{ugf} = \frac{g_{m,n}}{2\pi \cdot C_{out}}, \tag{5.25}$$

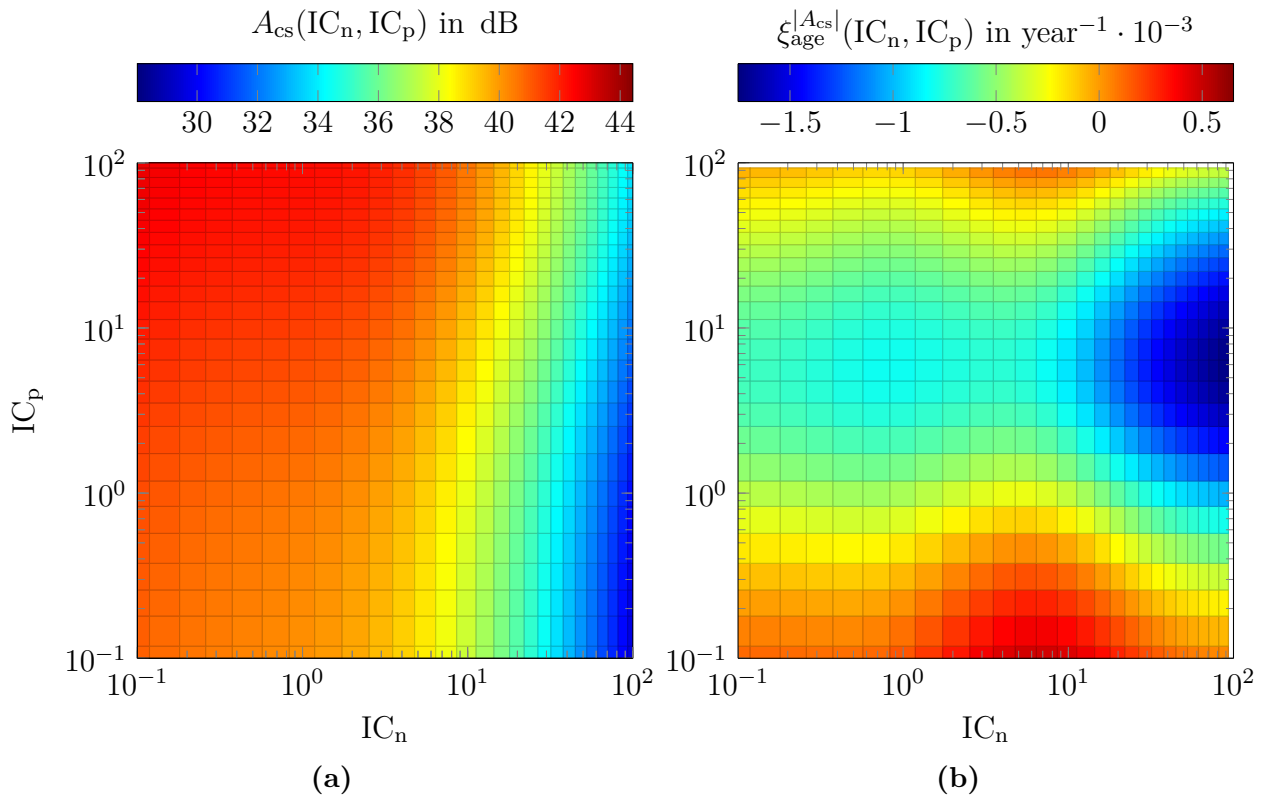


Figure 5.3: (a) DC-Gain-Map $|A_{cs}|(IC_n, IC_p)$ and (b) AAS-Map $\xi_{age}^{|A_{cs}|}(IC_n, IC_p)$ for a common source amplifier.

where C_{out} is the total capacitance present at the output node. Within the g_m/I_D sizing methodology, this expression is rearranged to

$$f_{ugf} = \frac{\left(\frac{g_{m,n}}{I_D}\right)^* \cdot I_D}{2\pi \cdot C_{out}}. \tag{5.26}$$

According to [96], the GBWP is the most important quality factor for any kind of amplifiers. However, it does not serve well as a figure of merit (FOM), since it depends on application specific design choices, such as the drain current and load capacitance. Therefore, a FOM is introduced, which is based on the GBWP but also remains independent on design choices. This FOM is defined as

$$FOM = \frac{\text{GBWP} \cdot C_{out}}{I_D} \tag{5.27}$$

$$[FOM] = \frac{\text{MHz pF}}{\text{mA}} \tag{5.28}$$

and quotes how much GBW can be obtained for a certain load capacitance and power consumption. A typical value for a single transistor lies in the range of 4500 $\frac{\text{MHz pF}}{\text{mA}}$

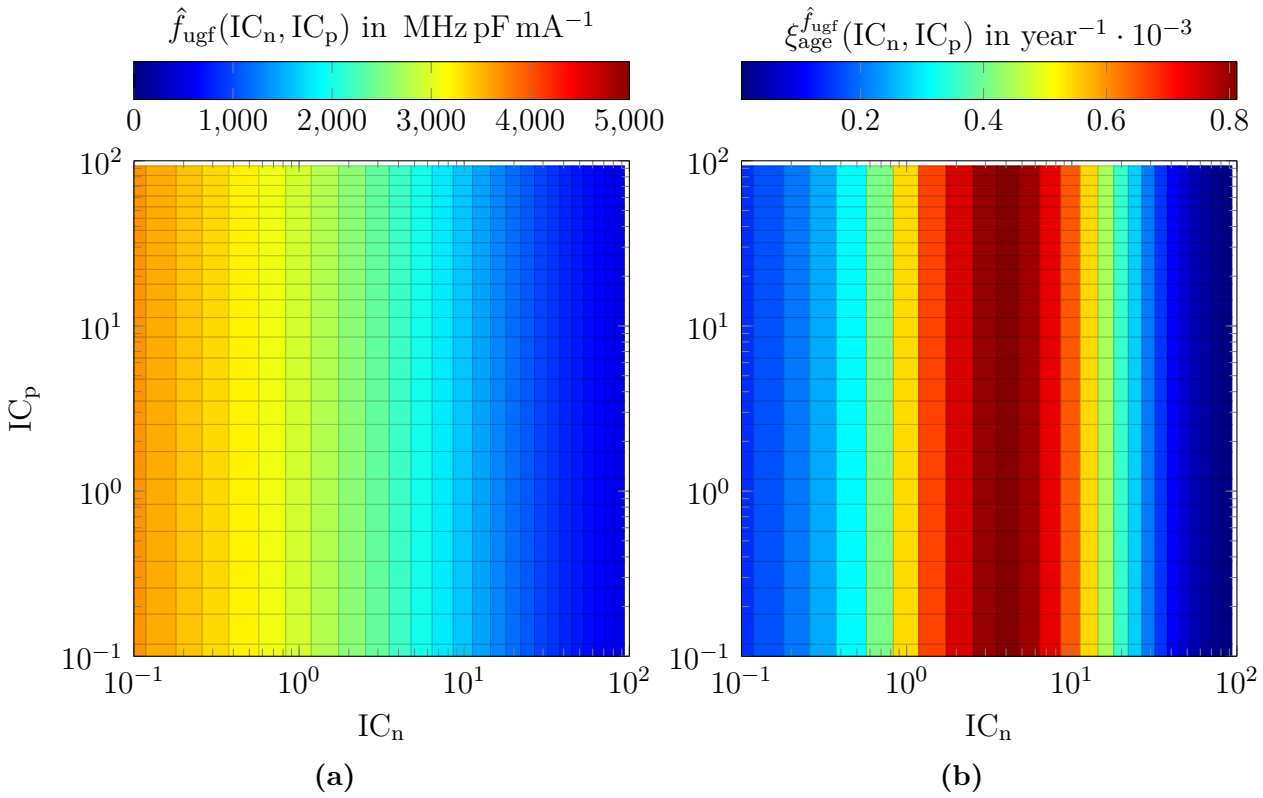


Figure 5.4: (a) The normalized unity-gain frequency-map $\hat{f}_{\mathrm{ugf}}(\mathrm{IC_n}, \mathrm{IC_p})$ and (b) AAS-Map $\xi_{\mathrm{age}}^{\hat{f}_{\mathrm{ugf}}}(\mathrm{IC_n}, \mathrm{IC_p})$ for a common source amplifier.

[96], dependent on the inversion coefficient. In the extend of this study this FOM is introduced as the normalized GBWP (or normalized unity-gain frequency) $\hat{f}_{\mathrm{ugf}}$

$$\hat{f}_{\mathrm{ugf}} = f_{\mathrm{ugf}} \cdot \frac{C_{\mathrm{L}}}{I_{\mathrm{D}}}. \tag{5.29}$$

The aging-aware sensitivity $\xi_{\mathrm{age}}^{\hat{f}_{\mathrm{ugf}}}$ is then given by

$$\xi_{\mathrm{age}}^{\hat{f}_{\mathrm{ugf}}} = \xi_{\mathrm{age}}^{\left(\frac{g_{\mathrm{m,n}}}{I_{\mathrm{D}}}\right)^*} - \xi_{\mathrm{age}}^{2\pi} = \xi_{\mathrm{age}}^{\left(\frac{g_{\mathrm{m,n}}}{I_{\mathrm{D}}}\right)^*}. \tag{5.30}$$

Since $\xi_{\mathrm{age}}^{\hat{f}_{\mathrm{ugf}}}$ is calculated without the use of addition or subtraction, a correction factor ϵ is not taken into account. The value of an AAS for a constant results to 0, therefore $\xi_{\mathrm{age}}^{2\pi}$ is not regarded in the final term. Comparing the expression for $\xi_{\mathrm{age}}^{|A_{\mathrm{cs}}|}$ in Eq. (5.22) and $\xi_{\mathrm{age}}^{\hat{f}_{\mathrm{ugf}}}$ in Eq. (5.30) shows, that both sensitivities are connected through $\xi_{\mathrm{age}}^{\left(\frac{g_{\mathrm{m,n}}}{I_{\mathrm{D}}}\right)^*}$. The AAS for the DC-Gain can written as

$$\xi_{\mathrm{age}}^{|A_{\mathrm{cs}}|} = \xi_{\mathrm{age}}^{\hat{f}_{\mathrm{ugf}}} - \left[\epsilon_{\left(\frac{\mathrm{gds,n}}{\mathrm{I_D}}\right)^*} \cdot \xi_{\mathrm{age}}^{\left(\frac{gds,n}{I_{\mathrm{D}}}\right)^*} + \epsilon_{\left(\frac{\mathrm{gds,p}}{\mathrm{I_D}}\right)^*} \cdot \xi_{\mathrm{age}}^{\left(\frac{gds,p}{I_{\mathrm{D}}}\right)^*}\right], \tag{5.31}$$

Table 5.1: Sample designs for a common source amplifier and corresponding sensitivities.

#	IC_n	IC_p	$\frac{W}{L}\big\vert_n$	$\frac{W}{L}\big\vert_p$	I_D	A_{fresh}	$A_{10\,yr}$	$\left\vert\frac{\Delta A}{A_{fresh}}\right\vert$	$\xi_{age}^{\vert A_{cs}\vert}$	$\hat{f}_{ugf,fresh}$	$\hat{f}_{ugf,10\,yr}$	$\left\vert\frac{\Delta \hat{f}_{ugf}}{\hat{f}_{ugf,fresh}}\right\vert$	$\xi_{age}^{\hat{f}_{ugf}}$
	–	–	–	–	µA	dB	dB	%	$\frac{1\cdot10^{-4}}{yr}$	$\frac{MHz\,pF}{mA}$	$\frac{MHz\,pF}{mA}$	%	$\frac{1\cdot10^{-4}}{yr}$
D_1	20	100	24.6	10.7	100	38.6	38.9	3.93	-1.35	1176	1192	1.38	3.42
D_2	10	30	35.6	49.1	100	39.6	39.1	3.21	-17.12	1522	1556	2.21	6.34
D_3	15	60	32.8	17.8	100	39.1	39.4	3.54	-10.42	1300	1324	1.78	4.64

which nicely reveals how certain characteristics interact with each other. This information is valuable for designers and helps to understand how degradation of one characteristic influences others.
Using the same constraints as for the DC-Gain, one can argue that the UGF and its associated AAS is entirely scalable through pairs of IC_n and IC_p. Both maps are shown in Fig. 5.4. Since the UGF is only dependent on the g_m/I_D of the input NMOS transistor, this value is ultimately only set through IC_n, which is shown in Fig. 5.4a. The AAS-map in Fig. 5.3b is also only dependent on IC_n and has its maximum around $IC_n \approx 4$, which corresponds to a moderate inversion. If IC_n is increased or decreased, then $\xi_{age}^{\hat{f}_{ugf}}$ minimizes until a a sensitivity value of almost $0\,yrs^{-1}$. This holds for an operation on strong inversion and also for very low values of IC_n in weak inversion. This observation is not seen for the fresh $\hat{f}_{ugf}$ within Fig. 5.4a, where the g_m/I_D is constantly increasing. As shown previously in the discussion for the DC-Gain, also here for the UGF the actual value is independent of its sensitivity.

5.3.3 Comparison of Designs

The generated AAS-maps from section 5.3.1 and 5.3.2 are now used in a sample design flow of a common source amplifier. It will be shown that these maps serve well to optimize the response to degradation for the given circuit. The design shall consider two target specifications, the DC-Gain and the normalized unity-gain frequency. The influence of HCD and NBTI is considered and the impact of degradation is simulated using the RelXpert simulator. These results are compared with the AAS-maps.
The study consists of four different design approaches of a common source amplifier. Each sample is designed to provide a DC-Gain of approximately 40 dB and should exceed a normalized unity-gain frequency of $\hat{f}_{ugf} > 1000\,MHz\,pF\,mA^{-1}$. The final widths of the transistors is found using the definition of the inversion coefficient IC, which is rearranged to

$$W = \frac{I_D \cdot L}{2n_0\mu_0 C_{OX} U_T^2 \cdot IC}. \tag{5.32}$$

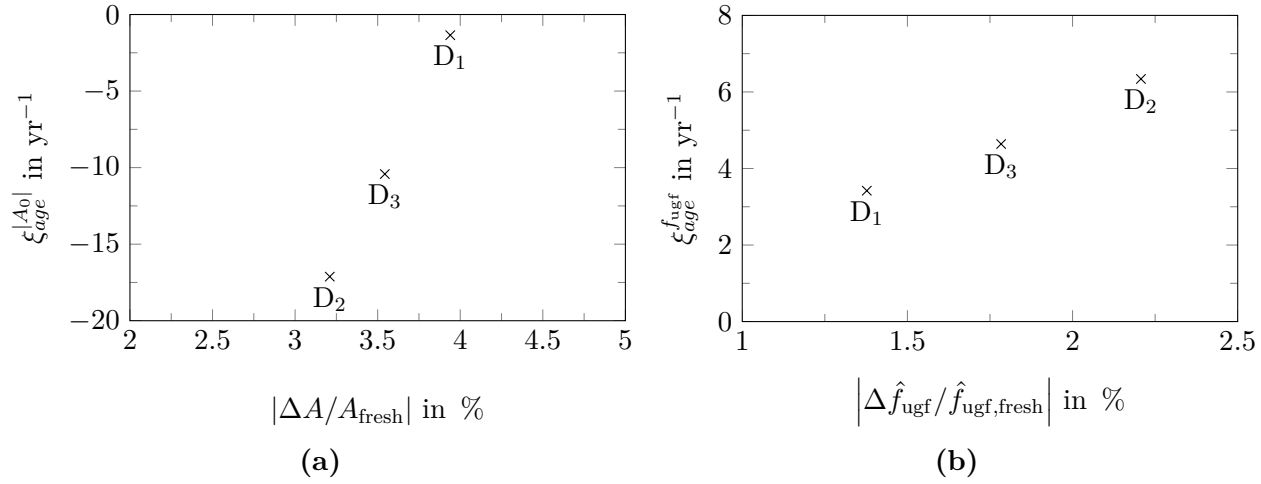


Figure 5.5: Aging-Aware sensitivities and corresponding degradation given in percentage. (a) DC-Gain A_{cs}. (b) unity-gain-frequency $\hat{f}_{\text{ugf}}$.

The drain current is fixed to a value of $I_{\text{D}} = 100\,\mu\text{A}$ and the channel length is held constant at $L = 1\,\mu\text{m}$. Therefore, the final channel widths are only dependent on the inversion coefficients IC_{n} and IC_{p}. These two coefficients are also used to design for a specific target specification. The data from Fig. 5.3a allows a selection of suitable pairs of IC_{n} and IC_{p}, which result in the desired gain of 40 dB. These pairs have to be checked with the data from Fig. 5.4a in order to fulfill the second target for $\hat{f}_{\text{ugf}}$. Using this methods, three different designs D_1 to D_3 have been chosen. The parameters of each design are shown in Tab. 5.1. The values computed using AAS and the corresponding degradation simulated through RelXpert are compared in Fig. 5.5a and Fig. 5.5b for the DC-Gain and the normalized unity-gain frequency respectively. Both figures relate the AAS to the relative simulated degradation.
The results from Tab. 5.1 and Fig. 5.5a show a reduction in degradation of the DC-Gain of about 1 % comparing D_1 to D_2 and D_2 D_3. The data from Fig. 5.5b however reveals the exact opposite behavior. The least degradation of the UGF is achieved in D_1. Each AAS value should be tied to a certain degradation, which is nicely mapped by the designs D_1, D_2 and D_3.

5.4 Validity of Aging-Aware Sensitivity Maps

In the previous section three different designs were used to compare the results for the AAS and the actual deviations simulated using SPICE and RelXpert. The question remains, if the AAS is able to compute feasible results under all conditions or if there exists a boundary, which limits the applicability. Note that AAS maps are based on the concept of operating point-dependent degradation, which is an approximative method,

since circuit degradation is determined by single transistor degradation only. The following shall demonstrate a validity analysis for the DC-Gain.
The drain current of a transistor I_D is dependent on internal parameters p_i and also on the terminal voltages V_S, V_B, V_G and V_D

$$I_\mathrm{D} = f_k(p_i, \dots, V_i, \dots), \tag{5.33}$$

where V_i represents a list containing all terminal voltages. A change in I_D induced by an impacting parameter ρ is given by

$$\frac{\partial I_\mathrm{D}}{\partial \rho} = \frac{\partial f_\mathrm{k}}{\partial \rho} + \sum_{j=1}^{N} \frac{\partial f_\mathrm{k}}{\partial V_j} \frac{\partial V_j}{\partial \rho}, \tag{5.34}$$

where $\frac{\partial f_\mathrm{k}}{\partial \rho}$ is the derivative of f_k with respect to the impacting parameter. However, a change in ρ may not only affect f_k, but also the terminal voltages, which in turn serve as input for f_k. The dependence of the terminal voltages on the impacting factors is considered by $\sum_{j=1}^{N} \frac{\partial f_\mathrm{k}}{\partial V_j} \frac{\partial V_j}{\partial \rho}$. The term $\frac{\partial f_\mathrm{k}}{\partial V_j}$ represents small signal conductances, such as g_m, and $\frac{\partial V_j}{\partial \rho}$ can be regarded as the sensitivity of the terminal voltages with respect to ρ. In terms of aging simulations, the impacting factor is the degradation time t_age. Further, if Eq. (5.34) is normalized to I_D, the general expression for the semi-normalized sensitivity is obtained and results to

$$\frac{1}{I_\mathrm{D}} \frac{\partial I_\mathrm{D}}{\partial t_\mathrm{age}} = \frac{1}{I_\mathrm{D}} \frac{\partial f_\mathrm{k}}{\partial t_\mathrm{age}} + \frac{1}{I_\mathrm{D}} \sum_{j=1}^{N} \frac{\partial f_\mathrm{k}}{\partial V_j} \frac{\partial V_j}{\partial t_\mathrm{age}}. \tag{5.35}$$

The AAS is therefore constructed using two terms. The first term is the direct derivative of the drain current, which is induced through the aging process. The second term also takes a change in terminal voltages into account, which is induced through the aging of the same and also of other components in the circuit. In order to calculate a precise solution of I_D at t_age, the terminal voltages need to be evaluated at that specific point in time. This cannot be achieved without the use of a SPICE solver, which for example is used by RelXpert. An AAS is calculated from results of the testbench in Fig. 4.3. Here, the terminal voltages do not change during an aging simulation, since the testbench configuration forces given node potentials. Therefore, $\frac{\partial V_j}{\partial t_\mathrm{age}}$ results to zero. Consequently, the AAS used in this work, is calculated by the first term only

$$\xi_\mathrm{age}^{I_\mathrm{D}} = \frac{1}{I_\mathrm{D}} \frac{\partial f_\mathrm{k}}{\partial t_\mathrm{age}}, \tag{5.36}$$

which is less accurate but removes the dependency on a SPICE solver. Although the AAS is constructed using this approximative method, it still provides a good measure for the sensitivity with respect to aging. The validity is shown using a simple approach.

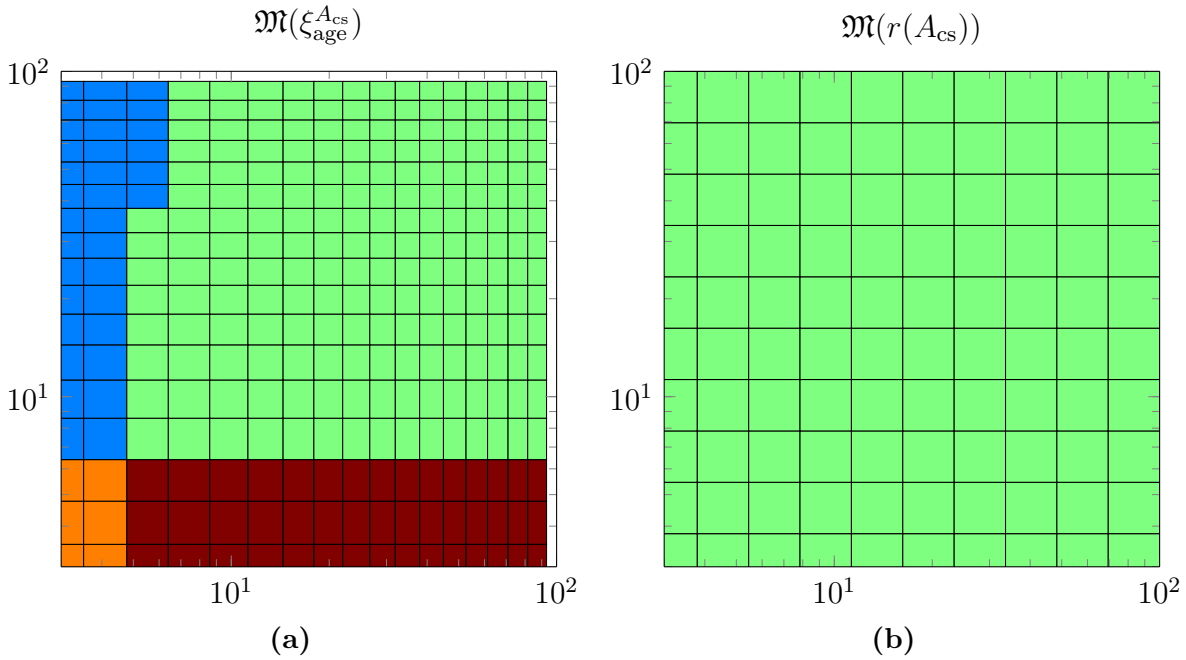


Figure 5.6: Monotonicity quadrant plots for the DC-Gain of a common source amplifier A_{cs}. (a) AAS-based results. (b) RelXpert-based results.

Let $\xi_{\text{age}}^{\alpha}$ be an AAS, which is only dependent on one inversion coefficient IC_o. It can be assumed, that due to its approximative nature the AAS will not provide a SPICE accurate result. However, it can still be used as a guideline in circuit design, if the following condition is met. The AAS should always provide a sensitivity value, which is proportional to the degradation, which is given from the RelXpert result. This is true, if both, $\xi_{\text{age}}^{\alpha}$ and the RelXpert solution $r(\alpha)$, have proportional slopes, which implies that both values rise and fall at the same IC_o. In a mathematical sense, this condition is met as long as the monotonicity of both values is identical for given intervals of IC_o. Let $[a, b]$ an interval, where a function $f(x)$ is monotone then $\mathfrak{M}_a^b(f(x))$ shall be a function, which reveals the type of monotonicity of $f(x)$ in $a \leq x \leq b$, such that

$$\mathfrak{M}_a^b(f(x)) = \begin{cases} + & \text{for monotonically increasing} \\ - & \text{for monotonically decreasing,} \end{cases} \tag{5.37}$$

then an AAS is valid with respect to the RelXpert solution, if

$$\mathfrak{M}_{\text{IC}_{\text{o,min}}}^{\text{IC}_{\text{o,max}}}(\xi_{\text{age}}^{\alpha}) = \mathfrak{M}_{\text{IC}_{\text{o,min}}}^{\text{IC}_{\text{o,max}}}(r(\alpha)). \tag{5.38}$$

In a two dimensional approach with e.g. $\xi^{\beta}_{\mathrm{age}}(\mathrm{IC_n}, \mathrm{IC_p})$, monotonicity has to be separately evaluated in the direction of $\mathrm{IC_n}$ and $\mathrm{IC_p}$, leading to four distinct solutions

$$\left(\mathfrak{M}^{\mathrm{IC_{n,max}}}_{\mathrm{IC_{n,min}}}(\beta), \mathfrak{M}^{\mathrm{IC_{n,max}}}_{\mathrm{IC_{n,min}}}(\beta)\right) := \begin{cases} \mathrm{Q_1} & \text{for } (+,+) \\ \mathrm{Q_2} & \text{for } (-,+) \\ \mathrm{Q_3} & \text{for } (-,-) \\ \mathrm{Q_4} & \text{for } (+,-) \end{cases}. \tag{5.39}$$

The validity expression from Eq. (5.38) still holds. An example of such a monotonicity analysis is shown in Fig. 5.6 for the DC-Gain of a common source amplifier. The different colored regions each correspond to a monotonicity of $\mathrm{Q_1}$ to $\mathrm{Q_4}$. It can be observed, that for low values of $\mathrm{IC_n}$ and $\mathrm{IC_p}$, there exist some difference in the solutions gained from the AAS in Fig. 5.6a and the RelXpert equivalent in Fig. 5.6b. However, for $\mathrm{IC_n} > 6$ and $\mathrm{IC_p} > 7$, both reveal the same result. Therefore, the AAS can be used in almost all cases, starting from the moderate inversion and covering the full strong inversion regime.

5.5 Optimum Aging-Aware Operating Point

So far the AAS is used to detect, if certain designs of a circuit are less susceptible to aging than others. However, this information can also be used to find the optimum conditions, at which the circuit degrades the least. The following study shows, how the concept of aging-aware sensitivity is used to provide a method, for determining the optimum operating point of a transistor in a circuit [48]. Since in most cases multiple characteristics define the behavior of a circuit, the concept of AAS has to be extended to be used in a multi-performance environment.

If multiple specifications are to characterize the behavior of a circuit, each of them has a unique AAS associated with it. Furthermore, each characteristic may be more important to the designer than others, e.g. the input referred noise is less important than the power consumption. Still, in order to optimize a design with respect to its sensitivity against aging, each AAS needs to be taken into account. Instead of using single AAS values for each characteristic, it is possible to combine each AAS to a multi-composed one, which represents the sensitivity of the circuit, rather than of an individual characteristics. The multi-composite AAS shall be defined as the sum of the absolute value for each sensitivity

$$\xi^{\tau}_{\mathrm{age}} = \sum_{i=1}^{n} \alpha_i \cdot \left|\xi^{\mathrm{p}_i}_{\mathrm{age}}\right|, \tag{5.40}$$

with $\sum_{i=1}^{n} \alpha_i = 1$. The factors α_i are used to weight each AAS according to its priority. The expression in Eq. (5.40) is an objective function, which can be used to find an optimum operating point for transistors in a circuit. The minimum of this function

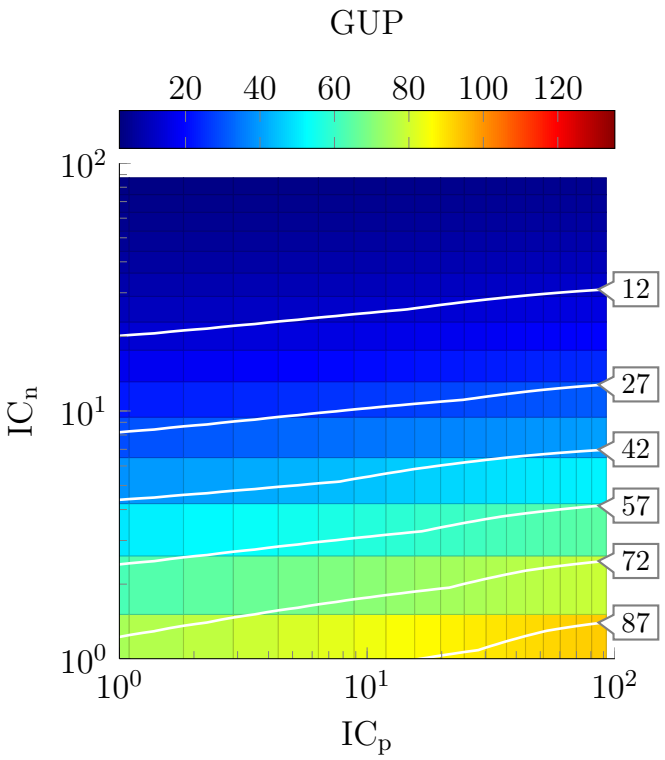


Figure 5.7: Fresh GUP in dependence on IC_p and IC_n. Contour lines are labeled with corresponding value.

leads to inversion coefficients, which will provide the least sensitivity with respect to aging. The results can be influenced by the weighting factors.
A composite AAS shall be built using the DC-Gain and normalized UGF as defined in Eq. (5.21) and Eq. (5.29) respectively. In order to provide a common design space, the gain-unity-gain product (GUP) is introduced, which is defined as

$$\mathrm{GUP} = \underbrace{\left| \frac{\left(\frac{gm,p}{I_{\mathrm{D,p}}}\right)^*}{\left(\frac{gds,p}{I_{\mathrm{D,p}}}\right)^* + \left(\frac{gds,n}{I_{\mathrm{D,n}}}\right)^*} \right|}_{\mathrm{p1}=A_{\mathrm{cs}}} \cdot \underbrace{\frac{\left(\frac{gm,p}{I_{\mathrm{D,p}}}\right)^*}{2 \cdot \pi}}_{\mathrm{p1}=\hat{f}_{\mathrm{ugf}}} \,. \tag{5.41}$$

Again, the GUP is dependent in $\mathbb{T}_{\mathrm{age}}$ of the transistors $\mathrm{M_n}$ and $\mathrm{M_p}$. The fresh GUP in dependence on the inversion coefficients IC_n and IC_p is shown in Fig. 5.7. Constant values of GUP are indicated by white contour lines. Each line is labeled with its associated value, so each pair of IC_n and IC_p resulting in the same contour line leads to the identical GUP. These lines shall provide boundaries, which will be used in further optimization routines. Optimum operating points can be found by the minimum of the total aging-aware sensitivity, such that

$$\min\left\{\xi^{\tau}_{\mathrm{age}}\right\} \Rightarrow \mathrm{d}_{\mathbb{T}\mathrm{OP,opt}}(\mathrm{M_n}),\ \mathrm{d}_{\mathbb{T}\mathrm{OP,opt}}(\mathrm{M_p}). \tag{5.42}$$

has to be fulfilled. The expression of $\xi^{\tau}_{\mathrm{age}}$ is gained through the description in Eq. (5.40) and the according specifications, namely $\xi^{\mathrm{p1}}_{\mathrm{age}} = \xi^{\mathrm{A_{cs}}}_{\mathrm{age}}$ and $\xi^{\mathrm{p2}}_{\mathrm{age}} = \xi^{\hat{f}_{\mathrm{ugf}}}_{\mathrm{age}}$, which results to

$$\xi^{\tau}_{\mathrm{age}} = \alpha_1 \cdot \left|\xi^{|A_{\mathrm{cs}}|}_{\mathrm{age}}\right| + \alpha_2 \cdot \left|\xi^{\hat{f}_{\mathrm{ugf}}}_{\mathrm{age}}\right| . \tag{5.43}$$

This description of $\xi^{\tau}_{\mathrm{age}}$ is able to map $\xi^{|A_{\mathrm{cs}}|}_{\mathrm{age}}$ and $\xi^{\hat{f}_{\mathrm{ugf}}}_{\mathrm{age}}$ onto a single value, which provides a measure for the age-dependent deviation of characteristics for a common source amplifier circuit. The number of incorporating characteristics is not limited, so the expression in Eq. (5.43) may be expanded to fit any number of different specifications. The value of $\xi^{\tau}_{\mathrm{age}}$ is not only dependent on $\mathrm{IC_p}$ and $\mathrm{IC_n}$, but also on the weighting factors α_1 and α_2. These weighting factors extend the degree of freedom within the analog design process. A visualization of $\xi^{\tau}_{\mathrm{age}}$ is given in Fig. 5.8 with different weighting factors. A configuration of the weighting factors of $(\alpha_1, \alpha_2) = (0.5, 0.5)$ may be considered as a neutral influence of each influencing characteristic, whereas the configuration $(\alpha_1, \alpha_2) = (0.2, 0.8)$ prioritizes $\hat{f}_{\mathrm{ugf}}$ over A_{cs}. The contour lines are depicted as a visual reference and are identical to the ones shown in Fig. 5.7. A detailed representation of the results for both weighting factor configurations can be found in the appendix in Tab. A.7 and Tab. A.8.
The minimum of $\xi^{\tau}_{\mathrm{age}}$ for each area, which is limited by the contour lines, and each found result corresponds to a pair of $\mathrm{IC_n}$ and $\mathrm{IC_p}$, which fulfills the specifications given by these limits. As discussed in section 5.4, this method only provides reasonable results for inversion coefficients greater than a certain minimum. Therefore the minimum found is constraint to this $\mathrm{IC_{min}}$. In most cases, the optimum found for $\mathrm{IC_p}$ is at the onset of deep strong inversion. Therefore, the energy-efficient moderate regime is only available to the NMOS. The choice of a different set of weighting factors influences the optimum result in one case, which is more precisely shown in Fig. 5.8a and Fig. 5.8b.

5.6 Summary

In this the section the concept of the aging-aware sensitivity is presented. Different sensitivity schemes used in integrated circuit design have been discussed and evaluated in regards of the use in an aging-aware environment. Based on a common source amplifier the AAS revealed how the DC-Gain and the normalized unity-gain frequency are sensitive with respect to NBTI and HCD degradation mechanisms. A validity analysis shows the applicability of the AAS and also investigates limits caused by its approximative degradation estimation. The AAS is used to optimize a given circuit incorporating two different design characteristics by finding an optimum aging-aware operating point for transistors.

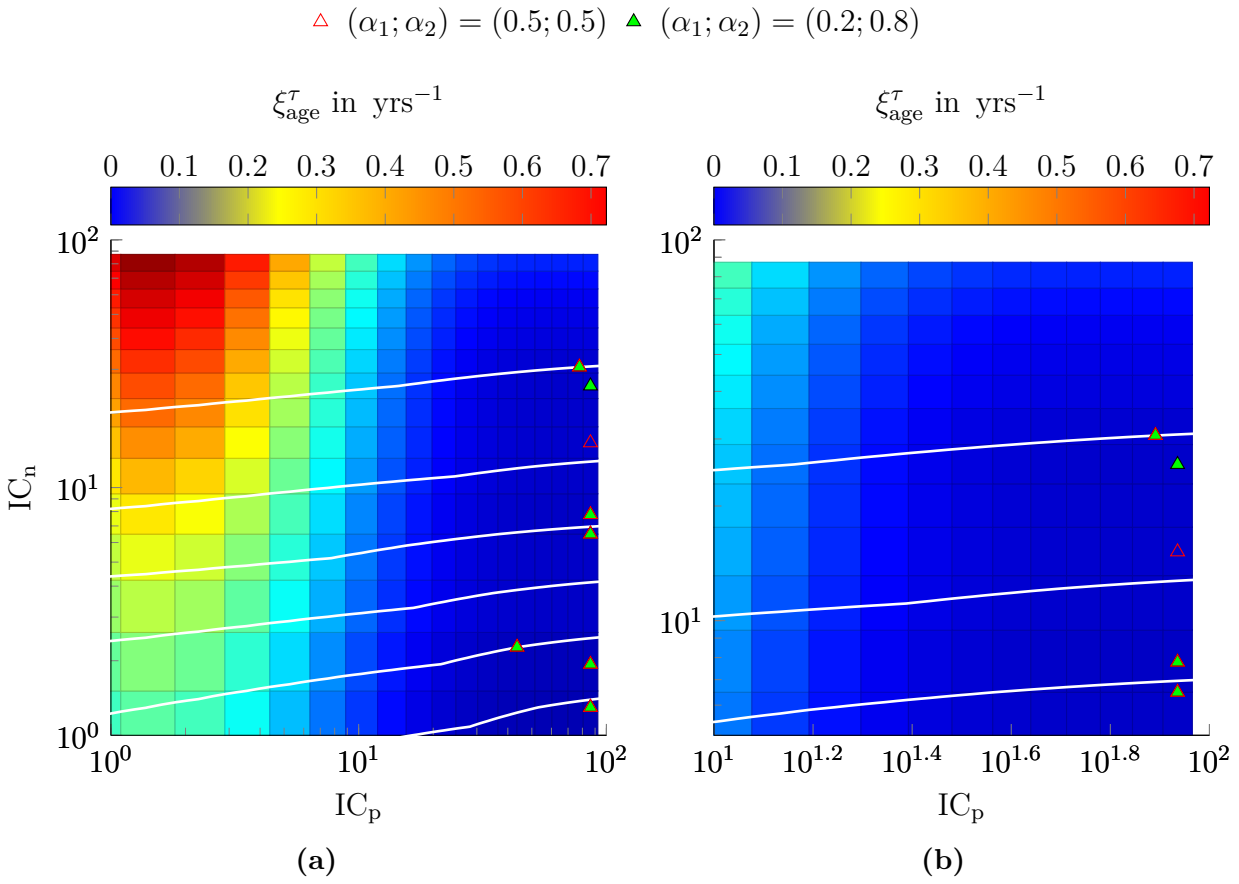


Figure 5.8: Multi-composite AAS of $\xi^{|A_{\text{cs}}|}_{\text{age}}$ and $\xi^{\hat{f}_{\text{ugf}}}_{\text{age}}$. Contour lines indicate values for GUP as depicted in Fig. 5.7. Markers show optimized designs with lower bound of $\text{IC}_{\text{min}} = 1$ for different configurations of $(\alpha_1;\alpha 2)$. Figure (b) shows resized area of (a).

6 Conclusion and Outlook

6.1 Conclusion

In this thesis different aging-aware design methods for the use with analog integrated circuits have been developed. These methods partly exploit the structural difference in certain circuit topologies and also build upon the idea of an operating point-dependent degradation of transistor parameters. Additionally, the GMID-Tool has been developed, which provides intuitive access to g_m/I_D parameters including extended parameters gained from the analysis of this work. A major contribution is also the concept of an aging-aware sensitivity (AAS) and its visualized representation (AAS-maps), which provide a useful guidance for minimizing age-dependent deviations in circuit characteristics at design time. Finally, a simulation flow has been presented, which allows the generation of the reliability function using an altered Monte Carlo analysis.
The basic idea of the structural degradation analysis for analog circuits lies in the fact, that most circuits rely on biasing circuitry. These provide stable voltages, which are converted through current mirrors into stable bias currents. Different biasing circuits, including a single transistor, a MOS only reference and a beta matching reference have been investigated. Each of these circuits biased different mirror circuits, which included a single transistor, a cascode mirror and also a wide swing structure. The results show, that for certain combinations, the deviation in the output current due to degradation can be reduced by more than 14 %. Interestingly, the proposed method is also able to be used in amplifier designs. A folded cascode amplifier has been modified in its bias and mirror circuitry, adapting the results of the previous study. This decreased the deviations in various characteristics (comparing best and worst configuration), such as for the DC-Gain (~10 % gain), the gain bandwidth (~15 % gain) and the common mode rejection ration (~10 % gain).
An important contribution of this thesis is the operating point-dependent degradation. This approach is based upon the g_m/I_D sizing methodology, which uses the state of inversion of the conducting channel, the so called inversion coefficient IC, as an indicator for the operating point of the device. While generating specific g_m/I_D related metrics within a dedicated testbench, a reliability simulator hooks into the simulation scheme and degrades the device. The resulting small signal g_m/I_D metrics are then available in

a fresh and also in an aged state, while the metrics themselves are mostly dependent on IC. Relating the degradation of each device in dependence on IC reveals unseen correlations of transistor parameters and their response to aging. It is shown, that the results can be used to optimize the aging behavior of a common source amplifier.
The results from the operating point-dependent degradation have been used to develop an innovative method, which can be used to design analog circuits with respect to its sensitivity to aging at a very early design stage. This approach, the aging-aware sensitivity, uses analytic description of performance characteristics for any kind of circuits, and is able to provide well approximated estimate on the deviation due to aging mechanisms. AAS is based on the definition of a semi-normalized sensitivity, while the underlying data is provided through operating point-dependent degradation. It has been shown, that the AAS concept allows a computation for any output characteristic, as long as it belongs to a linear system and is described by an analytic function. Also, the AAS of a circuit characteristics can always be arranged to be only dependent on single transistofsr sensitivities. This property makes the AAS a versatile method in analog circuits design, as single transistor AAS values only have to computed once and are then valid for any design, just like the underlying g_m/I_D methodology.
The AAS has been calculated for several circuit topologies and has been visualized through so called AAS-maps. These maps revealed interesting results, such that the actual value of performance is not necessarily coherent with the expected degradation. This is extremely useful, considering that this also holds for different designs, which yield identical output characteristics, but differ in the transistor operating point. A validity analysis has shown that the AAS is coherent with post-schematic reliability simulation results for a wide range of different inversion coefficients.
This thesis has used the results from the AAS scheme to not only calculate the estimated deviation, but also to optimize circuits towards higher reliability through optimum operating points. The optimum is found through the minimum of the AAS value and simulation results show a real benefit from this solution.
An increase in reliability needs to be evaluated, which is by generating the according reliability function. This is achieved by using a batch of statistically generated circuit samples and also an aging simulation. Each of the generated sampled needs to be aged and evaluated in terms of deviations in the output characteristics. Since aging simulations use a lot of simulation time, a modified approach has been proposed in this thesis. The variability-aware gradual aging (VAGA) simulation scheme uses the fact, that for moderate technology nodes a shift in the ditribution of process parameter is not expected. Therefore, modifying the simulation routine reduces the number of aging simulations tremendously (only one for each gradual aging step). VAGA has been used to verify the results for the structural degradation analysis.
In summary, the methods developed in this thesis for an aging-aware design of analog integrated circuits prove to be innovative and powerful. The methods have shown to be compatible with existing design methodologies and extend them naturally. With

these contributions, reliability and aging aspects of analog circuits is no longer only regarded at the end of a design flow, but is shifted to the very beginning of it, while still delivering accurate estimates of performance deviations. The results were benchmarked against existing reliability simulators and showed to be very competitive, while at the time the aging simulation effort is reduced to a minimum.

6.2 Outlook

The operating point-dependent degradation is only evaluated for the center of process variation, which in terms of process corners is the nominal corner. Since the reliability function can only be generated using statistically distributed samples, the AAS needs to be expanded to support the influence of process variability. The underlying g_m/I_D methodology would have to be altered to be used in a statistical context. A first approach could be made by interchanging each parameter by its statistical distribution (e.g. Gaussian). Each individual distribution is then used to construct statistical circuit characteristics. This would not only allow an estimate on expected deviations of circuit characteristics, but could provide an estimate for the yield and even more importantly the reliability at a given point in time.
This thesis has used two different technology nodes for all circuit related simulations (350 nm and 150 nm). These nodes have been used in the industry for years, yielding well defined process parameters, which is why these nodes are often used in safety critical environments. However, smaller technology nodes have come up, which have not only shrunk the minimal channel length, but also changed the fabrication steps of a traditional transistor (FinFET [58]). The operating point-dependent degradation as well as the aging-aware sensitivity will have to be evaluated against modern nodes. At his point, technology nodes up to 65 nm should perform well with the developed methods, as the general fabrication process has not changed until this node. The 45 nm node introduced HKMG transistors, which is prone to PBTI [128] and lets the NMOS suffer from two degradation modes (PBTI and HCD). Since these effects are excited under different stress scenarios, the extraction process for AAS parameters would need to be adopted. Also, since the FinFET technology allows ultra short channel lengths, quantization effects of the channel dopants lead to quantized BTI modes, which is known as random telegraph noise (RTN) [34]. This new degradation mode has to be checked against the current implementation of the AAS.

A Appendix

A.1 Transistor Dimensions for Structural Degradation Analysis using Current Mirrors and Bias Circuits

Within Tab. A.1 the chosen transistor dimensions for the structural degradation analysis are shown. Default width and length for all transistors is given as NMOS and PMOS. Any transistor not following the default values is specifically named. Dimensions given are valid for the folded cascode amplifier, the mos only reference, the beta matching reference and the cascode current mirrors in sink and source configuration.

Table A.1: Dimensions for all NMOS and PMOS transistors.

Device	W	L	Units	Device	W	L	Units
	Default Values				Folded Cascode OTA		
NMOS	5	1	µm	MC	12.5	5	µm
PMOS	12.5	1	µm	MF	5	5	µm
I_{bias}	12		µA	M9, M10	50	1	µm
	Mos Only Reference				Beta Matching Reference		
M_N	3	49	µm	S1	10	60	µm
M_P	5	23	µm	S3	5	0.5	µm
				R	12k		Ω
	Cascode as Source				Cascode as Sink		
M3	5	4	µm	M3	12.5	4	µm

A.2 Results for the Structural Degradation Analysis of a Folded Cascode Amplifier for different Bias Circuit and Reference Circuit Combination

The results for the degradation analysis of a folded cascode amplifier using different combinations of bias circuits and current mirror structures are presented in Tab. A.2 and Tab. A.3. Data from Tab. A.2 shows the degradation of amplifier characteristics, which is observed by choosing a single PMOS current mirror. The bias voltage is either generated by a beta matching reference or a mos only reference. The same characteristics are also given in Tab. A.2, which presents the observed degradation generated using a wide swing current mirror. Again, data is shown for a beta matching reference as well as for a mos only reference.

Table A.2: Degraded Circuit Characteristics for Single PMOS Bias

Specifications	200d	400d	600d	800d	1000d
Single PMOS biased by BMR					
Δ PM	0.89	1.24	1.41	1.61	1.84
Δ Slew Rate Rise	-19.55	-26.74	-31.97	-36.19	-39.81
Δ Slew Rate Fall	-55.55	-70.57	-79.67	-86.00	-90.55
Δ DC Gain	0.64	-1.52	-4.42	-8.19	-13.43
Δ GBW	-40.62	-57.50	-69.99	-79.57	-87.14
Δ PSRR	-50.07	-55.49	-59.92	-64.35	-69.51
Δ CMRR	3.83	6.71	8.91	10.58	10.32
Single PMOS biased by MOR					
Δ PM	0.77	1.14	1.30	1.40	1.57
Δ Slew Rate Rise	-11.27	-15.46	-18.48	-20.98	-23.07
Δ Slew Rate Fall	-48.97	-62.94	-71.71	-78.04	-82.95
Δ DC Gain	1.41	0.09	-1.65	-3.76	-6.36
Δ GBW	-33.33	-49.96	-59.05	-68.39	-75.58
Δ PSRR	-53.86	-59.86	-64.36	-68.35	-72.35
Δ CMRR	2.06	3.77	5.17	6.24	5.15

Table A.3: Degraded Circuit Characteristics for Wide Swing Bias

Specifications	200d	400d	600d	800d	1000d
Wide Swing biased by BMR					
Δ PM	0.81	1.25	1.47	1.61	1.86
Δ Slew Rate Rise	-19.55	-26.74	-31.97	-36.19	-39.81
Δ Slew Rate Fall	-58.75	-75.02	-84.48	-90.89	-95.38
Δ DC Gain	0.38	-2.03	-5.37	-9.93	-16.81
Δ GBW	-41.14	-58.75	-71.63	-81.76	-90.18
Δ PSRR	-46.32	-52.22	-57.13	-62.26	-68.78
Δ CMRR	5.83	9.90	12.69	14.85	15.13
Wide Swing biased by MOR					
Δ PM	0.72	1.14	1.31	1.40	1.57
Δ Slew Rate Rise	-11.38	-15.56	-18.52	-21.02	-23.05
Δ Slew Rate Fall	-52.10	-67.49	-76.90	-83.56	-88.57
Δ DC Gain	1.18	-0.32	-2.32	-4.78	-7.85
Δ GBW	-33.24	-50.18	-60.90	-69.92	-78.18
Δ PSRR	-51.05	-57.74	-62.78	-67.36	-72.01
Δ CMRR	2.81	5.05	6.83	8.27	7.68

A.3 Results of the Miller-Opamp Design

The results for the degradation analysis using the GMID-Tool for the Miller-Opamp design is shown in Tab. A.4. Data is shown for several transistors, including the input transistors and also all devices, which are part of the investigated analytical expressions of the circuit characteristics. Each transistor parameter is given as a result from the GMID-Tool as well es from SPICE solvers. In case of aged parameter values, RelXpert has been used as the SPICE equivalent reliability simulator. The deviation observed between the fresh and aged state is given for the GMID-Tool and also for the SPICE/RelXpert solution.

Table A.4: Results for the Miller-Opamp design and corresponding degradation with RelXpert and GMID-Tool.

	$M_{1/2}$			$M_{3/4}$			M_6			M_7				
Parameter	$\frac{g_m}{I_{DS}}$	g_m	g_{ds}	$\frac{g_m}{I_{DS}}$	g_m	g_{ds}	$\frac{g_m}{I_{DS}}$	g_m	g_{ds}	$\frac{g_m}{I_{DS}}$	g_m	g_{ds}	$\lvert A_{diff,I} \rvert$	$\lvert A_{diff} \rvert$
unit	-	µS	nS	-	mS	nS	-	mS	µS	-	mS	µS	dB	dB
GMIDTool (fresh)	14.06	161.8	45.0	9.56	109.9	96.4	6.99	1.80	9.61	10.89	2.8	14.3	61.2	102.7
GMIDTool (aged)	14.07	161.7	44.9	9.61	110.5	96.6	6.99	1.80	9.55	11.29	2.9	14.6	61.2	102.9
Δ in %	0.07	0.06	0.00	0.52	0.54	0.31	0.00	0.00	0.63	3.54	3.45	1.85	0.02	0.79
SPICE (fresh)	14.04	160.8	43.1	9.58	109.9	120.3	6.92	1.79	9.51	10.05	2.6	14.5	59.8	100.5
RelXpert (aged)	14.06	159.9	42.8	9.61	109.2	119.5	6.96	1.79	9.42	10.09	2.6	14.4	59.8	100.6
Δ in %	0.14	0.56	0.70	0.31	0.46	0.67	0.57	0.00	0.96	0.40	0.78	0.62	0.00	0.52

A.4 Degraded System Characteristics (VAGA)

Data presented in Tab. A.5 shows the results from a variability-aware gradual aging analysis taken on a folded cascode amplifier. Each characteristic is shown in terms of its mean value as well as standard deviation. Simulation has been performed for an aging time of 10 yrs and a step size of 2 yrs.

Table A.5: Degraded System Characteristics.

Specification	fresh	2 yrs	4 yr	6 yr	8 yr	10 yr
Mean Values	t_0	t_1	t_2	t_3	t_4	t_5
PM (°)	89.37	89.50	89.57	89.63	89.68	89.73
SR_{rise} ($^{MV}/_{s}$)	1.90	1.31	1.07	0.91	0.79	0.70
SR_{fall} ($^{MV}/_{s}$)	2.79	1.88	1.53	1.29	1.11	0.97
Gain ($^{V}/_{V}$)	66.33	66.04	65.76	65.48	65.18	64.87
GBW (MHz)	4.14	3.06	2.59	2.25	1.98	1.76
PSRR (dB)	73.89	71.94	70.78	69.91	69.13	68.44
Standard Deviation						
PM (°)	0.072	0.081	0.088	0.091	0.097	0.102
SR_{rise} ($^{MV}/_{s}$)	0.262	0.221	0.200	0.184	0.170	0.158
SR_{fall} ($^{MV}/_{s}$)	0.402	0.340	0.308	0.282	0.260	0.240
Gain ($^{V}/_{V}$)	0.311	0.451	0.561	0.668	0.773	0.876
GBW (MHz)	0.410	0.438	0.413	0.407	0.390	0.369
PSRR (dB)	10.06	10.21	9.72	9.67	9.75	9.92

A.5 Dimensions for Miller Amplifier

Tab. A.6 shown the channel dimensions, inversion coefficient and drain current for each transistor within the Miller amplifier design.

Table A.6: Design parameters for the Miller Amplifier.

Device	IC	I_D	W	L_0
$M_{1/2}$	3.5	11 µA	304	10.05
$M_{3/4}$	9.6	11 µA	60	14.05
M_5	21.2	22 µA	100	10.05
M_6	16.4	260 µA	150	1.05
M_7	7.5	260 µA	64	0.55
M_8	21.2	11 µA	50	10.05

A.6 Optimum Operating Points from AAS Analysis

Tab. A.7 and Tab. A.8 show results for the optimum operating points for a common source amplifier for different weighting factors.

Table A.7: Aging-Aware Inversion Coefficients for given GUP ($\alpha_1 = 0.5$, $\alpha_2 = 0.5$).

#	GUP low	GUP high	IC_n	IC_p	GUP	IC_n	IC_p	GUP
			$IC_{min} = 0$			IC_{min} =1e-1		
1	0	12	5.7e-2	2.4e+1	9.6	2.2e+1	2.7e+1	12.4
2	12	27	5.7e-2	7.8e+0	26.3	2.1e+1	1.4e+1	22.6
3	27	42	2.1e+1	7.8e+0	35.3	2.1e+1	7.8e+0	35.3
4	42	57	6.6e-2	2.3e+0	55.4	1.7e+1	3.3e+0	57.5
5	57	72	6.6e-2	1.3e+0	67.9	2.0e+1	1.9e+0	72.2
6	72	87	8.5e-2	7.1e-1	79.1	1.9e+1	1.1e+0	85.7
7	87	102	1.0e-1	5.3e-2	101.0	1.0e-1	1.0e-1	97.6
8	102	117	2.1e-3	3.9e-3	109.2	1.6e+1	1.6e-1	110.5

Table A.8: Aging-Aware Inversion Coefficients for given GUP ($\alpha_1 = 0.2$, $\alpha_2 = 0.8$).

#	GUP low	GUP high	IC_n	IC_p	GUP	IC_n	IC_p	GUP
			$IC_{min} = 0$			IC_{min} =1e-1		
1	0	12	5.7e-2	2.4e+1	9.6	2.2e+1	3.4e+1	10.0
2	12	27	5.7e-2	7.8e+0	26.3	2.1e+1	1.4e+1	22.6
3	27	42	8.5e-2	4.7e+0	37.9	2.1e+1	7.8e+0	35.3
4	42	57	8.5e-2	2.6e+0	52.5	1.7e+1	3.3e+0	57.5
5	57	72	6.6e-2	1.3e+0	67.9	2.0e+1	1.9e+0	72.2
6	72	87	8.5e-2	7.1e-1	79.1	1.9e+1	1.1e+0	85.7
7	87	102	1.0e-1	5.3e-2	101.0	1.8e+1	7.1e-1	95.0
8	102	117	2.1e-3	3.9e-3	109.2	1.6e+1	1.6e-1	110.5

A List of Abbreviations

AAS	Aging-aware sensitivity
ADC	Analog to digital converter
ADE	Analog design environment (from Cadence)
ANP	Aged netlist parser
AP	Action potential
BERT	The Berkeley reliability simulator
BMR	Beta matching reference
BSIM	Berkeley short-channel IGFET model
BTI	Bias temperature instability
CHC	Channel hot carrier
CHE	Channel hot electron
CMOS	Complementary metal oxide semiconductor
CMRR	Common mode rejection ratio
DAHC	Drain Avalanche Hot Carrier
EDA	Electronic design aid
EKV	Enz-Krummenacher-Vittoz
EM	Electromigration
EMI	Electromagnetic interference
ENOB	Effective number of bits
FEM	Finite element method
FOM	Figure of merit
GBW	Gain-bandwidth
GBWP	Gain-bandwidth product
GUP	Gain-unity-gain product
HCD	Hot carrier degradation
HCI	Hot carrier injection
HKMG	High-k metal gates
IGFET	Insulated-gate field-effect transistor

LDD	Layout dependence effect
LER	Line edge roughness
LFP	Local field potential
LNA	Low noise amplifier
LSB	Least significant bit
LWR	Line width roughness
MGG	Metal gate granularity
MOR	MOS only reference
MOS	Metal oxide semiconductor
MTTF	Mean time to failure
NBTI	Negative bias temperature instability
NMOS	N-channel CMOS transistor
NMS	Neural measurement system
OTA	Operational transconductance amplifier
PBTI	Positive bias temperature instability
PIP	Polysilicon-insulator-polysilicon (capacitor)
PM	Phase margin
PMOS	P-channel CMOS transistor
PSRR	Power supply rejection ratio
PVT	Process-voltage-temperature (variation)
PVTA	Process-voltage-temperature-aging (variation)
RDF	Random dopant fluctuation
SEU	Single event upset
SGHE	Secondary Generated Hot Electron
SHE	Substrate Hot Electron
SHH	Substrate hot hole
SILC	Stress-induced leakage current
SPICE	Simulation program with integrated circuit emphasis
SR	Slew rate
SRF	Slew rate falling edge
SRR	Slew rate rising edge
TDDB	Time-dependent dielectric breakdown
UGF	Unit-gain frequency
VAGA	Variability-aware gradual aging
WCD	Worst-case distance

A List of Symbols

Symbol	Description	Unit
$(\cdot)^*$	Parameter value from $g_\mathrm{m}/I_\mathrm{D}$ testbench	
A_cs	Gain of common source amplifier	dB
A_LNA	Gain of low noise amplifier	dB
A_Miller	Gain of Miller amplifier	dB
A_P	Area parameter in Pelgrom's model	$\mathrm{V}\cdot\sqrt{\mu\mathrm{m}^2}$
β	Gain factor	$\mathrm{A\,V}^{-1}$
C_f	Filter capacitance	F
c_i	i-th sample of a topology τ	
C_in	Input capacitance	F
C_OX	Oxide capacitance per area	$\mathrm{fF}/\,\mu\mathrm{m}^2$
$\mathbf{c}_\tau$	Vector containing m samples of a topology τ	
$\mathbf{d}_\mathbb{T}^k$	Vector containing all transistor parameters of a specific transistor k	
d_i	i-th transistpor parameter	
ΔD	Generation of interface traps	cm^{-2}
$D_h^\mathcal{P}$	Primitive sensitivity of $\mathcal{P}$ with respect to h	$[\mathcal{P}]/[h]$
ΔV_th	Mismatch in threshold voltage	V
E_a	Activation energy	$\mathrm{J}\cdot\mathrm{mol}^{-1}$
$\xi_\mathrm{age}^{\mathrm{d}_i}$	AAS of transistor paramater d_i	$1/\,\mathrm{yrs}$
ϵ_{d_i}	Correction factor for AAS of d_i	
E_m	Maximum channel electrical field	$\mathrm{V}/\,\mathrm{m}$
ϵ_OX	Relative oxide permitivity	
f_ugf	Unity-gain frequency	MHz
$\hat{f}_\mathrm{ugf}$	Normalized unity-gain frequency	$\frac{\mathrm{MHz\,pF}}{\mathrm{mA}}$
g_ds	Output transconductance	
g_m	Transdoncductance	S
$g_\mathrm{m}/I_\mathrm{D}$	Transconductance efficiency	$1/\,\mathrm{V}$

I_0	Constant specific current for inversion coefficient	A
$I_\square$	Sheet current or current per square	$\mathrm{A}/\,\mu\mathrm{m}^2$
IC	Inversion coefficient	
I_D	Drain current	A
$I_\mathrm{D,min}$	Minimum drain current	A
I_DS	Drain to source current	A
I_G	Gate current	A
I_GB	Gate to bulk current	A
I_GD	Gate to drain current	A
I_GS	Gate to source current	A
I_spec	Specific current for inversion coefficient	A
J	Current densitiy	$\mathrm{A}/\,\mu\mathrm{m}^2$
k	Shape parameter in Weibull function	
k	Boltzmann constant	J/K
κ	Constant ratio of a current mirror	
L	Channel length	µm
λ	Scale parameter in Weibull function	
l_c	Characteristic length of the saturation region	µm
L_eff	Effective channel length	µm
$\mathrm{M_i}$	Name of transistor i	
μ	Mobility of channel	$\mathrm{m^2V^{-1}s^{-1}}$
μ_0	Low field mobility of channel	$\mathrm{m^2V^{-1}s^{-1}}$
N_0	Number of initially existing Si–H bonds	cm^{-2}
N_H	Concentration of hydrogen at the interface	cm^{-2}
N_it	Number of interface traps	cm^{-2}
$\mathbf{p}(c_j)$	Vector of n performance value	
p_i	i-th performanc value	
q	Charge of an electron	$\approx -1.6 \cdot 10^{-19}\mathrm{C}$
$R_\tau(t)$	Reliability function of a topology τ	
R	Resistance	Ω
S_f	Shape factor	
$\mathbf{s_p}$	Set of n specification limits	
S_P	Variation paramater in Pelgrom's model	$\mathrm{V}/\,\mathrm{m}$
$\mathsf{S}_\mathcal{P}^h$	Normalized sensitivity	
$\zeta_h^\mathcal{P}$	Semi-normalized sensitivity	$[\mathcal{P}]$
$\xi_h^\mathcal{P}$	Semi-normalized sensitivity	$1/[h]$
T	Temperature	°K
t_age	Aging time	yrs

$\mathbf{t}_{\text{age}}$	Vector of aging points in time	
$\mathbb{T}_{\text{age}}$	Age-dependent operating point	
τ	Circuit topology (e.g. schematic)	
t_{bd}	Time to breakdown	s
$\mathbb{T}_{\text{g}_\text{m}/\text{I}_\text{D}}$	g_m/I_D operating point tuple	
t_{OX}	Gate oxide thickness	nm
t_{sim}	Transient simulation time	s
$\mathbb{T}_{\text{quadratic}}$	Quadratic operating point tuple	
$V_{\text{AP,max}}$	Amplitude of an action potential	µV
V_{bias}	Bias voltage	V
V_{D}	Drain Voltage	V
V_{dd}	Supply voltage	V
V_{DS}	Drain to source voltage	V
$V_{\text{DS,sat}}$	Saturation voltage	V
V_{G}	Gate voltage	V
V_{GS}	Gate to source voltage	V
V_{in}	Input Voltage	V
V_{LSB}	Voltage of a least significant bit	V
V_{-}	Negative input voltage	V
V_{o}	Output voltage	V
$V_{\text{o,min}}$	Minimum output voltage	V
V_{out}	Output voltage	V
V_{ov}	Overdrive voltage	V
V_{+}	Positive input voltage	V
V_{ref}	Reference voltage	V
V_{S}	Source voltage	V
v_{T}	Temperature voltage kT/q	mV
V_{th}	Threshold voltage	V
$V_{\text{th,n}}$	Threshold voltage NMOS	V
$V_{\text{th,p}}$	Threshold voltage PMOS	V
W	Channel width	µm
W_{eff}	Effective channel width	µm
$\frac{W}{L}\big\vert_{\text{i}}$	Ratio of channel width and length of transistor i	
ω_{T}	Gain bandwidth	MHz
Y_τ	Yield of a topology τ	

A Bibliography

[1] M. V. Aarset. How to identify a bathtub hazard rate. *IEEE Transactions on Reliability*, 36(1):106–108, 1987.

[2] K. Agarwal and S. Nassif. Characterizing process variation in nanometer CMOS. In *Design Automation Conference, 2007. DAC '07. 44th ACM/IEEE*, pages 396–399, June 2007.

[3] M. Alam. A critical examination of the mechanics of dynamic NBTI for PMOSFETs. In *Electron Devices Meeting, 2003. IEDM '03 Technical Digest. IEEE International*, pages 14.4.1–14.4.4, Dec 2003.

[4] M. Alam, H. Kufluoglu, D. Varghese, and S. Mahapatra. A comprehensive model for PMOS NBTI degradation: Recent progress. *Microelectronics Reliability*, 47(6):853 – 862, 2007. Modelling the Negative Bias Temperature Instability.

[5] M. Alam and S. Mahapatra. A comprehensive model of PMOS NBTI degradation. *Microelectronics Reliability*, 45(1):71 – 81, 2005.

[6] D. Ang and S. Wang. Recovery of the NBTI-stressed ultrathin gate p-MOSFET: The role of deep-level hole traps. *Electron Device Letters, IEEE*, 27(11):914–916, Nov 2006.

[7] F. Arnaud, L. Pinzelli, C. Gallon, M. Rafik, P. Mora, and F. Boeuf. Challenges and opportunity in performance, variability and reliability in sub-45 nm CMOS technologies. *Microelectronics Reliability*, 51(9-11):1508 – 1514, 2011. Proceedings of the 22th European Symposium on the Reliability of Electron Devices, Failure Physics and Analysis.

[8] N. Arora. *Mosfet modeling for VLSI simulation: theory and practice*. World Scientific, 2007.

[9] S. Aur, D. E. Hocevar, and P. Yang. HOTRON-A circuit hot electron effect simulator. In *IEEE International Conference of Computer-Aided Design*, pages 256–259, 1987.

[10] R. J. Baker. *CMOS: circuit design, layout, and simulation*, volume 18. John Wiley & Sons, 2011.

[11] D. Binkley. Tradeoffs and optimization in analog CMOS design. In *Mixed Design of Integrated Circuits and Systems, 2007. MIXDES '07. 14th International Conference on*, pages 47–60, June 2007.

[12] D. M. Binkley. *Tradeoffs and Optimization in Analog CMOS Design*. John Wiley & Sons, Ltd, 2008.

[13] J. R. Black. Electromigration;a brief survey and some recent results. *Electron Devices, IEEE Transactions on*, 16(4):338–347, Apr 1969.

[14] A. Brown, N. Idris, J. Watling, and A. Asenov. Impact of metal gate granularity on threshold voltage variability: A full-scale three-dimensional statistical simulation study. *Electron Device Letters, IEEE*, 31(11):1199–1201, Nov 2010.

[15] M. Bucher, D. Kazazis, F. Krummenacher, D. Binkley, D. Foty, and Y. Papananos. Analysis of transconductances at all levels of inversion in deep submicron CMOS. In *Electronics, Circuits and Systems, 2002. 9th International Conference on*, volume 3, pages 1183–1186 vol.3, 2002.

[16] C. Buneo, M. Jarvis, A. Batista, and R. Andersen. Properties of spike train spectra in two parietal reach areas. *Experimental Brain Research*, 153:134–139, 2003.

[17] H. Cai, K. Liu, and L. de Barros Naviner. Reliability-aware delay faults evaluation of CMOS flip-flops. In *Mixed Design of Integrated Circuits Systems (MIXDES), 2014 Proceedings of the 21st International Conference*, pages 385–389, June 2014.

[18] E. Cartier and A. Kerber. Stress-induced leakage current and defect generation in nFETs with hfo2/tin gate stacks during positive-bias temperature stress. In *Reliability Physics Symposium, 2009 IEEE International*, pages 486–492, April 2009.

[19] S. Chakravarthi, A. Krishnan, V. Reddy, C. Machala, and S. Krishnan. A comprehensive framework for predictive modeling of negative bias temperature instability. In *Reliability Physics Symposium Proceedings, 2004. 42nd Annual. 2004 IEEE International*, pages 273–282, April 2004.

[20] J. Chen, G. Shi, A. Tai, and F. Lee. A size sensitivity method for interactive MOS circuit sizing. In *New Circuits and Systems Conference (NEWCAS), 2012 IEEE 10th International*, pages 169–172, June 2012.

[21] F. R. Chouard. *Device Aging in Analog Circuits for Nanoelectronic CMOS Technologies*. PhD thesis, Munic, Technical University of Munic, Diss., 2012, 2012.

[22] D. M. Colombo, G. I. Wirth, and C. Fayomi. Design methodology using inversion coefficient for low-voltage low-power CMOS voltage reference. In *Proceedings of the 23rd Symposium on Integrated Circuits and System Design*, SBCCI '10, pages 43–48, New York, NY, USA, 2010. ACM.

[23] A. Comsol. COMSOL multiphysics user's guide. *Version: September*, 2005.

[24] A. Cunha, M. Schneider, and C. Galup-Montoro. An MOS transistor model for analog circuit design. *Solid-State Circuits, IEEE Journal of*, 33(10):1510–1519, Oct 1998.

[25] G. Dacey and I. Ross. Unipolar "field-effect" transistor. *Proceedings of the IRE*, 41(8):970–979, Aug 1953.

[26] A. Deif. Sensitivity analysis from the state equations by perturbation techniques. *Applied Mathematical Modelling*, 5(6):405 – 408, 1981.

[27] M. Dietrich and J. Haase. *Process Variations and Probabilistic Integrated Circuit Design*. Springer Science & Business Media, 2011.

[28] J. Duque-Carrillo, R. Perez-Aloe, and J. Valverde. Biasing circuit for high input swing operational amplifiers. *Solid-State Circuits, IEEE Journal of*, 30(2):156–159, Feb 1995.

[29] C. Duvvury and S. Aur. Hot-carrier degradation. *Hot Carrier Design Considerations for MOS Devices and Circuits*, page 120, 2012.

[30] Y. El-Mansy. Modelling weak avalanche multiplication currents in IGFETs and SOS transistors for CAD. In *Electron Devices Meeting, 1975 International*, volume 21, pages 31–34, 1975.

[31] C. Enz. A short story of the EKV MOS transistor model. *Solid-State Circuits Society Newsletter, IEEE*, 13(3):24–30, Summer 2008.

[32] C. Enz, F. Krummenacher, and E. Vittoz. An analytical MOS transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications. *Analog Integrated Circuits and Signal Processing*, 8(1):83–114, 1995.

[33] C. C. Enz and E. A. Vittoz. *Charge-based MOS transistor modeling: the EKV model for low-power and RF IC design*. John Wiley & Sons, 2006.

[34] M.-L. Fan, V.-H. Hu, Y.-N. Chen, P. Su, and C.-T. Chuang. Impacts of random telegraph noise on finfet devices, 6t sram cell, and logic circuits. In *Reliability Physics Symposium (IRPS), 2012 IEEE International*, pages CR.1.1–CR.1.6, April 2012.

[35] N. Georgieva, S. Glavic, M. Bakr, and J. Bandler. Feasible adjoint sensitivity technique for em design optimization. *Microwave Theory and Techniques, IEEE Transactions on*, 50(12):2751–2758, Dec 2002.

[36] G. Gielen, P. De Wit, E. Maricau, J. Loeckx, J. Martín-Martínez, B. Kaczer, G. Groeseneken, R. Rodríguez, and M. Nafría. Emerging yield and reliability challenges in nanometer cmos technologies. In *Proceedings of the Conference on Design, Automation and Test in Europe*, DATE '08, pages 1322–1327, New York, NY, USA, 2008. ACM.

[37] S. Gorke, R. Riebeling, F. Kraus, and R. Reichel. Flexible platform approach for fly-by-wire systems. In *Digital Avionics Systems Conference (DASC), 2013 IEEE/AIAA 32nd*, pages 2C5–1–2C5–16, Oct 2013.

[38] H. Graeb. Itrs 2011 analog eda challenges and approaches. In *Proceedings of the Conference on Design, Automation and Test in Europe*, DATE '12, pages 1150–1155, San Jose, CA, USA, 2012. EDA Consortium.

[39] G. Gramegna, A. D'Aquila, and B. Marletta. High-precision biasing circuit for a cascoded cmos stage, particularly for low noise amplifiers, May 21 2002. US Patent 6,392,490.

[40] T. Grasser, T. Aichinger, G. Pobegen, H. Reisinger, P.-J. Wagner, J. Franco, M. Nelhiebel, and B. Kaczer. The permanent component of NBTI: Composition and annealing. In *Reliability Physics Symposium (IRPS), 2011 IEEE International*, pages 6A.2.1–6A.2.9, April 2011.

[41] T. Grasser and B. Kaczer. Evidence that two tightly coupled mechanisms are responsible for negative bias temperature instability in oxynitride MOSFETs. *Electron Devices, IEEE Transactions on*, 56(5):1056–1062, May 2009.

[42] T. Grasser, B. Kaczer, W. Goes, H. Reisinger, T. Aichinger, P. Hehenberger, P.-J. Wagner, F. Schanovsky, J. Franco, P. Roussel, and M. Nelhiebel. Recent advances in understanding the bias temperature instability. In *Electron Devices Meeting (IEDM), 2010 IEEE International*, pages 4.4.1–4.4.4, Dec 2010.

[43] T. Grasser, B. Kaczer, P. Hehenberger, W. Gos, R. O'Connor, H. Reisinger, W. Gustin, and C. Schunder. Simultaneous extraction of recoverable and permanent components contributing to bias-temperature instability. In *Electron Devices Meeting, 2007. IEDM 2007. IEEE International*, pages 801–804, Dec 2007.

[44] T. Grasser, H. Reisinger, P.-J. Wagner, F. Schanovsky, W. Goes, and B. Kaczer. The time dependent defect spectroscopy (TDDS) for the characterization of the bias temperature instability. In *Reliability Physics Symposium (IRPS), 2010 IEEE International*, pages 16–25, May 2010.

[45] R. R. Harrison and C. Charles. A low-power low-noise CMOS amplifier for neural recording applications. *Solid-State Circuits, IEEE Journal of*, 38(6):958–965, 2003.

[46] A. Hastings. *The art of analog layout*. Prentice Hall, 2006.

[47] P. Heim and M. Jabri. MOS cascode-mirror biasing circuit operating at any current level with minimal output saturation voltage. *Electronics Letters*, 31:690–691(1), April 1995.

[48] N. Hellwege, N. Heidmann, S. Paul, and D. Peters-Drolshagen. Optimum operating points of transistors with minimal aging-aware sensitivity. In *IEEE SBCCI 2015*, August 2015.

[49] N. Hellwege, N. Heidmann, D. Peters-Drolshagen, and S. Paul. Optimierte zuverlässigkeitsanalyse eines cmos-operationsverstärkers zur aufnahme von neuronalen signalen. *ITG-Fachbericht-ANALOG 2013-Entwicklung von Analogschaltungen mit CAE-Methoden*, 2013.

[50] N. Hellwege, N. Heidmann, D. Peters-Drolshagen, and S. Paul. Reliability analysis for integrated circuit amplifiers used in neural measurement systems. In *Design, Automation Test in Europe Conference Exhibition (DATE), 2013*, pages 713–716, March 2013.

[51] N. Hellwege, N. Heidmann, D. Peters-Drolshagen, and S. Paul. Using operating point-dependent degradation and gm/id method for aging-aware design. In *Integrated Reliability Workshop Final Report (IRW), 2013 IEEE International*, pages 113–116, Oct 2013.

[52] N. Hellwege, N. Heidmann, D. Peters-Drolshagen, and S. Paul. Variability-aware gradual aging for generating reliability figures of a neural measurement system. In *Mixed Design of Integrated Circuits and Systems (MIXDES), 2013 Proceedings of the 20th International Conference*, pages 380–385, June 2013.

[53] N. Hellwege, N. Heidmann, D. Peters-Drolshagen, and S. Paul. AAS-maps: Aging-aware sensitivity-maps for reliability driven analog circuit design. In *Reliability Physics Symposium, 2014 IEEE International*, pages 5D.3.1–5D.3.6, June 2014.

[54] N. Hellwege, N. Heidmann, D. Peters-Drolshagen, and S. Paul. Analyse arbeitspunktabhängiger degradierung bei mos-transistoren durch nbti und hci. *eda-Workshop 14*, 2014.

[55] N. Hellwege, N. Heidmann, D. Peters-Drolshagen, and S. Paul. An aging-aware transistor sizing tool regarding BTI and HCD degradation modes. In *Mixed Design of Integrated Circuits and Systems (MIXDES), 2015 Proceedings of the 22th International Conference*, June 2015.

[56] J. Henrie and R. Shapley. LFP power spectra in v1 cortex: the graded effect of stimulus contrast. *J. Neurophysiol.*, pages 479 – 490, 2005.

[57] P. Heremans, R. Bellens, G. Groeseneken, and H. Maes. Consistent model for the hot-carrier degradation in n-channel and p-channel MOSFETs. *Electron Devices, IEEE Transactions on*, 35(12):2194–2209, Dec 1988.

[58] D. Hisamoto, W.-C. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. Anderson, T.-J. King, J. Bokor, and C. Hu. FinFET-a self-aligned double-gate MOSFET scalable to 20 nm. *Electron Devices, IEEE Transactions on*, 47(12):2320–2325, Dec 2000.

[59] C. Hu. The berkeley reliability simulator bert: an IC reliability simulator. *Microelectronics Journal*, 23(2):97 – 102, 1992.

[60] C. Hu, S. C.Tam, F.-C. Hsu, P.-K. Ko, T.-Y. Chan, and K. Terrill. Hot-electron-induced MOSFET degradation-model, monitor, and improvement. *Electron Devices, IEEE Transactions on*, 32(2):375–385, Feb 1985.

[61] D. Ielmini, M. Manigrasso, F. Gattel, and G. Valentini. A unified model for permanent and recoverable NBTI based on hole trapping and structure relaxation. In *Reliability Physics Symposium, 2009 IEEE International*, pages 26–32, April 2009.

[62] K. O. Jeppson and C. M. Svensson. Negative bias stress of MOS devices at high electric fields and degradation of NMOS devices. *Journal of Applied Physics*, 48(5):2004–2014, 1977.

[63] P. G. Jespers. *The Gm/ID design methodology for CMOS analog low power integrated circuits*. Springer US, 2007.

[64] B. Kaczer, T. Grasser, P. Roussel, J. Martin-Martinez, R. O'Connor, B. O'Sullivan, and G. Groeseneken. Ubiquitous relaxation in BTI stressing - new evaluation and insights. In *Reliability Physics Symposium, 2008. IRPS 2008. IEEE International*, pages 20–27, April 2008.

[65] B. Kaczer, P. Roussel, T. Grasser, and G. Groeseneken. Statistics of multiple trapped charges in the gate oxide of deeply scaled MOSFET devices - application to nbti. *Electron Device Letters, IEEE*, 31(5):411–413, May 2010.

[66] T. Karnik and P. Hazucha. Characterization of soft errors caused by single event upsets in CMOS processes. *Dependable and Secure Computing, IEEE Transactions on*, 1(2):128–143, April 2004.

[67] K. Kim and A. A. Iliadis. Latch-up effects in CMOS inverters due to high power pulsed electromagnetic interference. *Solid-State Electronics*, 52(10):1589 – 1593,

2008. Papers Selected from the International Semiconductor Device Research Symposium 2007 - {ISDRS} 2007.

[68] S.-D. Kim, H. Wada, and J. Woo. TCAD-based statistical analysis and modeling of gate line-edge roughness effect on nanoscale mos transistor performance and scaling. *Semiconductor Manufacturing, IEEE Transactions on*, 17(2):192–200, May 2004.

[69] J. Kludt, K. Weide-Zaage, M. Ackermann, and V. Hein. Deformation of octahedron slotted metal tracks. In *Integrated Reliability Workshop Final Report (IRW), 2013 IEEE International*, pages 161–165. IEEE, 2013.

[70] G.-A. Klutke, P. C. Kiessler, and M. Wortman. A critical look at the bathtub curve. *IEEE Transactions on Reliability*, 52(1):125–129, 2003.

[71] A. Krishnan, C. Chancellor, S. Chakravarthi, P. Nicollian, V. Reddy, A. Varghese, R. Khamankar, and S. Krishnan. Material dependence of hydrogen diffusion: implications for NBTI degradation. In *Electron Devices Meeting, 2005. IEDM Technical Digest. IEEE International*, pages 4 pp.–691, Dec 2005.

[72] H. Kufluoglu and M. A. Alam. A geometrical unification of the theories of NBTI and HCI time-exponents and its implications for ultra-scaled planar and surround-gate MOSFETs. *IEDM Tech. Dig*, 113, 2004.

[73] A. A. Kuhn, T. Trottenberg, A. Kivi, A. Kupsch, G.-H. Schneider, and P. Brown. The relationship between local field potential and neuronal discharge in the subthalamic nucleus of patients with parkinson's disease. *Experimental neurology*, 194(1):212–220, 2005.

[74] K. Kuhn. Reducing variation in advanced logic technologies: Approaches to process and design for manufacturability of nanoscale CMOS. In *Electron Devices Meeting, 2007. IEDM 2007. IEEE International*, pages 471–474, Dec 2007.

[75] W. Kuo, W.-T. K. Chien, and T. Kim. *Reliability, Yield, and Stress Burn-in: A Unified Approach for Microelectronics Systems Manufacturing & Software Development*. Springer Science & Business Media, 1998.

[76] L. Lewyn, T. Ytterdal, C. Wulff, and K. Martin. Analog circuit design in nanoscale CMOS technologies. *Proceedings of the IEEE*, 97(10):1687–1714, Oct 2009.

[77] J. Lienig. Introduction to electromigration-aware physical design. In *Proceedings of the 2006 International Symposium on Physical Design*, ISPD '06, pages 39–46, New York, NY, USA, 2006. ACM.

[78] E. Maricau and G. Gielen. Computer-aided analog circuit design for reliability in nanometer CMOS. *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on*, 1(1):50–58, March 2011.

[79] E. Maricau and G. Gielen. *Analog IC Reliability in Nanometer CMOS*. Springer Science & Business Media, 2013.

[80] E. Maricau, P. D. Wit, and G. Gielen. An analytical model for hot carrier degradation in nanoscale CMOS suitable for the simulation of degradation in analog IC applications. *Microelectronics Reliability*, 48(8-9):1576–1580, 2008. 19th European Symposium on Reliability of Electron Devices, Failure Physics and Analysis (ESREF 2008).

[81] J. Martin-Martinez, R. Rodríguez, M. Nafria, and X. Aymerich. Time-dependent variability related to BTI effects in MOSFETs: Impact on CMOS differential amplifiers. *Device and Materials Reliability, IEEE Transactions on*, 9(2):305–310, 2009.

[82] K. Mistry, C. Allen, C. Auth, B. Beattie, D. Bergstrom, M. Bost, M. Brazier, M. Buehler, A. Cappellani, R. Chau, et al. A 45nm logic technology with high-k+ metal gate transistors, strained silicon, 9 Cu interconnect layers, 193nm dry patterning, and 100% pb-free packaging. In *Electron Devices Meeting, 2007. IEDM 2007. IEEE International*, pages 247–250. IEEE, 2007.

[83] R. Muller, S. Gambini, and J. Rabaey. A 0.013mm2, 5µW, dc-coupled neural signal acquisition ic with 0.5V supply. *Journal of Solid-State Circuits*, 47(1), January 2012.

[84] C. T. Nordhausen, E. M. Maynard, and R. A. Normann. Single unit recording capabilities of a 100 microelectrode array. *Brain Research*, 726(1-2):129 – 140, 1996.

[85] L.-T. Pang and B. Nikolic. Measurements and analysis of process variability in 90 nm CMOS. *Solid-State Circuits, IEEE Journal of*, 44(5):1655–1663, May 2009.

[86] C. Parthasarathy, M. Denais, V. Huard, G. Ribes, D. Roy, C. Guerin, F. Perrier, E. Vincent, and A. Bravaix. Designing in reliability in advanced CMOS technologies. *Microelectronics Reliability*, 46(9):1464–1471, 2006.

[87] M. J. Pelgrom, A. C. Duinmaijer, A. P. Welbers, et al. Matching properties of MOS transistors. *IEEE Journal of solid-state circuits*, 24(5):1433–1439, 1989.

[88] L. T. Pillage, R. A. Rohrer, and C. Visweswariah. *Electronic circuit and system simulation methods*. McGraw-Hill New York, 1995.

[89] G. Pollissard-Quatremere, G. Gosset, and D. Flandre. A modified gm/id design methodology for deeply scaled CMOS technologies. *Analog Integrated Circuits and Signal Processing*, pages 1–14, 2013.

[90] T. Quarles, A. Newton, D. Pederson, and A. Sangiovanni-Vincentelli. SPICE3 version 3.3 user's manual. *University of California, Berkeley, CA*, 94720, 1993.

[91] S. Ramey, C. Prasad, M. Agostinelli, S. Pae, S. Walstra, S. Gupta, and J. Hicks. Frequency and recovery effects in high-k BTI degradation. In *Reliability Physics Symposium, 2009 IEEE International*, pages 1023–1027, April 2009.

[92] H. Reisinger, T. Grasser, K. Ermisch, H. Nielen, W. Gustin, and C. Schlunder. Understanding and modeling AC BTI. In *Reliability Physics Symposium (IRPS), 2011 IEEE International*, pages 6A.1.1–6A.1.8, April 2011.

[93] C. RelXpert. Users manuals BSIMPro+/RelXpert/UltraSim. *Cadence Design Systems Inc*, 2010.

[94] A. Rofougaran and A. Abidi. A table lookup FET model for accurate analog circuit simulation. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 12(2):324–335, Feb 1993.

[95] T. Sakurai and A. Newton. A simple MOSFET model for circuit analysis. *Electron Devices, IEEE Transactions on*, 38(4):887–894, Apr 1991.

[96] W. M. Sansen. *Analog design essentials*, volume 859. Springer Science & Business Media, 2007.

[97] G. T. Sasse. *Reliability engineering in RF CMOS*. PhD thesis, University of Twente, Enschede, July 2008.

[98] S. Schmale, J. Hoeffmann, B. Knoop, G. Kreiselmeyer, H. Hamer, D. Peters-Drolshagen, and S. Paul. Exploiting correlation in neural signals for data compression. *22nd European Signal Processing Conference (EUSIPCO 2014)*, Sept. 2014.

[99] S. Schmale, B. Knoop, J. Hoeffmann, D. Peters-Drolshagen, and S. Paul. Joint compression of neural action potentials and local field potentials. *47th Asilomar Conference on Signals, Systems and Computers*, Nov. 2013.

[100] D. K. Schroder and J. A. Babcock. Negative bias temperature instability: Road to cross in deep submicron silicon semiconductor manufacturing. *Journal of Applied Physics*, 94(1):1–18, 2003.

[101] B. Sheu, W.-J. Hsu, and B. Lee. An integrated-circuit reliability simulator-RELY. *Solid-State Circuits, IEEE Journal of*, 24(2):473–477, April 1989.

[102] T. Shima and H. Tamada. Table look-up MOSFET modeling system using a 2-d device simulator and monotonic piecewise cubic interpolation. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 2(2):121–126, April 1983.

[103] W. Shockley. A unipolar "field-effect" transistor. *Proceedings of the IRE*, 40(11):1365–1376, Nov 1952.

[104] W. Shockley. Problems related top-n junctions in silicon. *Cechoslovackij fiziceskij zurnal B*, 11(2):81–121, 1961.

[105] F. Silveira, D. Flandre, and P. Jespers. A gm/id based methodology for the design of cmos analog circuits and its application to the synthesis of a silicon-on-insulator micropower OTA. *Solid-State Circuits, IEEE Journal of*, 31(9):1314–1319, Sep 1996.

[106] P. Stolk, F. Widdershoven, and D. Klaassen. Modeling statistical dopant fluctuations in MOS transistors. *Electron Devices, IEEE Transactions on*, 45(9):1960–1971, Sep 1998.

[107] K.-W. Su, Y.-M. Sheu, C.-K. Lin, S.-J. Yang, W.-J. Liang, X. Xi, C.-S. Chiang, J.-K. Her, Y.-T. Chia, C. Diaz, and C. M. Hu. A scaleable model for STI mechanical stress effect on layout dependence of MOS electrical characteristics. In *Custom Integrated Circuits Conference, 2003. Proceedings of the IEEE 2003*, pages 245–248, Sept 2003.

[108] E. Takeda, N. Suzuki, and T. Hagiwara. Device performance degradation to hot-carrier injection at energies below the Si-SiO2energy barrier. In *Electron Devices Meeting, 1983 International*, volume 29, pages 396–399, 1983.

[109] K. Takeuchi, T. Fukai, T. Tsunomura, A. Putra, A. Nishida, S. Kamohara, and T. Hiramoto. Understanding random threshold voltage fluctuation by comparing multiple fabs and technologies. In *Electron Devices Meeting, 2007. IEDM 2007. IEEE International*, pages 467–470, Dec 2007.

[110] R. Thakker, C. Sathe, A. Sachid, M. Baghini, V. Rao, and M. Patil. A novel table-based approach for design of FinFET circuits. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 28(7):1061–1070, July 2009.

[111] E. Tolstosheeva, J. Hoeffmann, J. Pistor, D. Rotermund, T. Schellenberg, D. Boll, T. Hertzberg, V. Gordillo-Gonzalez, S. Mandon, D. Peters-Drolshagen, M. Schneider, K. Pawelzik, A. Kreiter, S. Paul, and W. Lang. Towards a wireless and fully-implantable ECoG system. In *Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS EUROSENSORS XXVII), 2013 Transducers Eurosensors XXVII: The 17th International Conference on*, pages 384–387, June 2013.

[112] Y. P. Tsividis and C. McAndrew. *Operation and Modeling of the MOS Transistor*. Oxford University Press, 2011.

[113] T. Tsunomura, J. Nishimura, A. Kumar, A. Nishida, S. Inaba, K. Takeuchi, T. Hiramoto, and T. Mogami. Suppression of VT variability degradation induced by NBTI with RDF control. In *VLSI Technology (VLSIT), 2011 Symposium on*, pages 150–151, June 2011.

[114] K. Tu. Recent advances on electromigration in very-large-scale-integration of interconnects. *Journal of applied physics*, 94(9):5451–5473, 2003.

[115] B. Tudor, J. Wang, C. Sun, Z. Chen, Z. Liao, R. Tan, W. Liu, and F. Lee. MOSRA: An efficient and versatile mos aging modeling and reliability analysis solution for 45nm and below. In *Solid-State and Integrated Circuit Technology (ICSICT), 2010 10th IEEE International Conference on*, pages 1645–1647, Nov 2010.

[116] D. Varghese, D. Saha, S. Mahapatra, K. Ahmed, F. Nouri, and M. Alam. On the dispersive versus arrhenius temperature activation of nbti time evolution in plasma nitrided gate oxides: measurements, theory, and implications. In *Electron Devices Meeting, 2005. IEDM Technical Digest. IEEE International*, pages 684–687, Dec 2005.

[117] R. Viereckl, D. Ahlemann, J. Assmann, F. Kuhnert, C. Stuermer, and S. Bratzel. Connected car studie 2014. Technical report, PricewaterhouseCoopers Aktiengesellschaft Wirtschaftsprüfungsgesellschaft, Strategy&, Center of Automotive Management, 2014.

[118] E. A. Vittoz. Micropower techniques, 1994.

[119] J. Vlach and K. Singhal. *Computer methods for circuit analysis and design*. Springer, 1983.

[120] W. Wang, V. Reddy, A. Krishnan, R. Vattikonda, S. Krishnan, and Y. Cao. Compact modeling and simulation of circuit reliability for 65-nm CMOS technology. *Device and Materials Reliability, IEEE Transactions on*, 7(4):509–517, Dec 2007.

[121] M. White and Y. Chen. *Scaled cmos technology reliability users guide*. Pasadena, CA: Jet Propulsion Laboratory, National Aeronautics and Space Administration, 2008.

[122] D. Wolters and J. Van Der Schoot. Kinetics of charge trapping in dielectrics. *Journal of applied physics*, 58(2):831–837, 1985.

[123] H.-S. Wong and Y. Taur. Three-dimensional "atomistic" simulation of discrete random dopant distribution effects in sub-0.1 /spl mu/m mosfet's. In *Electron Devices Meeting, 1993. IEDM '93. Technical Digest., International*, pages 705–708, Dec 1993.

[124] E. Y. Wu and J. Sune. Power-law voltage acceleration: A key element for ultra-thin gate oxide reliability. *Microelectronics Reliability*, 45(12):1809 – 1834, 2005.

[125] Q. Wu, M. Porti, A. Bayerl, M. Lanza, J. Martin-Martinez, R. Rodriguez, M. Nafria, X. Aymerich, and E. Simoen. Impact of NBTI and CHC stress on the nanoscale electrical properties of strained and non-strained MOSFETs. In *Electron Devices (CDE), 2015 10th Spanish Conference on*, pages 1–4, Feb 2015.

[126] M. Xie and C. Lai. Reliability analysis using an additive weibull model with bathtub-shaped failure rate function. *Reliability Engineering & System Safety*, 52(1):87 – 93, 1996.

[127] H. Yaegashi, K. Oyama, A. Hara, S. Natori, and S. Yamauchi. Overview: Continuous evolution on double-patterning process. In *SPIE Advanced Lithography*, pages 83250B–83250B. International Society for Optics and Photonics, 2012.

[128] S. Zafar, Y. Kim, V. Narayanan, C. Cabral, V. Paruchuri, B. Doris, J. Stathis, A. Callegari, and M. Chudzik. A comparative study of nbti and pbti (charge trapping) in sio2/hfo2 stacks with fusi, tin, re gates. In *VLSI Technology, 2006. Digest of Technical Papers. 2006 Symposium on*, pages 23–25, 2006.

[129] W. Zhao, F. Liu, K. Agarwal, D. Acharyya, S. Nassif, K. Nowka, and Y. Cao. Rigorous extraction of process variations for 65-nm CMOS design. *Semiconductor Manufacturing, IEEE Transactions on*, 22(1):196–203, Feb 2009.

[130] Z. Zumsteg, C. Kemere, S. O'Driscoll, G. Santhanam, R. Ahmed, K. Shenoy, and T. Meng. Power feasibility of implantable digital spike sorting circuits for neural prosthetic systems. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 13(3):272 –279, sept. 2005.

Scientific Publications

The following list contains scientific publications and conference presentations contributed by the author (list is ordered anti-chronologically).

Conferences

1. Sascha Heinßen, Nico Hellwege, Nils Heidmann, Steffen Paul, Dagmar Peters-Drolshagen: *Robust Digital Calibration Engine for MEMS Inertial Sensor Systems*, IEEE SENSORS 2015, Busan, South Korea, November 2015

2. Theodor Hillebrand, Nico Hellwege, Nils Heidmann, Steffen Paul, Dagmar Peters-Drolshagen: *Charge-based Stochastic Aging Analysis of CMOS Circuits*, IEEE IIRW 2015, Fallen Leaf Lake, USA, Oktober 2015

3. Theodor Hillebrand, Nico Hellwege, Nils Heidmann, Dagmar Peters-Drolshagen, Steffen Paul: *Stochastic Analysis of Degradation and Variations in CMOS-Transistors*, ZuE 2015 - Zuverlässigkeit und Entwurf, Siegen, Germany, September 2015

4. Nico Hellwege, Sascha Heinßen, Kris Niederkleine, Nils Heidmann, Steffen Paul, Dagmar Peters-Drolshagen: *Process Variability Monitor for Embedded MEMS Ineratial Sensors Exploiting Digital Calibration Coefficients*, Eurosensors XXIX, Freiburg, Germany, September 2015

5. Nico Hellwege, Nils Heidmann, Steffen Paul, Dagmar Peters-Drolshagen: *Optimum Operating Points of Transistors with minimal Aging-Aware Sensitivity*, IEEE SBCCI 2015, Bahia, Brazil, August 2015

6. Nico Hellwege, Nils Heidmann, Marco Erstling, Dagmar Peters-Drolshagen, Steffen Paul: *An Aging-Aware Transistor Sizing Tool Regarding BTI and HCD Degradation Modes*, IEEE MIXDES, Torun, Poland, Juni 2015

7. Nils Heidmann, Nico Hellwege, Steffen Paul, Dagmar Peters-Drolshagen: *Variability-Aware Aging Modeling for Reliability Analysis of an Analog Neural Measurement System*, IEEE European Test Symposium, Cluj-Napoca, Romania, Mai 2015

8. Nils Heidmann, Nico Hellwege, Steffen Paul, Dagmar Peters-Drolshagen: *Behavioral Modeling of an Analog Neural Recording Front-End for System Simulation and Reliability Analysis*, edaWorkshop15, Dresden, Germany, Mai 2015

9. Nils Heidmann, Nico Hellwege, Steffen Paul, Dagmar Peters-Drolshagen: *NBTI and HCD Aware Behavioral Models for Reliability Analysis of Analog CMOS Circuits*, IEEE IRPS, Monterey, USA, April 2015

10. Georg Georgakos, Domenik Helms, Nils Heidmann, Nico Hellwege, et. al. : *RELY - Neue Mehtoden zum Entwurf von SoCs für kontrollierbare hohe Zuverlässigkeit für Anwendungen wie Transport, Medizin und Automatisierung*, newsletter edacentrum vol. 1, Hannover, Dezember 2014

11. Nasim Pour Aryan, Nils Heidmann, Martin Wirnshofer, Nico Hellwege, Jonas Pistor, Dagmar Peters-Drolshagen, Georg Georgakos, Steffen Paul and Doris Schmitt-Landsiedell: *Power Efficient Digital IC Design for a Medical Application with High Reliability Requirements*, PATMOS, Palma de Mallorca, Spain, September 2014

12. Nils Heidmann, Nico Hellwege, Jonas Pistor, Dagmar Peters-Drolshagen, Steffen Paul: *A Modular Analog Front-End for the Recording of Neural Spikes and Local Field Potentials within a Neural Measurement System*, Eurosensors XXVIII, Brescia, Italy, September 2014

13. Jonas Pistor, Nils Heidmann, Janpeter Höffmann, Steffen Paul: *Programmable Current Source for Implantable Neural Stimulation Systems*, Eurosensors XXVIII, Brescia, Italy, September 2014

14. Nils Heidmann, Nico Hellwege, Maike Taddiken, Dagmar Peters-Drolshagen, Steffen Paul: *Analog Behavioral Modeling for Age-Dependent Degradation of Complex Analog Circuits*, IEEE MIXDES, Lublin, Poland, Juni 2014

15. Nico Hellwege, Nils Heidmann, Dagmar Peters-Drolshagen, Steffen Paul: *AAS-Maps: Aging-Aware Sensitivity Maps for Reliability Driven Analog Cicuit Design*, IEEE IRPS, Hawaii, USA, Juni 2014

16. Nico Hellwege, Nils Heidmann, Dagmar Peters-Drolshagen, Steffen Paul: *Analyse arbeitspunkt-abhängiger Degradierung bei MOS Transistoren durch NBTI und HCI*, edaWorkshop14, Hannover, Germany, Mai 2014

17. Nils Heidmann, Nico Hellwege, Tim Höhlein, Thomas Westphal, Dagmar Peters-Drolshagen, Steffen Paul: *Modeling of an Analog Recording System Design for ECoG and AP Signals*, Design, Automation & Test in Europe Conference & Exhibition (DATE), Dresden, Germany, März 2014

18. Nico Hellwege, Nils Heidmann, Dagmar Peters-Drolshagen, Steffen Paul: *Using Operating Point-Dependent Degradation and gm/Id Method for Aging-Aware Design*, IEEE IIRW 2013, Fallen Leaf Lake, USA

19. Nils Heidmann, Alexander Dannies, Nico Hellwege, Dagmar Peter-Drolshagen, Walter Lang, Steffen Paul: *A Low-Power Wireless UHF / LF Sensor Network with Web-Based Remote Supervision-Unit - Implementation in the Intelligent Container*, IEEE SENSORS 2013, Baltimore, USA

20. Nico Hellwege, Nils Heidmann, Dagmar Peters-Drolshagen, Steffen Paul: *Variability-aware Gradual Aging for Generating Reliability Figures of a Neural Measurement System* IEEE MIXDES 2013, Gdynia, Poland, Juni 2013

21. Nico Hellwege, Nils Heidmann, Dagmar Peters-Drolshagen, Steffen Paul: *Optimierte Zuverlässigkeitsanalyse eines CMOS-Operationsverstärkers zur Aufnahme von neuronalen Signalen*, ITG-Fachbericht-ANALOG 2013 - Entwicklung von Analogschaltungen mit CAE-Methoden, Aachen, Germany, März 2013

22. Nico Hellwege, Nils Heidmann, Dagmar Peters-Drolshagen, Steffen Paul: *Reliability Analysis for Integrated Circuit Amplifiers Used in Neural Measurement Systems*, Design, Automation & Test in Europe Conference & Exhibition (DATE), Grenoble, France, März 2013

Presentations

- Steffen Paul, Nico Hellwege, Nils Heidmann, Dagmar Peters-Drolshagen: *Reliability-Driven Analog Circuit Design using gm/Id Method and Cross Layer Modelling of Aging*, Design Automation Conference (DAC), Workshop, San Francisco, USA, Juni 2015

- Nils Heidmann, Nico Hellwege, Maike Taddiken, Dagmar Peters-Drolshagen, Steffen Paul: *AD-RSM: Using Age-Dependent Degradation and Response Surface Models for Analog Behavioral Modeling*, Design Automation Conference (DAC), WIP session, San Francisco, USA, Juni 2014