

UNIVERSITÄT DORTMUND

Fakultät für Elektrotechnik und Informationstechnik



**Zuverlässigkeit von
sub- μm -CMOS-Schaltungen bei
Bias-Temperature-Stress (BTS)**

Der Fakultät für Elektrotechnik und Informationstechnik
der Universität Dortmund vorgelegte

Dissertation

zur Erlangung des akademischen Grades
Doktor der Ingenieurwissenschaften

von

Dipl.-Ing. Christian Schlünder

Dortmund, Februar 2006

Inhaltsverzeichnis

1	EINLEITUNG	1
2	BIAS-TEMPERATURE-STRESS	3
2.1	BELASTUNGSBEDINGUNGEN IN SCHALTUNGEN	4
2.1.1	<i>Homogene Belastungsbedingungen in Digitalschaltungen</i>	7
2.1.2	<i>Homogene Belastungsbedingungen in Analogschaltungen</i>	8
2.1.3	<i>Inhomogene Belastungsbedingungen in Analog- und HF-Schaltungen</i>	11
2.2	PHYSIKALISCHE MECHANISMEN BEI BELASTUNG	14
2.2.1	<i>Belastungsinduzierte Grenzflächenzustände</i>	14
2.2.2	<i>Einbau fester Ladungen in das Gateoxid.....</i>	16
2.3	DEGRADATION VON TRANSISTOREIGENSCHAFTEN	17
2.4	ZEITLICHER VERLAUF DER DEGRADATION	17
2.5	AUSWIRKUNG DER DEGRADATION AUF SCHALTUNGSEIGENSCHAFTEN	20
2.5.1	<i>Auswirkung der Degradation nach homogener Belastung in Digitalschaltungen.....</i>	22
2.5.2	<i>Auswirkung der Degradation nach homogener Belastung in Analogschaltungen</i>	22
2.5.3	<i>Auswirkung der Degradation nach inhomogener Belastung.....</i>	23
2.6	LEBENSDAUEREXTRAPOLATION.....	24
2.6.1	<i>‘Worst-Case’ und ‘Duty-Cycle’ Ansatz</i>	24
2.6.2	<i>Beschleunigung.....</i>	27
2.6.2.1	Beschleunigung durch erhöhte Spannung	28
2.6.2.2	Beschleunigung durch erhöhte Temperatur.....	29
2.6.3	<i>Extrapolation auf Betriebsbedingungen</i>	30
2.6.3.1	Extrapolation der Belastungsspannung	30
2.6.3.2	Extrapolation der Belastungstemperatur	32
3	CHARAKTERISIERUNG UND BELASTUNG	34
3.1	AUFBAU UND ABLAUF DER BELASTUNGSEXPERIMENTE	34
3.2	DEFINITION WICHTIGER ELEKTRISCHER TRANSISTORPARAMETER	36
3.2.1	<i>Gateoxiddicke.....</i>	36
3.2.1.1	Physikalische Gateoxiddicke.....	37
3.2.1.2	Elektrisch wirksame Gateoxiddicke.....	39
3.2.2	<i>Einsatzspannung.....</i>	42
3.2.3	<i>Weitere Transistorparameter.....</i>	43
3.3	TRANSISTORCHARAKTERISIERUNG	45
3.3.1	<i>Eingangs- und Ausgangskennlinien.....</i>	45
3.3.2	<i>Einsatzspannung.....</i>	46
3.3.2.1	Steilheitsmethode	46
3.3.2.2	Zielstrommethode	48
3.3.2.3	DIBL-Effekt.....	49
3.3.3	<i>Charge-Pumping</i>	51

4	EXPERIMENTELLE ERGEBNISSE	61
4.1	HOMOGENE BELASTUNG.....	61
4.2	INHOMOGENE BELASTUNG.....	67
5	MODELLIERUNG	75
5.1	HERLEITUNG DER MODELLGLEICHUNGEN	76
5.2	VERIFIKATION.....	79
5.3	VEREINFACHTE MODELLIERUNG	84
5.4	RICHTLINIEN FÜR DIE LEBENSDAUERVORHERSAGE	85
6	SCHALTUNGSSIMULATION	88
6.1	METHODIK DER SIMULATION GESCHÄDIGTER SCHALTUNGEN.....	88
6.2	SIMULATIONSABLAUF FÜR EINE BELASTETE SCHALTUNG.....	89
6.2.1	<i>Bestimmung der in der Modellgleichung verwendeten Parameter</i>	89
6.2.2	<i>Bestimmung der individuellen Belastungen der Transistoren</i>	90
6.2.3	<i>Berechnung der zu erwartenden Änderung der Modellparameter</i>	92
6.2.4	<i>Neusimulation des Schaltungsverhaltens</i>	93
7	MAßNAHMEN ZUR VERMINDERUNG DER DEGRADATION	95
7.1	PROZESSTECHNISCHE MAßNAHMEN.....	95
7.1.1	<i>Einfluss verschiedener Prozessierungsschritte</i>	96
7.1.2	<i>Verwendung von Deuterium als Wasserstoffersatz</i>	97
7.2	SCHALTUNGSTECHNISCHE MAßNAHMEN	99
7.2.1	<i>Wahl des Transistorarbeitspunktes in belastungsintensiven Betriebsmoden</i>	100
8	BIAS-TEMPERATURE-STRESS-PROBLEMATIK ZUKÜNFTIGER CMOS-PROZESSE	102
8.1	MODERNE ‚DUAL WORKFUNCTION‘-SPEICHERTECHNOLOGIEN.....	103
8.2	SILICON ON INSULATOR	105
8.3	ZUKÜNFTIGE TRANSISTOR-BAUFORMEN	107
9	ZUSAMMENFASSUNG	112

Anhang

A1 PROZESSDATEN	115
A2 MESSAUFBAU UND TESTSTRUKTUREN	116
A3 MODELLBERECHNUNGEN	120
VERWENDETE SYMBOLE UND ABKÜRZUNGEN	122
LITERATURVERZEICHNIS	125
LEBENS LAUF	135
DANKSAGUNG	136

1 Einleitung

Seit der Erfindung des MOS-Feldeffekttransistors und seiner ersten Realisierung auf Siliziumbasis im Jahre 1960 [Heil35, Shockley48, Kahng60] erlangte die Halbleitertechnologie eine immer weiter wachsende Bedeutung. Mithilfe der Großintegrationstechnik [Goser90, Goser91] gelingt es heute, immer komplexere Schaltungen mit weiter steigenden Transistoranzahlen in einem Chip zu integrieren. Heutzutage werden Halbleiterprodukte mit teilweise mehr als einer Milliarde Transistoren innerhalb eines Chips produziert [ITRS04, IFX04, Intel04]. Die fortschreitende Verkleinerung der Transistorstrukturgrößen stellt den Schlüssel des Erfolges der Halbleitertechnik dar.

Mit den zunehmend geringeren Strukturgrößen steigen die technischen Anforderungen an die Fertigung von Halbleiterprodukten. Während bislang bei dem so genannten ‚Shrink‘-Prozess hauptsächlich die technischen Möglichkeiten, immer kleinere Strukturgrößen funktional und reproduzierbar herzustellen, große Herausforderungen darstellten, erlangen im tiefen Submikrometerbereich Zuverlässigkeitsaspekte der Transistoren eine zunehmende Bedeutung [Schlunder05_2].

Bei der Verkleinerung der Transistorgrößen werden die Spannungen, mit denen die Bauelemente betrieben werden, nicht im selben Maße wie die geometrischen Abmessungen verringert. Die elektrischen Felder innerhalb der Bauelemente steigen somit an. Speziell die Belastung der Gateoxide nimmt so mit dem Wechsel zu jeder neuen Technologiegeneration bedingt durch die steigenden Feldstärken weiter zu [ITRS04, Thewes99_1].

Sollen die in einer modernen Halbleitertechnologie gefertigten Schaltungen Anforderungen an Zuverlässigkeit für eine geforderte Lebensdauer erfüllen, müssen die auftretenden physikalischen Schädigungsmechanismen verstanden und mithilfe von Modellen beschrieben werden können.

In dieser Arbeit soll speziell auf den bei temperaturunterstützter Belastung der Gateoxide von MOS-Feldeffekt-Transistoren erfolgenden ‚Bias-Temperature-Stress‘ eingegangen werden. Die dabei auftretenden Schädigungsmechanismen und die daraus resultierenden Folgen für das einzelne Bauelement und die Schaltung stehen im Fokus der Arbeit.

Zuerst werden in Kapitel 2 die Belastungsbedingungen bei ‚Bias-Temperature-Stress‘ (BTS) eingeführt. Es wird die kausale Kette von den Belastungsbedingungen in einer Schaltung, den einsetzenden Schädigungsmechanismen über die Parameterdegradation bis letztendlich zur Beeinträchtigung der Schaltungsfunktionen und der Lebensdauer einer Schaltung diskutiert. Die dafür notwendige Methodik wird von der messtechnischen Untersuchung der Transistoren bis zu den für die Lebensdaueraussage notwendigen Extrapolationstechniken erläutert.

In Kapitel 3 werden die verschiedenen experimentellen Versuche vorgestellt. Beginnend mit einer Beschreibung des prinzipiellen Ablaufes aller Belastungsexperimente werden darauf folgend Belastung und Charakterisierung der Bauelemente im Einzelnen erklärt. Dabei wird sowohl auf die Bedeutung der einzelnen Transistorparameter, als auch auf die zugrunde liegende Messtechnik für die Parameterextraktion eingegangen. Insbesondere werden homogene und inhomogene Belastungsversuche durchgeführt und miteinander verglichen. In Kapitel 4 werden die experimentellen Ergebnisse diskutiert.

Kapitel 5 ist der physikalischen Modellierung der Transistordegradation nach ‚Bias-Temperature-Stress‘ gewidmet. Eine Gegenüberstellung zu bisherigen Modellen und eine Bewertung der Modellvorhersagequalität schließen dieses Kapitel ab.

In Kapitel 6 werden computergestützte Simulationen von Schaltkreisen behandelt. Nach einem kurzen Überblick der bisher angebotenen Funktionalität derartiger Simulatoren wird die Simulationen geschädigter Schaltungen betrachtet. Bei der schrittweisen Beschreibung der zugrunde liegenden Methodik eines Simulationsablaufes einer belasteten Schaltung wird ein von der bisherigen Praxis abweichender Ansatz erörtert. Darauf aufbauend werden die daraus resultierenden neuen Möglichkeiten und Konsequenzen diskutiert.

Kapitel 7 thematisiert Maßnahmen, mit deren Hilfe die belastungsbedingte Degradation von Schaltungen nach ‚Bias-Temperature-Stress‘ reduziert werden kann. Der Ansatzpunkt dafür kann im prozesstechnischen oder im schaltungstechnischen Bereich liegen. Bei den prozesstechnischen Maßnahmen soll insbesondere auf die Verwendung des Wasserstoff-Isotops Deuterium während der Prozessierung der Bauelemente eingegangen werden. Der positive Einfluss auf die Gateoxidbelastbarkeit wird anhand von experimentellen Vergleichsmessungen vorgestellt und diskutiert. Bei den schaltungstechnischen Maßnahmen wird beispielhaft ein Konzept vorgestellt, das einzelne Bauelemente vor starker Belastung schützen kann. Kapitel 8 gibt einen Ausblick auf die BTS-Problematik zukünftiger Halbleitertechnologien. In Kapitel 9 wird die Arbeit zusammenfasst.

2 Bias-Temperature-Stress

MOS-Feldeffekt-Transistoren durchlaufen während ihres Einsatzes in elektrischen Schaltungen unterschiedliche Betriebsmoden. Je nach Arbeitspunkt stellen diese Betriebszustände mehr oder weniger starke Belastungen für den Transistor dar und können zu einer Degradation des Bauelementes führen. Darunter versteht man eine im Laufe der Zeit auftretende Änderung seiner elektrischen Parameter bedingt durch verschiedene mikroskopische Schädigungen, die durch physikalische Mechanismen in den Bauelementen hervorgerufen werden. Diese Degradation des Bauelementes kann die Funktion einer Schaltung beeinträchtigen und letztendlich auch deren Ausfall bedingen.

In dieser Arbeit soll das Hauptaugenmerk auf Transistorschädigungen und ihren Auswirkungen liegen, die durch ‚Bias-Temperature-Stress‘ (BTS) hervorgerufen werden. In diesem Kapitel wird auf Grundlagen und die Bedeutung belastungsbedingter Degradation für die Lebensdauer von Bauelementen moderner CMOS-Technologien eingegangen. Zuerst werden die Belastungsbedingungen erläutert, die in einem p-MOS Transistor die Schädigungsmechanismen auslösen. Die schaltungstechnischen Voraussetzungen, die dabei innerhalb einer integrierten Schaltung vorherrschen, werden anhand typischer Beispiele für Digital-, Analog- und Hochfrequenzschaltungen näher vorgestellt. Daraufhin werden die physikalischen Vorgänge während der Schädigung näher betrachtet. Ihr Einfluss auf die elektrischen Parameter eines Transistors und letztendlich die Auswirkungen der Bauelementedegradation auf Betrieb und Funktion elektrischer Schaltungen werden darauf aufbauend diskutiert. Vervollständigt wird die Betrachtung von BTS durch eine Einführung in die Methodik für eine Vorhersage der Schaltungslebensdauer. Dafür wird auch auf die für die verschiedenen Extrapolationen notwendigen mathematischen Modelle eingegangen.

2.1 Belastungsbedingungen in Schaltungen

Die in dieser Arbeit untersuchten Degradationsphänomene eines Feldeffekttransistors basieren auf der Kombination eines vertikalen elektrischen Feldes innerhalb des Gateoxids des Transistors bei gleichzeitig erhöhter Temperatur [Jeppson77, Ogawa95, Ogawa96, Schlünder99, Chakravarthi04]. Man nennt diese Belastungsbedingung ‚Bias-Temperature-Stress‘ (BTS). Degradation nach einer solchen Belastung ist sowohl beim n-MOSFET als auch beim p-MOS Transistor zu beobachten, sie ist jedoch bei p-Kanal Transistoren unter gleichen Feldstärken stärker ausgeprägt [Makabe00].

In der Literatur finden sich neben den in dieser Arbeit verwendeten Abkürzung BTS bzw. BT-Belastung oft die Begriffe ‚Negative-Bias-Temperature-Stress‘ (NBTS) oder ‚Negative-Bias-Temperature-Instability‘ (NBTI). Sie beschreiben die Belastung beziehungsweise die Instabilität speziell bei negativen Gate-zu-Source-Spannungen. Das zusätzliche Wort ‚Negative‘ steht hier also für das Vorzeichen des vertikalen elektrischen Feldes über dem Gateoxid.

Ein starkes laterales elektrisches Feld, wie es z.B. bei Channel-Hot-Carrier-Belastungen vorliegt [Hu85, Heremans91, Brox94, Thewes95, Thewes97_1, Walter98, Thewes99_2], ist bei einfachen Standardversuchen nicht vorhanden. Unter Vernachlässigung der relativ komplexen Dotierungsprofile kann unter diesen Umständen von einem über dem gesamten Oxidbereich homogenem elektrischen Feld ausgegangen werden. BT-Belastungen führen dementsprechend im Gegensatz zu Hot-Carrier-Belastungen in erster Näherung zu einer homogenen Schädigung innerhalb der gesamten aktiven Fläche des Bauelementes. Aus diesem Sachverhalt ergibt sich ebenfalls, dass die Schädigung durch BTS in erster Näherung auch keine Kanallängenabhängigkeit aufweist.

Die teilweise unter identischen Belastungsbedingungen beobachtbaren stärkeren Degradationen bei Transistoren mit sehr geringen Kanalweiten oder sehr geringen Kanallängen lassen sich auf Randeffekte zurückführen, die das Degradationsverhalten beeinflussen. Eine Schwächung der Randbereiche der Oxidfläche gegenüber BT-Belastungen lässt sich dabei auf zwei unterschiedliche Effekte zurückführen.

Bedingt durch die bei modernen Technologien zwischen benachbarten MOSFETs eingesetzte Grabenisolation (Shallow Trench Isolation - STI) werden durch die

hierfür notwendigen Fertigungsschritte die Gateoxide in den Außenbereichen längs der Kanallänge zusätzlichen Belastungen ausgesetzt. In diesen Bereichen können die benachbarten mit Oxid aufgefüllten Gräben zu Beeinträchtigungen des Aufwuchsprozesses und zu einer geringeren Oxidqualität führen.

Die vorgeschädigten Oxidbereiche zeigen unter BT-Belastungen stärkere Degradationen als die restlichen Bereiche. Der Anteil des durch den STI-Prozess beeinflussten Gateoxides ist weitestgehend unabhängig von der gezeichneten Weite des Transistors. Das bedeutet, dass der prozentuale Anteil dieser Bereiche bei schmalen Transistoren besonders hoch ist und sich daher bei solchen Bauelementen am deutlichsten auswirkt. Die stärkere Degradation der Randbereiche führt daher besonders bei Transistoren mit geringen Weiten zu stärker ausgeprägten Parameterdegradationen [Chung04].

Ein ähnlicher, die Randbereiche des Gateoxids am Anfang und am Ende des Kanals beeinträchtigender Effekt kann zu einer Verstärkung der Degradation bei sehr kurzen Kanallängen führen [Yamamoto99]. Die Ätzschritte zur Strukturierung der Gateelektrode und die anschließende Ionenimplantation können zu einer Schwächung der Oxidbereiche in der Nähe der Diffusionsgebiete führen.

Ähnlich zu den beschriebenen Beeinträchtigungen der Gateoxidrandgebiete längs des Kanals, wird hier ein von der Länge des Kanals unabhängiger Anteil des Oxides am Anfang und Ende des Kanalbereiches geschädigt. Analog zu den STI-geschädigten Bereichen, die sich besonders bei schmalen Transistoren auswirken, führt der Einfluss der Bereiche in der Nähe der Diffusionsgebiete besonders bei sehr kurzen Kanallängen zu einer Verstärkung der Degradation bei BTS.

In Abbildung 2.1 sind die Betriebsbedingungen eines p-Kanal-Transistors skizziert, welche eine BT-Belastung hervorrufen können. Am Gateanschluss des Transistors liegt gegenüber den Source- und Drainknoten ein betragsmäßig relativ großes negatives Potenzial an. Dies führt zu einem großen Gateoxidfeld. Eine horizontale Feldkomponente herrscht in diesem Fall nicht vor, die Belastungsbedingung kann dementsprechend als homogen bezeichnet werden. Das elektrische Feld führt in unmittelbarer Nähe der Grenzfläche zwischen Silizium und Gateoxid im Silizium zu starker Inversion, d.h., mit Hilfe von Löchern wird ein leitender Kanal gebildet, es fließt jedoch kein Strom zwischen Source und Drain, da diese Transistorknoten auf gleichem Potenzial liegen.

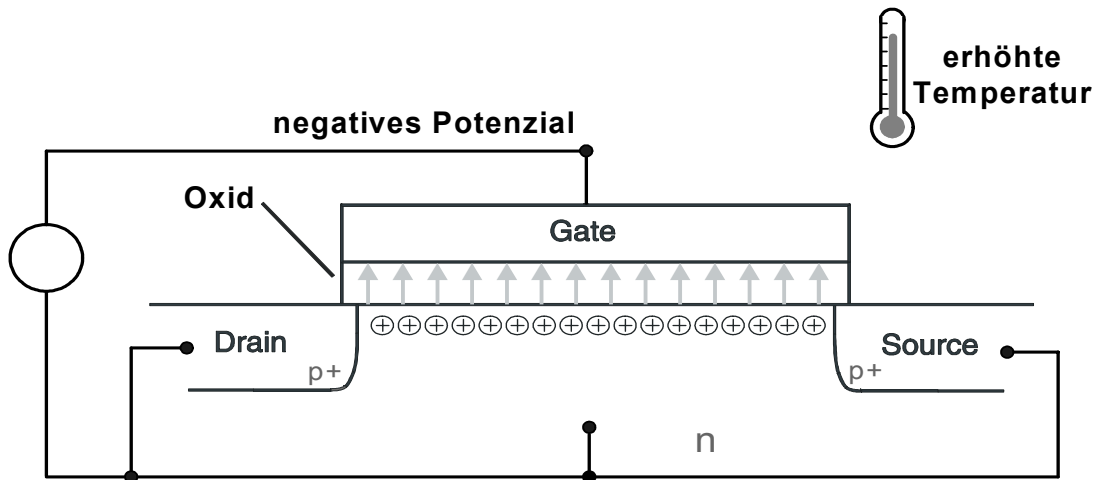


Abbildung 2.1: Bei homogenem Negative-Bias-Temperature-Stress vorherrschende Betriebsbedingungen beim p-Kanal MOS-Transistors

Bedingt z.B. durch die immer schwieriger abzuführende Verlustwärme erreichen heutige Schaltungen hohe Temperaturen. Aber auch äußerlich einwirkende Hitzequellen führen zu einem Anstieg der Chiptemperatur. Dem Einfluss von Sonnenbestrahlung ausgesetzte mikroelektronische Schaltungen können ebenfalls Werte von über 80°C erreichen. Liegt sowohl das zuerst beschriebene elektrische Feld als auch eine erhöhte Temperatur gleichzeitig vor, ist die Bedingung für Bias-Temperature-Stress erfüllt. Die Transistoren einer Schaltung, die dieser Belastungssituation ausgesetzt sind, erleiden Degradationen ihrer elektrischen Parameter. Diese Belastungsbedingungen kommen innerhalb des normalen Betriebes realer und häufig eingesetzter Schaltungen vor.

In den folgenden Unterkapiteln 2.1.1 und 2.1.2 werden die Belastungsbedingungen für homogene BT-Belastungen mithilfe von Beispielen digitaler und analoger Schaltungen verdeutlicht. In Kapitel 2.1.3 werden darüber hinaus Bedingungen erörtert, die zu einer BT-Belastung führen können, die von dem beschriebenen homogenen Fall abweicht. Die Voraussetzungen für diese bislang in der Literatur nicht beschriebene inhomogene BT-Belastungssituation werden ebenfalls anhand von Beispielschaltungen für Analog- und Hochfrequenzapplikationen erläutert. Die Auswirkungen der Degradation nach BTS werden in Kapitel 2.3 und 2.5 behandelt.

2.1.1 Homogene Belastungsbedingungen in Digitalschaltungen

Abbildung 2.2 zeigt eine typische CMOS-Inverterschaltung. Die Gateanschlüsse beider Transistoren sind miteinander verbunden und dienen als gemeinsamer Eingang. Je nach Eingangsspannung agiert jeweils ein Transistor als Schaltelement und der jeweils andere als Lastelement [Schumacher87].

Liegt am Eingang der Inverterschaltung ein niedriges Potenzial an (Massepotenzial), leitet der p-Kanal-Transistor und der n-MOS Transistor sperrt. Am Ausgang des Inverters steht nun die Betriebsspannung an, d.h., am p-Kanal-Transistor liegen in diesem Schaltzustand des Inverters sowohl an den Knoten Drain und Source als auch am Substrat jeweils Betriebsspannung an, während das Gatepotenzial auf Masse liegt. Der p-MOS Transistor befindet sich folglich in starker Inversion.

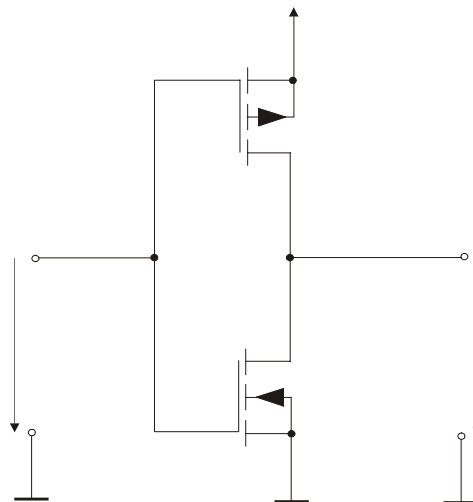


Abbildung 2.2: CMOS Inverter

Wenn zusätzlich eine hohe Temperatur vorhanden ist, sind die Bedingungen für Bias-Temperature-Stress erfüllt. Befindet sich der Transistor über längere Zeit in einem derartigen Belastungszustand, führt dies zu einer Degradation seiner elektrischen Eigenschaften. Die Schaltungsparameter des Inverters verändern sich dadurch und können die Funktion der gesamten Schaltung beeinträchtigen, in der diese Baugruppe eingesetzt ist (siehe Kapitel 2.5.1).

2.1.2 Homogene Belastungsbedingungen in Analogschaltungen

In Abbildung 2.3 ist ein zweistufiger CMOS-Operationsverstärker mit CMOS-Differenzstufe mit Stromspiegellast¹ dargestellt. Die Funktion der Schaltung liegt in der Verstärkung der Differenz der beiden zugeführten Eingangssignale. Die Differenzstufe ist die erste verstärkende Stufe und – in verschiedenen Ausführungsformen – grundsätzlicher Bestandteil von Operationsverstärkern. Ihre Funktion soll kurz erläutert werden: Die n-Kanal Transistoren M1 und M2 realisieren einen Stromspiegel, dessen Ein- und Ausgangstransistor innerhalb der zwei Inverter (M1, M3 und M2, M4) als Last dient. Der Transistor M5 stellt eine Konstantstromquelle dar. Mit seiner Hilfe wird ein definierter Strom in die zusammengeführten Sourceknoten der Eingangstransistoren M3 und M4 eingepreßt.

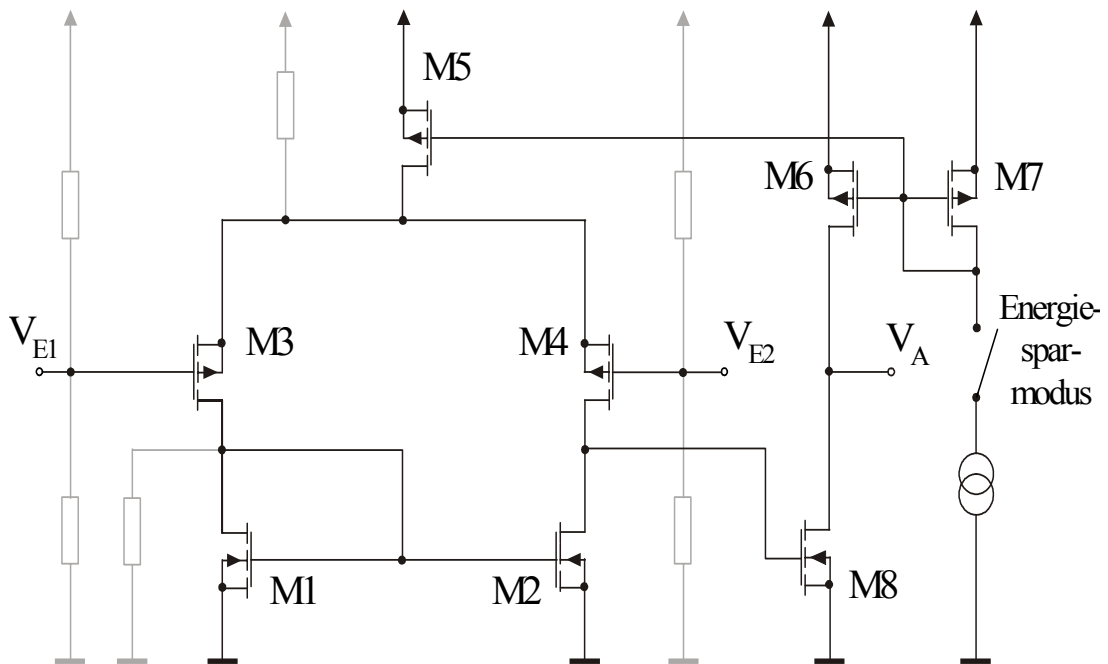


Abbildung 2.3: Zweistufiger CMOS-Operationsverstärker mit CMOS-Differenzstufe mit Stromspiegellast. Pfade für parasitäre Leck- und Sperrströme, die für das Auftreten von BT-Belastung eine Rolle spielen, werden durch die grau eingezeichneten Widerstände verdeutlicht.

¹ In der Literatur findet sich auch häufig der Name 'single-ended' für eine Differenzstufe dieser Ausführungsform.

Unter der Voraussetzung, dass die Transistoren beider Inverter identisch sind und die Gatespannung von M3 und M4 gleich sind, teilt sich der Strom exakt zur Hälfte in beide Zweige auf. Auf diesem Wege stellt sich an den Lastelementen die gleiche Spannung ein. Wird die Gatespannung an nur einem einzelnen Eingangselement variiert (V_{E1} oder V_{E2}), verändert sich der Strom innerhalb des Zweiges und der Spannungsabfall über dem Lastelement nimmt zu oder ab.

Da der zur Verfügung stehende Strom durch die Konstantstromquelle vorgegeben ist, ändert sich der Strom innerhalb des anderen Zweiges ebenfalls entsprechend. Diese Stromänderungen werden am gemeinsamen Drainknoten von M2 und M4 in Spannungsänderungen umgesetzt. Durch die bei der abgebildeten Differenzstufe eingesetzte Stromspiegellast wird am Ausgang ein bereits einphasiges Signal mit hoher Stufenverstärkung bereitgestellt.

Wird beispielsweise die Eingangsspannung V_{E1} variiert, wird das aus den Differenzströmen entstehende Signal noch zusätzlich über die Stromspiegelanordnung verstärkt in den rechten Zweig übertragen. Liegt dagegen ein Gleichtaktsignal an beiden Eingängen der Differenzstufe an, kommt es zu keiner Änderung in den beiden Signalzweigen. Eine ideale Differenzstufe verstärkt nur Differenzsignale unabhängig vom jeweiligen Gleichtaktpegel [Schumacher87].

Eine reale Differenzstufe weist dagegen Ausgangsspannungsfehler auf, die sich auf Parametervariationen („Mismatch“) zwischen den Transistoren der idealerweise identischen Inverter zurückführen lassen. Diese Abweichungen führen dazu, dass der Strom der Quelle sich unter der Bedingung $V_{E1} = V_{E2}$ nicht exakt in beide Zweige aufteilt. Bereits ohne ein angelegtes Differenzsignal herrscht so eine Unsymmetrie in den beiden Signalzweigen vor.

Dieser Fehler lässt sich durch die Eingangsoffsetspannung beschreiben. Diese Spannung gibt den Wert an, der der Eingangsdifferenzspannung überlagert werden muss, damit das Ausgangssignal null ist bzw. auf dem Wert seines Ruhearbeitspunktes liegt. Wird die Differenzstufe in einer monolithischen Technik hergestellt, lassen sich Transistorpaare in hoher Qualität herstellen. Das Mismatchverhalten lässt sich durch Layout-Maßnahmen, z.B. durch Vergrößerung der Fläche und durch Kreuzkoppelungen, weiter reduzieren [Baker98, Gregor92, Linnenbank98]. Ganz vermeiden lässt es sich jedoch aus prinzipiellen statistischen Gründen nicht [Laksh86, Pelgrom89, Oehm93, Stolk98].

Die Symmetrie einer Differenzstufe kann jedoch durch BT-Belastungen massiv beeinträchtigt werden. Um Energie zu sparen, werden in vielen Anwendungen (Bsp.: Telekommunikationsbausteine) Schaltungskomponenten bei vorübergehender

Nichtbenutzung in einen ‚Energiesparmodus‘ versetzt. In der Literatur wird dieser Schaltungszustand oft auch als ‚Power-Down-Modus‘ bezeichnet. Dabei wird nicht die Betriebsspannung der Schaltung ausgeschaltet, was bei Wiedereinschalten erhebliche Zeit und Leistungsverluste durch Umladung der Versorgungsleitungen mit sich bringen würde, sondern es wird wie angedeutet durch den Schalter in Abbildung 2.3 der Bias-Strom der Schaltung abgeschaltet. Da die gesamte Schaltung dann – von Leckströmen abgesehen – querstromfrei ist, erzeugt sie eine geringere Verlustleistung. Bei einem Einsatz dieser Schaltungskomponente in einem mobilen elektronischen Gerät ermöglicht dieser ‚Power-Down-Modus‘ eine Verlängerung der Betriebszeiten des Gerätes bei Batteriebetrieb.

Die auch im ‚Power-Down-Modus‘ anliegende Betriebsspannung führt zu parasitären Stromflüssen in der Schaltung. Mögliche Pfade dieses Stromes, der durch geringe Leck- bzw. Sperrströme der Transistoren hervorgerufen wird, sind in Abbildung 2.3 durch die grau eingezeichneten Widerstände angedeutet. Auf diesem Wege können sich an allen Schaltungsknoten, so auch an den Source-, Drain- und Substratknoten der p-Kanal-Eingangstransistoren (M3, M4) unterschiedliche Potenziale aufbauen. Da die Gateanschlüsse der Transistoren mit benachbarten Schaltungsteilen verbunden sind, können sich auch dort innerhalb des Power-Down-Modus verschiedene Potenziale einstellen.

Die Belastungsbedingung der Eingangstransistoren kann dementsprechend auch asymmetrisch ausfallen. Wird einer der Transistoren oder beide aufgrund der anliegenden Spannungen in Inversion getrieben, können die in Kapitel 2.1 beschriebenen BTS-Voraussetzungen vorliegen, falls eine erhöhte Temperatur durch externe Einflüsse gegeben ist. Die verschiedenartigen Belastungsbedingungen, die die beiden Eingangselemente der Differenzstufe vorfinden, können dann zu einem asymmetrischen Degradationsverhalten führen.

Während ein Bauelement eventuell nur leicht oder gar nicht einer BT-Belastung ausgesetzt ist, kann der andere Eingangstransistor stark geschädigt werden. Auf die Auswirkungen dieser nicht einheitlichen Transistordegradation auf die Funktion der Differenzstufe wird im Kapitel 2.5.2 eingegangen.

2.1.3 Inhomogene Belastungsbedingungen in Analog- und HF-Schaltungen

Bei der bisherigen Betrachtung von Bias-Temperature-Stress wird in der Literatur während der Belastung immer von einem ausschließlich vertikalen elektrischen Feld über dem Gateoxid des Transistors ausgegangen. An den Source- und Drainanschlüssen des Bauelementes liegt dabei ein identisches Potenzial und es fließt kein Strom zwischen Drain und Source. In Abbildung 2.4a) werden diese Feldverhältnisse anhand eines skizzierten Querschnittes eines p-Kanal Transistors verdeutlicht.

Neben diesen beschriebenen homogenen Belastungsbedingungen mit ausschließlich vertikalem elektrischen Feld können in realen Schaltungsanwendungen, insbesondere bei bestimmten Analog- und Hochfrequenz-Applikationen, jedoch auch inhomogene Fälle auftreten.

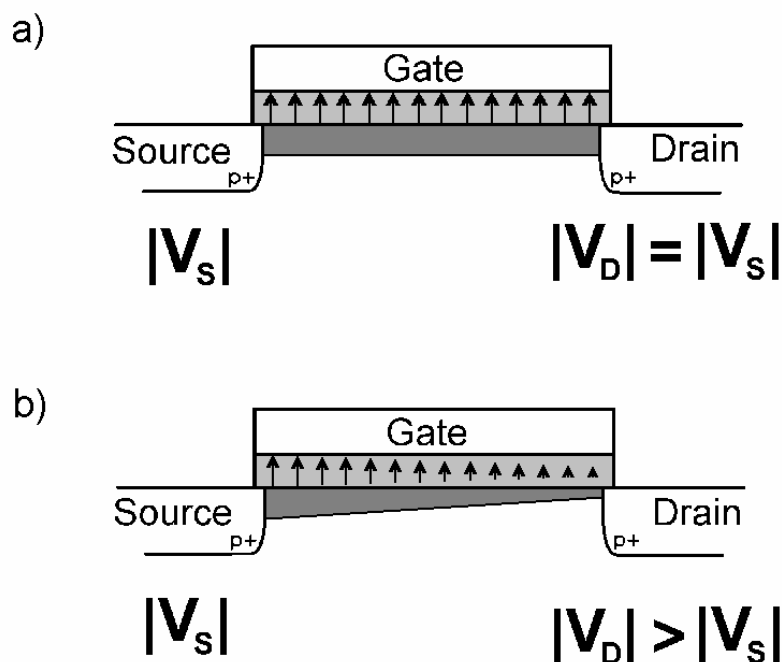


Abbildung 2.4: Schematischer Querschnitt eines p-MOSFETs

a): unter homogenen BT-Belastungsbedingungen mit $V_D = V_S$ und

b): unter inhomogenen BT-Belastungsbedingungen mit $|V_D| > |V_S|$

Tritt an den Transistorknoten Drain und Source während der Belastungssituation jedoch eine Potenzialdifferenz auf, entsteht zusätzlich zu dem rein vertikalen elektrischen Feld auch eine horizontale Feldkomponente. Der Transistor befindet sich dann in Arbeitspunkten mit hohen Gatespannungen und moderaten Drain-zu-Source Spannungen. Dieser Fall ist schematisch in Abbildung 2.4b) dargestellt.

Die in Abbildung 2.3 auf Seite 8 gezeigte Schaltung des zweistufigen Operationsverstärkers kann auch hier als Beispiel für eine solche Belastungssituation herangezogen werden. Wird die Schaltung als zeitkontinuierlicher Komparator mit durchgängig eingeschaltetem Bias-Strom eingesetzt, können derartige inhomogene Belastungszustände auftreten:

Werden stark unterschiedliche Eingangssignale an die Transistoren M3 und M4 gelegt, sodass dadurch z.B. der Transistor M4 ausgeschaltet wird, wird der Transistor M3 von dem gesamten Strom durchflossen, den die Stromquelle M5 liefert, und arbeitet daher im Sättigungs- oder Triodenbereich. Wenn die Schaltung gleichzeitig einer erhöhten Temperatur ausgesetzt ist, ist dieser Transistor in dieser Situation einer BT-Belastung ausgesetzt. Dies führt zu einer Degradation seiner elektrischen Parameter.

Inhomogene BT-Belastungszustände können auch noch in anderen Applikationen beobachtet werden. In einigen Hochfrequenz-CMOS-Schaltungen z.B. arbeiten die Transistoren mit einer hohen Frequenz mit großen Signalen. In Abbildung 2.5 ist ein typischer spannungsgesteuerter Oszillator (VCO) als ein Beispiel für solche Schaltungen dargestellt. Anwendungen mit erhöhten Temperaturen führen auch hier zu inhomogenen BT-Belastungen. In diesem Fall sind die Transistoren M1 und M2 der Belastungssituation ausgesetzt. Es liegen hohe Gatespannungen an bei gleichzeitig moderaten Drain-zu-Source-Spannungen. Unter dieser Belastung kommt es zu einer Degradation der elektrischen Parameter dieser Transistoren.

Auf die Auswirkungen der Degradation nach inhomogenen BT-Belastungen auf die Funktion der beschriebenen Analog- und Hochfrequenzschaltungen wird in Kapitel 2.5.3 eingegangen.

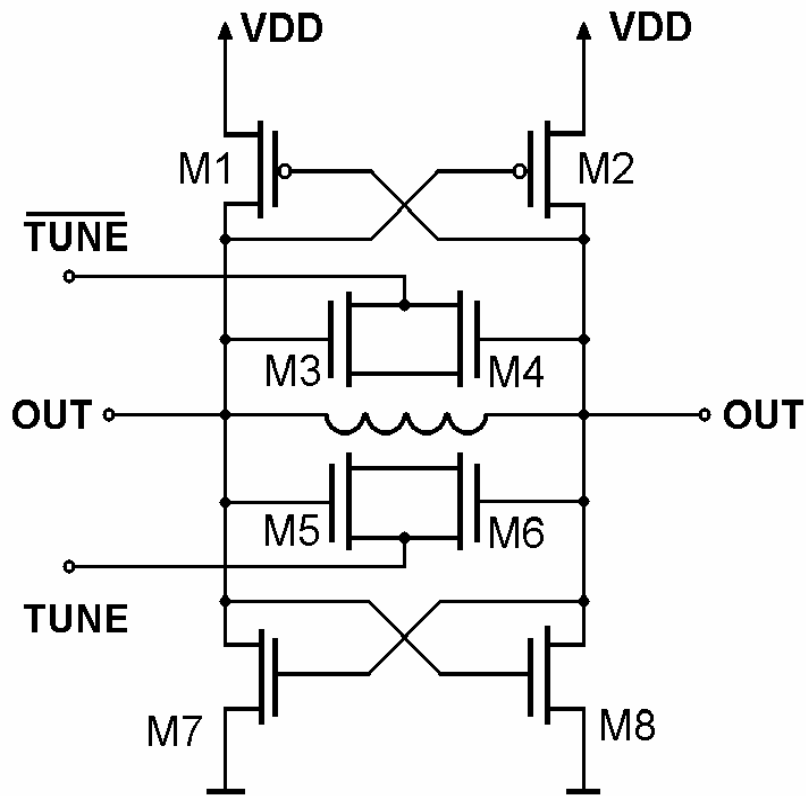


Abbildung 2.5: Spannungsgesteuerter Oszillator (VCO). Die Transistoren M1 und M2 können im Betrieb inhomogenen BT-Belastungszuständen ausgesetzt sein.

2.2 Physikalische Mechanismen bei Belastung

Die durch BT-Belastungen hervorgerufenen Degradationserscheinungen von p-Kanal Transistoren basieren auf mikroskopischen Änderungen der Beschaffenheit des Gateoxides und der Grenzfläche zwischen dieser isolierenden Schicht und dem Siliziumsubstrat. Das vorherrschende elektrische Feld und die erhöhte Temperatur während der Belastungssituation führen zu Schäden in unmittelbarer Nähe des Si-SiO₂-Überganges. In der Literatur werden hauptsächlich zwei Arten von Schädigungen für die Degradationserscheinungen von p-Kanal Transistoren verantwortlich gemacht. Beide Schädigungsmechanismen sollen in diesem Kapitel vorgestellt werden.

2.2.1 Belastungsinduzierte Grenzflächenzustände

Die für die elementare Funktionsweise des MOS-Transistors notwendige Isolierung der Steuerelektrode vom Substratmaterial wird durch eine Isolationsschicht aus Siliziumdioxid erreicht. Bei modernen Technologien enthält die Trennschicht für eine Verbesserung der dielektrischen Eigenschaften oft Zusätze wie z.B. Stickstoff. Diese dünne Schicht wird im Regelfall durch thermische Oxidation des Siliziums gebildet. An der Grenzfläche zwischen Si-Substrat und Oxid geht das periodische Gitter des Si-Einkristalls in einer einige Atomlagen dicken Schicht in die amorphe Struktur des Siliziumdioxides über (siehe Abbildung 2.6).

In dieser Grenzschicht kommt es zur Ausbildung von donator- und akzeptorartigen Störstellen, da die unterschiedlichen Gitterstrukturen der Materialien zum Auftreten von nicht abgesättigten Bindungen führen („dangling bonds“). Es ist unvermeidlich, dass die Grenzflächenschicht bereits direkt nach der Prozessierung derartige so genannte Grenzflächenzustände enthält. Grenzflächenzustände können Ladungsträger einfangen, für eine gewisse Zeit binden und wieder emittieren.

Da die elektrische Performance des Bauelementes durch Grenzflächenzustände verringert wird, wird versucht, die Anzahl der Grenzflächezustände bzw. deren Dichte möglichst gering zu halten. Nach dem eigentlichen thermischen Aufwachsen des Oxides werden dafür bei späteren Ausheilschritten diese Grenzflächenzustände zu einem großen Teil elektrisch neutralisiert.

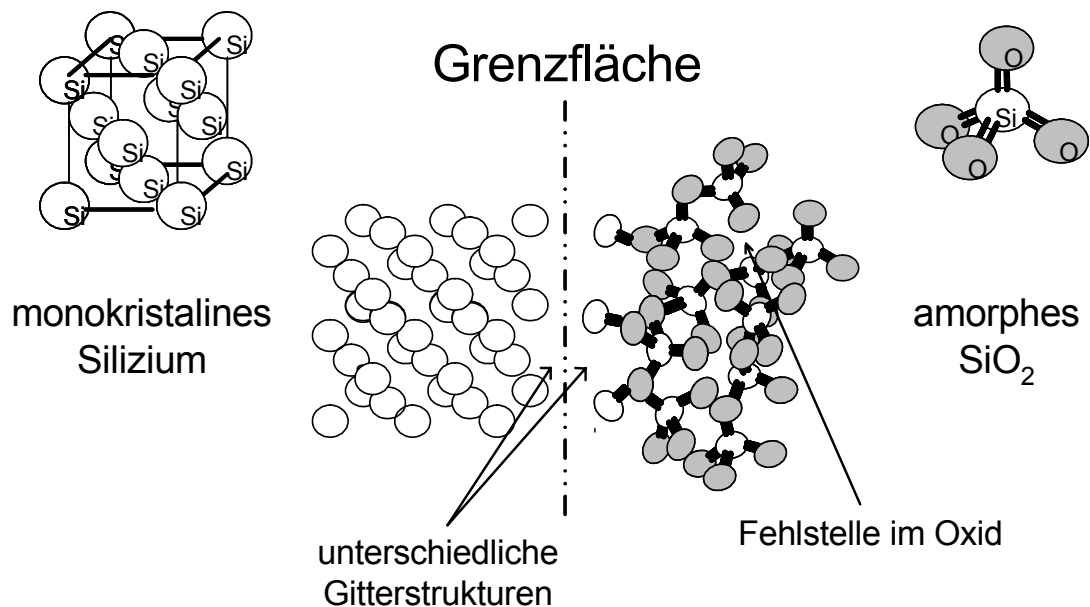


Abbildung 2.6: Schematische Darstellung der Struktur von monokristallinem Silizium des Substrats und amorphem Gateoxid. Unterschiedliche Gitterkonstanten führen zu nicht abgesättigten Bindungen [Denais04].

Während der dem Gateoxidprozess und anderen späteren Prozessschritten folgenden Temperaturbehandlungen, die auch in Wasserstoffatmosphäre erfolgen können, diffundiert Wasserstoff aus verschiedenen Reservoirs an die Grenzfläche. Die Quelle des Wasserstoffs kann sowohl das eingesetzte Prozessgas darstellen, als auch umliegende Isolations- oder Passivierungsschichten, die bereits überschüssigen Wasserstoff enthalten können. Die dabei entstehenden Bindungen stellen keine Fangstellen für Ladungsträger mehr dar, sie sind elektrisch neutral und beeinflussen daher die elektrischen Eigenschaften des Bauelementes nicht mehr. Die Anzahl der Grenzflächenzustände eines Transistors kann mit Hilfe von Wasserstoff stark reduziert werden, eine vollständige Neutralisierung aller offenen Bindungen gelingt jedoch nicht.

In Abbildung 2.7 ist die Anordnung von Grenzflächenzuständen in der Grenzschicht zwischen Siliziumsubstrat und Gateoxid eines Transistors skizziert. Die kleinen Quadrate symbolisieren donatorartige Grenzflächenzustände, die mit positiven Ladungsträgern wechselwirken. Unter den in Kapitel 2.1 beschriebenen Belastungsbedingungen können die mit Wasserstoff abgesättigten Bindungen

sukzessive wieder aufbrechen und neue offene Bindungen werden erzeugt. Der frei werdende Wasserstoff diffundiert von der Bruchstelle weg. Die nun offene Bindung kann daraufhin wieder als Grenzflächenzustand agieren, Ladungsträger zeitlich binden und wieder emittieren. [Jeppson77, Ogawa95, Ogawa96].

2.2.2 Einbau fester Ladungen in das Gateoxid

Idealerweise ist die isolierende Oxidschicht zwischen der Gateelektrode und dem Siliziumsubstrat elektrisch neutral. In Defekten innerhalb des Siliziumoxides können jedoch Ladungsträger eingefangen werden und dort verbleiben. Im Gegensatz zu den Umladevorgängen der zuvor beschriebenen Grenzflächenzustände mit relativ geringen Zeitkonstanten ist der Einbau dieser festen Ladungen in das Oxidmaterial jedoch durch wesentlich größere Zeitkonstanten charakterisiert. Die meisten dieser Ladungen verbleiben auch bei einem deaktivierten Transistor im Gateoxid.

In Abbildung 2.7 sind derartige feste positive Ladungen und ihre Position in unmittelbarer Nähe zur SiO_2 -Grenzschicht mithilfe kleiner Kreise skizziert. Neben der bereits beschriebenen Erhöhung der Grenzflächenzustandsdichte führen BT-Belastungen auch zum Einbau fester Ladungen in das Oxid [Schlünder99].

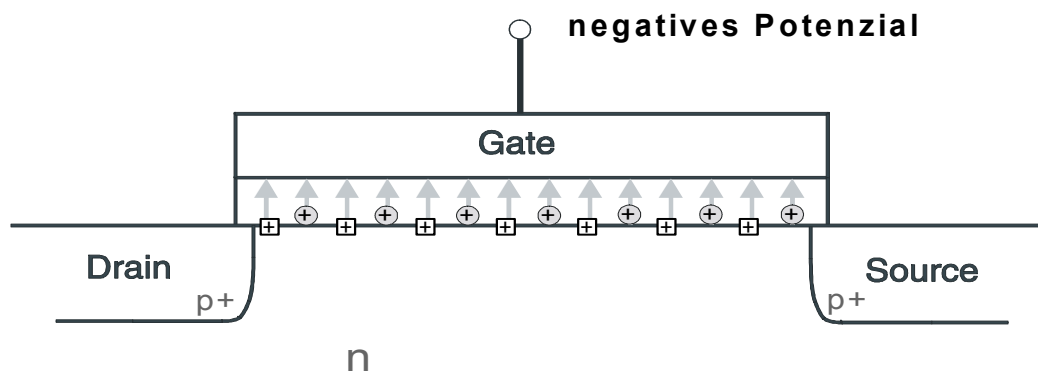


Abbildung 2.7: Positiv besetzte Grenzflächezustände und feste positive Ladungen innerhalb des Oxides wirken der Gatespannung eines p-Kanal Transistors entgegen.

2.3 Degradation von Transistoreigenschaften

Positiv besetzte Grenzflächenzustände wirken der Gatespannung eines p-Kanal MOS-Transistors entgegen. Sie decken einen Teil der Gateladungen ab, die dann für die Steuerung des Kanals nicht mehr zur Verfügung stehen. Als Folge nimmt der Betrag der Schwellenspannung zu. Zusätzlich kann die Beweglichkeit der Ladungsträger im Kanal vermindert werden, da der belastungsinduzierte Anstieg der Grenzflächenzustandsdichte zu einer stärkeren Streuung der Kanalladungsträger entlang der Grenzfläche führen kann.

Eingebaute positive Oxidladungen wirken sich ähnlich zu dem Einfluss eines besetzten donatorartigen Grenzflächenzustandes aus. Die positiven Ladungsträger schirmen den Stromfluss des Transistors steuernde negative Ladungen auf der Gateelektrode ab. Sie wirken also ebenfalls der Gatespannung entgegen. Folglich nimmt der Betrag der Schwellenspannung auch durch diesen Mechanismus zu.

Zusammenfassend lässt sich sagen, dass die während einer BT-Belastung entstehenden Schäden zu einer Verschiebung der Einsatzspannung zu betragsmäßig größeren Werten, zu einer Beeinträchtigung der Beweglichkeit der Ladungsträger und damit zu einer Drainstromverminderung führen.

2.4 Zeitlicher Verlauf der Degradation

Der zeitliche Verlauf der Degradation der elektrischen Parameter nach BTS folgt in guter Näherung einem Potenzgesetz. Wird die Parameterverschiebung als Funktion der Zeit in ein Diagramm mit logarithmischer Skalierung sowohl für die x-Achse als auch für die y-Achse aufgetragen, kann der Verlauf der Kurve durch eine Regressionsgerade angenähert werden.

Als Beispiel wird in Abbildung 2.8 die Verschiebung der Einsatzspannung als Funktion der Belastungszeit aufgetragen. Der Transistor wird bei $T_{stress} = 125^{\circ}\text{C}$ einer Belastungsspannung von $-2,4\text{V}$ am Gate ausgesetzt. Die Steigung der eingezeichneten Regressionsgerade bzw. der Wert des Exponenten der Funktion beträgt hier 0,247.

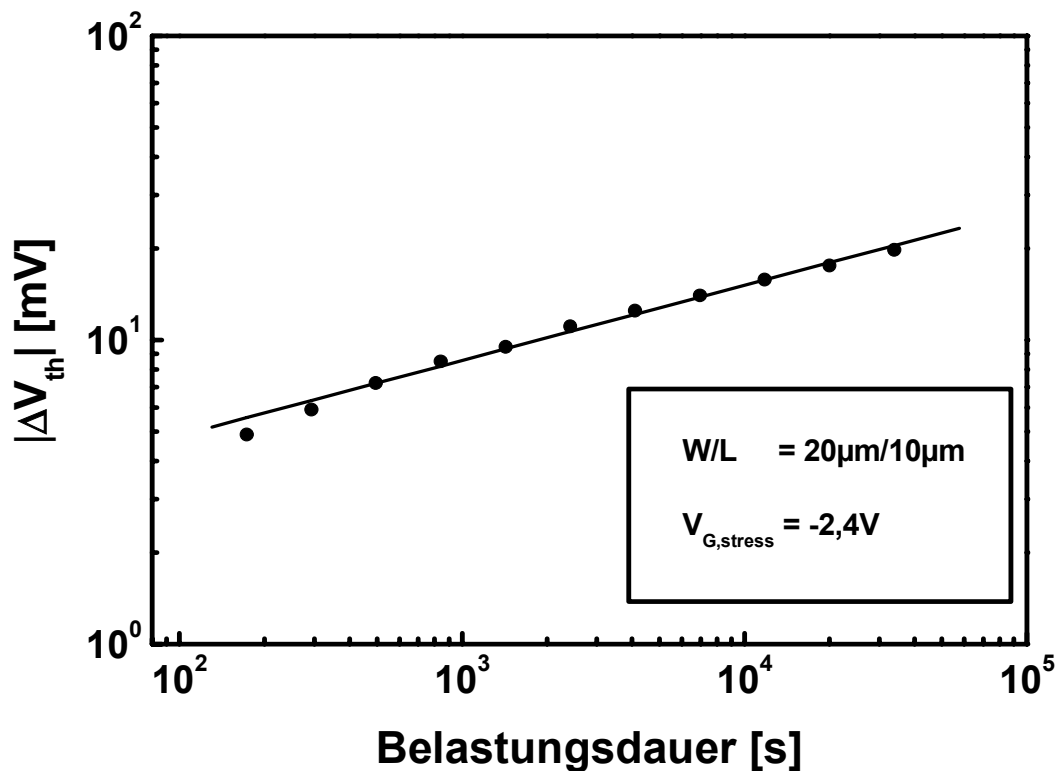


Abbildung 2.8: Degradation der Einsatzspannung als Funktion der Belastungsdauer. Die Punkte entsprechen den Messwerten, die Regressionsgerade wird mithilfe einer Potenzialfunktion gebildet. Technologie T10.

Dieses Steigungsverhalten lässt sich auf die zugrunde liegenden Schädigungsmechanismen zurückführen [Jeppson77, Ogawa95, Schlünder99, Chakravarthi04]. Der zeitliche Verlauf der Degradation ergibt sich aus dem Reaktionsverlauf des Aufbrechens der mit Wasserstoff abgesättigten Bindungen. Der dabei frei werdende mobile Wasserstoff kann in verschiedenen Zuständen auftreten. In Abhängigkeit davon ergeben sich unterschiedliche Zeitkonstanten für die einzelnen verschiedenen Reaktionsverläufe [Chakravarthi04].

In Tabelle 2.1 werden die möglichen Reaktionen und die damit verbundenen Zeitkonstanten der Degradation aufgelistet. Atomarer Wasserstoff (Modell I) führt zu der von Jeppson et al. und Ogawa et al. beschriebenen Steigung von 0,25 [Jeppson77, Ogawa95], molekularer Wasserstoff (Modell II) dagegen nur zu einer Steigung von 0,165 [Brower88, Stesmans00].

Tabelle 2.1: Mögliche stabile Formen des nach dem belastungsinduzierten Aufbrechen der Bindungen frei werdenden Wasserstoffes und Steigung einer Regressionsgeraden bei doppelt logarithmischer Darstellung der Degradation als Funktion der Zeit [Chakravarthi04]:

Modell	Reaktion	Wasserstoff-Konfiguration	Steigung der Regressionsgeraden
I	$\text{SiH} + \text{h}^+ \Leftrightarrow \text{Si}^+ + \text{H}^0$	H^0	0,25
II	$\text{SiH} + \text{h}^+ \Leftrightarrow \text{Si}^+ + 0,5 \text{H}_2^0$	H_2	0,165
III	Modell I und II	H^0, H_2	0,165 – 0,25
IV	$\text{SiH} + 2 \text{h}^+ \Leftrightarrow \text{Si}^+ + \text{H}^0$	H^+	0,25 – 0,5
V	Modell I, II und III	$\text{H}^0, \text{H}^+, \text{H}_2$	0,165 – 0,5
VI	Modell I und Einbau von H	H^0	> 0,25
VII	Modell I und Ausbau von H	H^0	< 0,25
VIII	Modell II und Einbau von H_2	H_2^0	> 0,165
IX	Modell II und Ausbau von H_2	H_2^0	< 0,165

Entstehen bei dem Aufbrechen der Brücken Wasserstoff-Protonen (Modell IV) diffundieren diese nicht nur von der Bruchstelle weg, sondern werden darüber hinaus durch das über dem Oxid vorherrschende elektrische Feld beschleunigt [Rashkeev01]. Die Steigung der Degradation kann in diesem Fall 0,25-0,5 betragen. Die bei unterschiedlichen BT-Belastungsexperimenten tatsächlich beobachteten Steigungen lassen auf eine Kombination der verschiedenen Reaktionsverläufe schließen. Auch der Einbau und Ausbau von Wasserstoff in Gitterfehlstellen innerhalb des Oxids beeinflusst den zeitlichen Verlauf der Degradation (Modell VI-IX). Unterschiedliche Anteile der verschiedenen Reaktionen am Schädigungsprozess führen zu abweichenden Steigungen. Eine derartige Überlagerung der Reaktionsverläufe kann jedoch nicht mit Überlagerungen von einfachen Potenzfunktionen beschrieben werden. Die zusammengesetzten Funktionen wiesen dabei immer einen größeren Exponenten als die der Einzelfunktionen auf. Die Steigung des zeitlichen Degradationsverlaufes könnte so nur zunehmen. Der größte Teil der in der Literatur veröffentlichten Daten und auch die für diese Arbeit experimentell ermittelten Daten zeigen Degradationsverläufe mit Steigungen im Bereich von 0,2 bis 0,3. Dies erlaubt folgende Schlussfolgerung:

Die unter BT-Belastung einsetzende Degradationsreaktion wird dominiert durch frei werdende neutrale Wasserstoffkonfigurationen (z.B. H^0) mit geringeren Anteilen von molekularem Wasserstoff, positiv geladenen Wasserstoff-Protonen und dem Einbau von Wasserstoff.

Neben diesem generellen zeitlichen Verlauf der Degradation kann darüber hinaus häufig eine Änderung der Steigung während der Belastung beobachtet werden. Mit längeren Belastungsdauern nimmt die Steigung der Degradation typischerweise leicht aber kontinuierlich ab. Das deutet auf eine Tendenz zur Sättigung der Degradation hin [Aono04, Chakravarthi04]. Dies kann ebenfalls nicht auf der Basis eines einfachen Potenzgesetzes beschrieben werden. Eine mögliche Erklärung für die fortschreitende Abschwächung des Degradationsprozesses kann in einer Änderung der beschriebenen Anteile der Reaktionspfade liegen. Ein weiterer Erklärungsansatz baut auf eine mögliche Diffusionsbarriere für Wasserstoff an der Grenzfläche zum polykristallinen Silizium der Gateelektrode, die den Transfer des Wasserstoffs im Laufe der Zeit reduzieren kann und so den Degradationsprozess verlangsamt. [Chakravarthi04, Chakravarthi03].

2.5 Auswirkung der Degradation auf Schaltungseigenschaften

Nachdem die Degradation eines p-Kanal-Transistors auf phänomenologischer Ebene beschrieben und einfache Beispiele für BT-Belastungssituationen innerhalb häufig eingesetzter Schaltungen vorgestellt wurden, sollen nun die Konsequenzen der Transistordegradation auf Schaltungsebene erläutert werden.

Die degradationsbedingten Veränderungen lassen sich aufgrund ihrer weitgehenden Längen- und Weitenunabhängigkeit in einem einfachen Modell erster Ordnung durch eine Verschiebung der Schwellenspannung ausdrücken (vgl. [Jeppson77, Ogawa95, Ogawa96]). In Abbildung 2.9 ist dies anhand eines Ersatzschaltbildes eines p-Kanal Transistors nach BT-Belastung dargestellt. Die zusätzliche Spannungsquelle ΔV_{th} wirkt der angelegten Gatespannung entgegen. Die erwähnte mögliche Beweglichkeitsreduktion wird bei diesem einfachen Modell vernachlässigt.

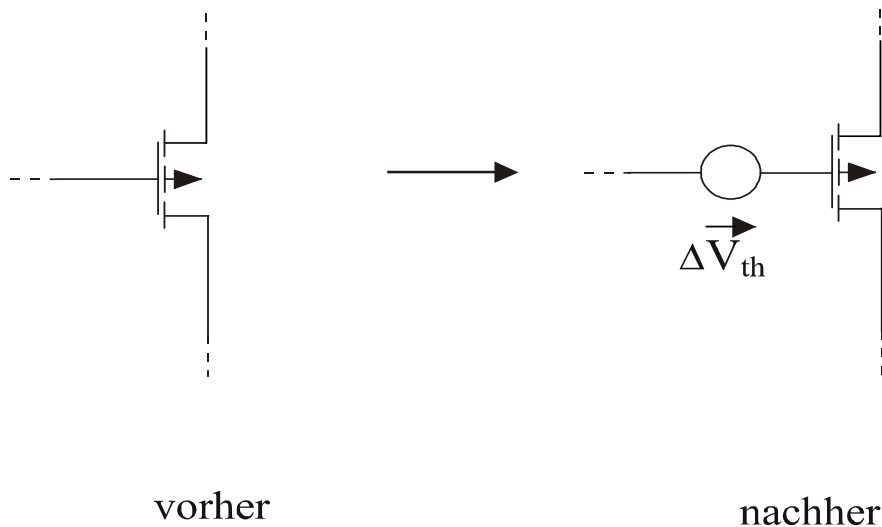


Abbildung 2.9: Einfache Modellierung der Auswirkung der BTS-induzierten Degradation mithilfe eines Ersatzschaltbildes

Die effektive Gatespannung ist gegenüber dem ungestressten Transistor reduziert, was dazu führt, dass der Transistor bei gleicher angelegter Gatespannung nach BT-Belastungen weniger Strom leitet. Diese Stromdegradation wirkt sich je nach Aufgabe des Transistors innerhalb einer Schaltung unterschiedlich aus. Insbesondere ist die Degradation nach BTS für analoge und für digitale Schaltungsanwendungen gesondert zu berücksichtigen.

Die durch BT-Belastungen hervorgerufenen Parameterdegradationen sind bei modernen CMOS-Technologien (unter 180nm) mittlerweile so groß, dass die BTS-Problematik nicht mehr nur bei Anwendungen eine Rolle spielt, die besondere Anforderungen z.B. an das Matching zweier Transistoren stellen und daher besonders empfindlich auf belastungsinduzierte Parameterverschiebungen der Transistoren untereinander reagieren. Die Verschiebung der Einsatzspannung und der Verlust an Stromtreibfähigkeit nach BTS können auch für einzelne, nicht paarweise eingesetzte Bauelemente so große Werte annehmen, dass die Funktion der Schaltung beeinträchtigt werden kann.

2.5.1 Auswirkung der Degradation nach homogener Belastung in Digitalschaltungen

Die Hauptanforderungen, die an Transistoren innerhalb einer Digitalanwendung gestellt werden, liegen in deren Stromtreibfähigkeit mit dem Ziel hoher möglicher Taktfrequenzen und geringer Laufzeiten. Um dieses Ziel zu erreichen, wird der Kanalbereich der Transistoren in Logikschaltung üblicherweise in Minimallänge ausgeführt. Eine Reduzierung des Stromes, den der Transistor bei gleicher Gatespannung führt, wirkt sich auf die geforderten Eigenschaften der Schaltung aus. Die Schaltverzögerungen nehmen bedingt durch die bei geringeren Strömen größeren Umladezeiten von Kapazitäten zu. Die Fähigkeit, nachfolgende Stufen zu treiben, nimmt ab, sodass beim Schaltungsentwurf geforderte Gatterlaufzeiten nicht mehr eingehalten werden.

Da die Transistoren in den unterschiedlichen Gattern darüber hinaus nicht notwendigerweise gleichermaßen degradieren (z.B. aufgrund unterschiedlicher anliegender Datenmuster), kann es zu einer Verschiebung der Gatterlaufzeiten untereinander kommen, die durch Toleranzwerte (Margins) beim Schaltungsentwurf berücksichtigt werden müssen. Die maximal mögliche Performance der Technologie kann also nicht genutzt werden.

2.5.2 Auswirkung der Degradation nach homogener Belastung in Analogschaltungen

Die Anforderungen, die an Transistoren innerhalb einer Analoganwendung gestellt werden, sind wesentlich breiter gefächert [Thewes97_2, Thewes99_1]. Die tolerierbaren Abweichungen der Parameter Schwellenspannung (V_{th}), Drainstrom (I_D), Steilheit (g_m) und differentieller Drain-Source-Leitwert (g_{DS}) sind geringer. Für Analoganwendungen werden üblicherweise Arbeitspunkte gewählt, bei denen die Gatespannung nur wenige 100mV über der Einsatzspannung des Transistors liegt. Eine Abweichung der Einsatzspannung führt somit zu einer relativ großen Abweichung des elektrischen Verhaltens des Transistors in einem derartigen Arbeitspunkt.

Anhand der bereits beschriebenen Differenzstufe (Abbildung 2.3) lässt sich dieser Sachverhalt leicht verdeutlichen. Wie bereits erwähnt, ist das Matching-Verhalten der Differenzstufe von entscheidender Bedeutung. Eine unterschiedliche Degradation

der Eingangselemente führt zu einer belastungsbedingten Eingangs-Offsetspannung. Da diese Baugruppe die erste verstärkende Stufe eines Operationsverstärkers darstellt, setzt sich deren Fehler fort und wird weiter verstärkt. Die Auswirkungen der durch BT-Belastungen hervorgerufenen Degradationen wirken sich also innerhalb einer Analoganwendung stärker aus und der entstehende Fehler lässt sich bei der weiteren Signalverarbeitung nicht einfach eliminieren, sondern wird sogar noch verstärkt. Eine Möglichkeit, den degradationsbedingten Fehler auf schaltungstechnischer Ebene zu bekämpfen, ist der Einsatz von ‚auto-zeroing‘-Techniken, wie z.B. Chopper-Verstärkern [TieSch93]. Dies erfordert jedoch einen höheren Aufwand bei der Schaltungsentwicklung und kann zu Takteinkopplungen führen. In Kapitel 7 werden weitere schaltungstechnische Alternativen erläutert, die die Degradation nach BTS und ihre Auswirkungen vermindern können.

2.5.3 Auswirkung der Degradation nach inhomogener Belastung

Die Auswirkungen auf die Funktion einer Schaltung, die durch die Degradation von Transistorparametern nach inhomogenen BT-Belastungen hervorgerufen werden, sind qualitativ vergleichbar mit den Auswirkungen nach homogenen BT-Belastungen.

Für den in Kapitel 2.1.3 beschriebenen zeitkontinuierlichen Komparator heißt das, dass sowohl die durch homogene als auch durch inhomogene BT-Belastungen hervorgerufene Degradation auf Schaltungsebene zu einer Verschiebung der Eingangs-Offset-Spannung führen kann. Die Stärke der Degradation fällt jedoch unter den beiden verschiedenen BT-Belastungen unterschiedlich aus.

Inhomogene BT-Belastungen der p-Kanal Transistoren des in Kapitel 2.1.3 beschriebenen spannungsgesteuerten Oszillators führen zu einer Abnahme ihrer Stromtreibfähigkeit. Dies kann in dieser Schaltung zu einem Verlust an (maximaler) Oszillationsfrequenz führen, zu einer Verschiebung des Bereiches der einstellbaren Frequenz zu niedrigeren Werten oder sogar zu einem generellen Verlust der Oszillation. Darüber hinaus kann es zu einer Erhöhung des Phasenrauschens kommen [Tiebout01].

2.6 Lebensdauerextrapolation

Je nach Einsatzzweck und –ort einer mikroelektronischen Schaltung können die Anforderungen an deren Zuverlässigkeit stark schwanken. An Halbleiterprodukte, die z.B. in sicherheitsrelevanten Anwendungen eingesetzt werden, werden höhere Zuverlässigkeitsanforderungen gestellt als in der Unterhaltungselektronik.

Nur in seltenen Fällen kann die Produktlebensdauer einer Schaltung durch einen echten vollständigen Test des Betriebes einer Baugruppe für den gesamten geforderten Lebensdauerzeitraum abgesichert werden. Produkte könnten bei diesem Weg erst Jahre nach Abschluss ihrer Entwicklung zum Verkauf angeboten werden.

Es müssen also andere Möglichkeiten in Betracht gezogen werden, die Lebensdauern von einzelnen Transistoren und gesamten Schaltungen bewerten zu können, ohne die gesamte abzusichernde Lebensdauer abwarten zu müssen. In den folgenden Unterkapiteln wird auf Methoden eingegangen, die für die Lebensdauerbeurteilung herangezogen werden. Sie verkürzen die Messungen auf ein sinnvolles praktikables Niveau und ermöglichen den Schritt von der Bewertung eines einzelnen Transistors auf die Bewertung einer gesamten Schaltung.

2.6.1 'Worst-Case' und 'Duty-Cycle' Ansatz

Innerhalb einer Schaltungsapplikation durchlaufen die eingesetzten Transistoren verschiedene Arbeitspunkte. Jeden einzelnen Transistor zu bewerten und abzusichern sprengt bei einer höheren Anzahl von Transistoren schnell die zeitlichen Möglichkeiten. Es muss ein Ansatz gewählt werden, der auf der einen Seite den Arbeitsaufwand bei Messung und Bewertung stark reduziert, auf der anderen Seite aber weiterhin die Möglichkeit bietet, die Lebensdauer einer Schaltung mit vielen Einzeltransistoren mit ausreichender Genauigkeit zu bewerten.

Hierfür wird der so genannte ‚worst-case‘ Ansatz zugrunde gelegt. Dafür wird unter den verschiedenen im Betrieb möglichen Arbeitspunkten der Transistoren jeweils nur der belastungskritischste betrachtet. Die Bestimmung dieses Belastungsarbeitspunktes ist eine wichtige Voraussetzung für die korrekte Bewertung und muss für jede eingesetzte Transistorbauart (Kanalart, Gateoxiddicke, Diffusionsprofile, Wannenkonzentration usw.) erfolgen.

Die extrapolierte Lebensdauer eines einzelnen Transistors in diesem ‚worst-case‘-Arbeitspunkt entspricht dann einer sicheren Minimum-Lebensdauer aller eingesetzten Transistoren dieses Typs. Die Lebensdauer, die der Transistor in diesem kritischsten Arbeitspunkt erreicht, wird in allen anderen Arbeitspunkten ebenfalls erreicht oder übertroffen.

Dieser Ansatz kann auch für den zeitlichen Ablauf und Wechsel verschiedener Belastungskriterien angewandt werden. Es wird jeweils der kritischste Fall angenommen. Der belastungsstärkste Arbeitspunkt wird dann für die gesamte Lebensdauer als aktiv angenommen. Wenn aus der Funktionsweise der Schaltung und der Produkthanforderung heraus bestimmt werden kann, dass die Transistoren nur zu einem Anteil der geforderten Produktlebensdauer der Belastungssituation ausgesetzt sein werden, wird für die Lebensdauerextrapolation ein so genannter ‚Duty-Cycle‘ eingeführt. Werden die Transistoren z.B. nur zwei Prozent der geforderten Lebensdauer der Belastung ausgesetzt, wird bei der Lebensdauerextrapolation ein Faktor von 50 berücksichtigt.

Abbildung 2.10 veranschaulicht diesen Ansatz am Beispiel eines Inverters. Der p-Kanal Transistor wird während des beispielhaft aufgetragenen zeitlichen Verlaufs der Ein- und Ausgangsspannungen des Inverters nur zu bestimmten Zeiten einer BT-Belastung ausgesetzt.

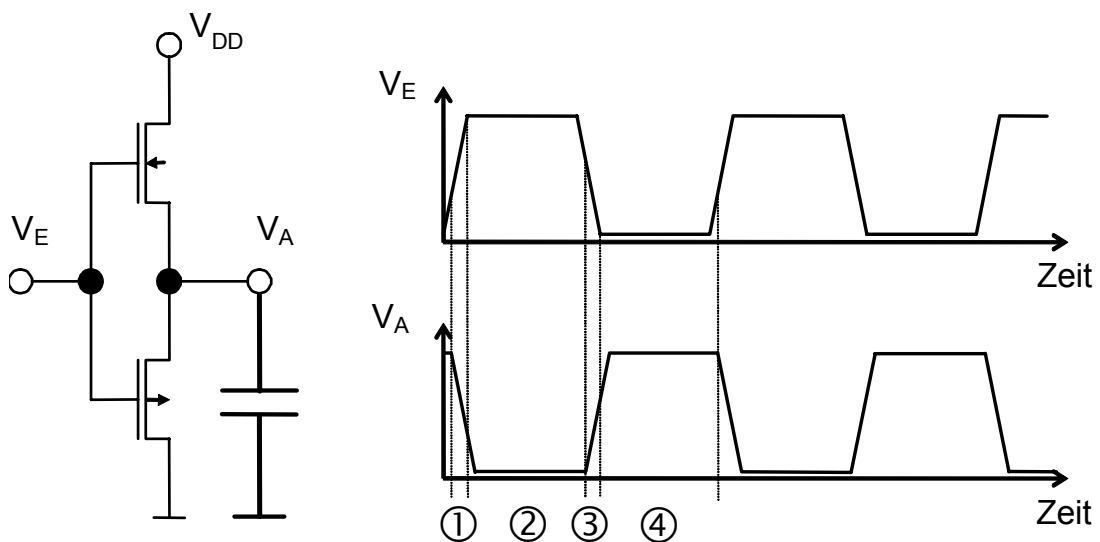


Abbildung 2.10: Inverterschaltung mit kapazitiver Last und beispielhafte zeitliche Auftragung der Ein- und Ausgangsspannung. Die vier gekennzeichneten Bereiche entsprechen unterschiedlichen Belastungssituationen.

Nur innerhalb des mit der Nummer 4 gekennzeichneten Bereiches, in dem die Eingangsspannung des Inverters auf Null bzw. Low-Pegel liegt, herrscht bei gleichzeitig erhöhter Umgebungstemperatur die BT-Belastungsbedingung für den p-Kanaltransistor vor (siehe Kapitel 2.1.1). In den übrigen Bereichen herrschen andere Belastungsbedingungen vor, die zu von BTS abweichenden Belastungen der Transistoren führen und gesondert beachtet werden müssen [Schlünder05_2].

Wird für den untersuchten Inverter angenommen, dass zu z.B. gleichen Anteilen Low- und High-Pegel am Eingang liegen (Bereiche 2 und 4), kann der Duty-Cycle für BT-Belastung des p-Kanal Transistors unter Vernachlässigung der Flanken (Bereiche 1 und 3) mit 50% angenommen werden. Bei der Lebensdauerextrapolation wird dann ein Faktor von 2 berücksichtigt.

Für alle Parameter, die bei den untersuchten Degradationsmechanismen Einfluss auf die Lebensdauer des Transistors haben, deren Wert aber für die geforderte Betriebsdauer nicht mithilfe eines Duty-Cycle Wertes genauer spezifiziert werden kann, wird jeweils der ‚worst-case‘ angenommen. Die Einflussgrößen bei BTS sind die Betriebsspannung, die Temperatur und die Belastungsdauer. Bei anderen Belastungsarten mit anderen Parameterabhängigkeiten können weitere Einflussgrößen wie z.B. die Transistorgeometrie dazukommen.

Werden die Anforderungen an die Lebensdauer für die Schaltung unter diesen Bedingungen erfüllt, wird die Schaltung auch im realen Einsatz die Lebensdauerziele erreichen. Einschränkend muss hier hinzugefügt werden, dass es sich hierbei um eine Extrapolation handelt, die grundsätzlich mit bestimmten Ungenauigkeiten behaftet ist.

Mithilfe dieser Vorgehensweise kann die Lebensdauererwartung einzelner Transistoren in einer Schaltung betrachtet und dann auf die Lebensdauer der gesamten Schaltung geschlossen werden. Die Lebensdauer einer Schaltung gilt als erreicht, sobald die elektrischen Parameter eines einzelnen Transistors aus zuvor festgesetzten Grenzen herauslaufen, da nach einer Drift der Parameter über die festgelegten Grenzen hinaus die Funktionalität der Schaltung nicht mehr gewährleistet werden kann.

Die zur Bewertung der Lebensdauer notwendigen Experimente müssen auf alle verschiedenen Transistortypen angewendet werden, die die Technologie zur Verfügung stellt, unabhängig davon, wie häufig sie in einer Schaltung eingesetzt werden. Die notwendigen Lebensdauerbewertungen können so auch bei komplexen Produkten mit sehr vielen eingesetzten Transistoren auf eine realisierbare Anzahl von vergleichsweise wenigen Einzelbewertungen reduziert werden.

2.6.2 Beschleunigung

Für die Bewertung einer geforderten Lebensdauer von z.B. 10 Jahren, die für die Transistoren abgesichert werden soll, müssen die Messzeiten auf Zeiträume im Bereich von Stunden bis maximal Wochen reduziert werden. Um in diesem Zeitbereich Belastungsuntersuchungen durchführen und bewerten zu können, werden beschleunigte Belastungen verwendet. Die Degradation der elektrischen Parameter der Transistoren wird dabei gegenüber realen Betriebsbedingungen beschleunigt. Hierfür werden die BTS-relevanten Belastungsparameter verändert. Bei beschleunigten Experimenten werden z.B. die Temperatur und die Belastungsspannung gegenüber realen Bedingungen angehoben.

Die gewählten erhöhten Belastungsbedingungen erfordern später für die Lebensdauerabschätzung eine Umrechnung der erreichten Lebensdauerwerte zurück auf die geforderten realen Betriebsbedingungen. So kann in vertretbaren Belastungszeiten die Parameterdegradation gemessen werden und die Belastungszeit bestimmt werden, die verstreicht, bis eine festgesetzte maximale Parameterverschiebung erreicht wird. Die Zeitspanne wird für jeden für die Beschleunigung der Degradation veränderten Parameter extrapoliert auf eine Zeitspanne, die den geforderten unbeschleunigten Betriebsbedingungen entspricht. In unserem Beispiel erfolgen also zwei Extrapolationen: Erstens eine Extrapolation der Lebensdauer, die während des Experimentes bei angelegter Belastungsspannung erreicht wurde, auf die Produktlebensdauer bei Betriebsspannung. Zweitens erfolgt eine Extrapolation der Temperatur und der damit assoziierten Lebensdauer auf die Bedingungen der regulär geforderten Temperatur.

Diese Extrapolationen basieren auf Modellen, die die Beschleunigungen der Degradationen als Funktion der verschiedenen Belastungsparameter beschreiben. Diese Modelle erfordern ein tiefes Verständnis der physikalischen Mechanismen, die die Schädigung des Transistors im Belastungsfall beschreiben. Die Extrapolationen der Lebensdauer werden oft über eine Zeitdekade hinaus durchgeführt. Unzureichend genaue Modelle generieren daher sehr große Fehler bei der Lebensdauerabschätzung. In Kapitel 2.6.3 werden die Extrapolationsmodelle in eigenen Unterkapiteln detailliert beschrieben.

Als sehr wichtig in diesem Zusammenhang muss auch das Verständnis der Beschleunigungsmechanismen angesehen werden. Damit die Vorgehensweise der Belastung in Zeitraffer mit anschließender Extrapolation auf Betriebsbedingungen

mit akzeptierbarem Fehler funktionieren kann, müssen mehrere Bedingungen erfüllt werden. Zum einen darf die Beschleunigung nur unter Bedingungen geschehen, unter denen der Beschleunigungsfaktor bekannt und möglichst einheitlich ist, zum anderen dürfen keine weitere Schädigungsmechanismen hinzukommen oder stark an Bedeutung gewinnen, die unter den geforderten Betriebsbedingungen keine Rolle spielen. Falls die letztgenannte Bedingung nicht erfüllt ist, beruht die gemessene Parameterdegradation auf den Beiträgen mehrerer Schädigungsmechanismen, die oft nicht eindeutig trennbar sind. Eine korrekte Bewertung der Degradation und die Extrapolation auf Betriebsbedingungen sind dann nicht mehr möglich.

2.6.2.1 Beschleunigung durch erhöhte Spannung

Bei der bislang stetigen Skalierung der Transistorparameter sind die für BT-Belastungen entscheidenden Parameter Spannung und Gateoxiddicke oft nicht in dem Maße mitskaliert worden, dass die elektrische Feldstärke über dem Oxid konstant geblieben wäre. Die Feldstärke über dem Oxid während des normalen Betriebes ist mit modernen Technologien zunehmend gestiegen [ITRS04, Thewes99_1]. Ein einheitlicher fester Wert für eine geeignete Belastungsfeldstärke für Transistoren verschiedener CMOS-Technologien kann daher nicht angegeben werden. Belastungsbedingungen und Schädigungsmechanismen müssen bei jeder neuen Technologiegeneration bewertet werden.

Der Bereich der möglichen Belastungsspannungen erstreckt sich über ein Intervall, das sinnvoll zu betragsmäßig hohen Werten durch das Auftreten zusätzlicher Schädigungsmechanismen und letztendlich durch den Spannungsdurchbruch des Gateoxides begrenzt wird. Belastungsuntersuchungen bei der maximal erlaubten Betriebsspannung des Transistors betrachten definitionsgemäß keine beschleunigte Degradation. In Spezialfällen, wenn z.B. nur relativ kurze Transistorlebensdauern abgesichert werden müssen, bieten sie jedoch den Vorteil, dass mögliche Vorhersageungenauigkeiten, bedingt durch Extrapolationen auf unterschiedliche Belastungsspannungen, wegfallen. Liegt die geforderte Lebensdauer in Bereichen bis etwa 100h, stellt die unbeschleunigte Belastung daher eine praktikable Möglichkeit dar, die Zuverlässigkeit des Bauelementes abzusichern. Beschleunigte Belastungsuntersuchungen, für die mithilfe von mindestens drei Belastungsspannungen hinreichend genau die Spannungsbeschleunigung ermittelt werden muss, bieten in diesen Fällen keinen zeitlichen Vorteil.

Sollen längere Lebensdauern für die Transistoren zugesichert werden, kommen üblicherweise beschleunigte Messungen zum Einsatz. Die Gatespannung wird während der Belastung gegenüber den bei normalem Betrieb möglichen Spannungen erhöht. Die Schädigung wird durch diese Maßnahme beschleunigt. Das Degradationsverhalten des Bauelementes kann mit geringerem Zeitaufwand beobachtet und bewertet werden.

2.6.2.2 Beschleunigung durch erhöhte Temperatur

Neben der Spannung, die für die Belastungsuntersuchungen erhöht wird und sich beschleunigend auf die Degradation auswirkt, steht ein weiterer für die Belastung variierbarer Parameter zur Verfügung. Die Temperatur des Transistors während der Belastungsuntersuchungen wirkt sich ebenfalls auf Intensität und Geschwindigkeit der Schädigung aus. Je nach zugrunde liegendem Schädigungsmechanismus kann eine Erhöhung oder eine Absenkung der Temperatur sich beschleunigend auf den Schädigungsmechanismus auswirken. Im Falle von BT-Belastungen wirkt sich eine Erhöhung der Temperatur beschleunigend aus.

Auch für die Beschleunigung über die Temperatur müssen die bereits bei der Spannungsbeschleunigung beschriebenen Regeln eingehalten werden. Wird die Temperatur zu stark erhöht, können zusätzliche schädigende Effekte auftreten.

Lässt sich ein derartiger zusätzlicher Schädigungsmechanismus bei den ermittelten Messwerten nicht eindeutig extrahieren, wird auch hier die Voraussetzung für eine korrekte Extrapolation auf die Betriebstemperatur genommen. Die berechnete Lebensdauer kann die echte Lebensdauer sowohl über- als unterschätzen. Tritt ein zusätzlicher, einen elektrischen Parameter des Transistors in dieselbe Richtung wie bei BTS verschiebender Schädigungsmechanismus auf, ergibt die Extrapolation auf Betriebsbedingungen eine zu niedrige Lebensdauer. Tritt dagegen ein zusätzlicher Mechanismus in Kraft, der eine Degradation der betrachteten Parameter in eine entgegen gesetzte Richtung bewirkt, überlagern sich die Schädigungseffekte und ergeben letztendlich eine schwächere Degradation der elektrischen Parameter. Eine folgende Extrapolation auf Betriebsbedingungen ergäbe eine überschätzte Lebensdauer.

2.6.3 Extrapolation auf Betriebsbedingungen

Für die Berechnung der Zeiten bis zum Erreichen einer maximal erlaubten Parameterverschiebung, die ein Transistor bei normalen, unbeschleunigten Belastungsbedingungen erreicht, werden die bei beschleunigten Belastungsbedingungen gemessenen Zeiten verwendet und extrapoliert. Dafür müssen Modelle herangezogen werden, die die unterschiedlich beschleunigende Wirkung der Veränderung der Belastungsparameter berücksichtigen. Je nach Schädigungsmechanismus müssen angemessene Modelle verwendet werden.

2.6.3.1 Extrapolation der Belastungsspannung

Das Verfahren der Extrapolation auf unterschiedliche Betriebsspannungen lässt sich sehr einfach anhand eines Diagramms beschreiben. Die Werte der Belastungsspannungen werden auf der x-Achse, die Werte der erreichten Lebensdauern auf der y-Achse aufgetragen. Das in Abbildung 2.11 schematisch skizzierte Diagramm zeigt die verschiedenen Extrapolationsschritte von den Messwerten der Belastungsexperimente bis zu der Lebensdauervorhersage eines Transistors. Die x-Werte der eingetragenen Punkte werden durch die bei den Belastungsmessungen angelegten Spannungen gebildet und direkt aufgetragen, die y-Werte der Punkte entsprechen den erreichten Zeiten, bis sich die maximal erlaubte Parameterdrift einstellt. Diese Zeitwerte stellen in einigen Fällen bereits eine Extrapolation dar. Das gewählte Lebensdauerkriterium wird in solchen Fällen in der Stresszeit des Experimentes nicht vollständig erreicht. Die Zeitwerte werden dann durch Extrapolation der beobachteten Parameterverschiebungen bis zu dem gewählten Lebensdauerkriterium von z.B. 100mV Einsatzspannungsverschiebung ermittelt, wie in Abbildung 2.11 oben links schematisch dargestellt.

Jede Belastungsmessung liefert so ein Wertepaar für einen Punkt in dem Lebensdauerdiagramm (unteres Diagramm in Abbildung 2.11). Die eingetragenen Punkte beschreiben dementsprechend die Lebensdauer eines Transistors als Funktion der Betriebsspannung. Mittels einer Regression kann durch die Punkte eine Ausgleichsgerade gelegt werden, die letztendlich die Funktion der Spannungsbeschleunigung beschreibt. Die Lebensdauer kann mithilfe dieser zweiten Extrapolation auf die gewünschte Betriebsspannung extrapoliert werden.

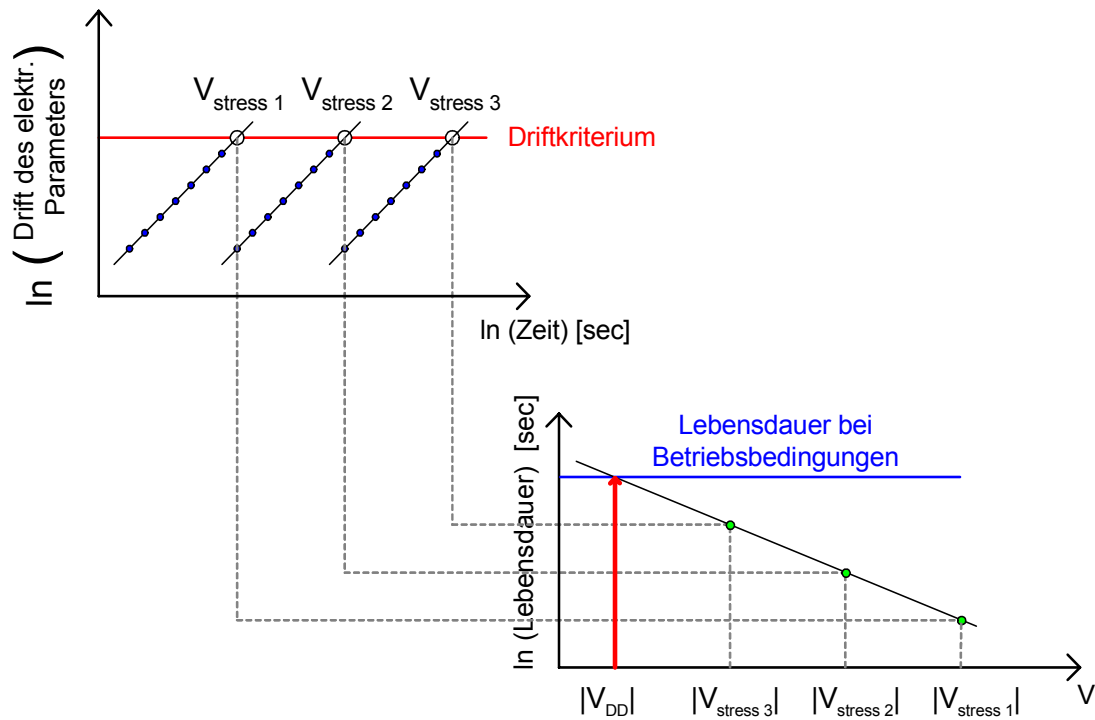


Abbildung 2.11: Schematische Darstellung zur Extrapolation der Lebensdauer bei beschleunigten Untersuchungen auf die Lebensdauer bei der gewünschten Betriebsspannung. Die Belastungsspannungen $V_{\text{stress } 1}$, $V_{\text{stress } 2}$, $V_{\text{stress } 3}$ entsprechen in diesem Beispiel den negativen Belastungs-Gatespannungen.

Die Punkte der verschiedenen Belastungsmessungen zeigen oft Schwankungen und damit Abweichungen von der ermittelten Extrapolationsfunktion. Die Abweichungen können durch Fehler bei den weiter oben beschriebenen Hochrechnungen der Parameterdrift entstehen oder können auf realen statistischen Schwankungen basieren. Um dadurch bedingte Fehler bei der Lebensdauer vorhersage zu minimieren, sollten mindestens drei unterschiedliche Belastungsspannungen gewählt werden und pro Belastungsspannung mehrere Transistoren untersucht werden. Die ermittelten Punkte sollten ferner über mehrere Dekaden in der Zeit verteilt sein.

2.6.3.2 Extrapolation der Belastungstemperatur

Der Einfluss der Temperatur auf die Degradation wird im Falle von BTS an p-Kanal Transistoren durch das Arrhenius Modell beschrieben:

$$\tau \propto e^{-\frac{E_a}{k \cdot T}} \quad (2.1)$$

Hierbei steht k für die Boltzmannkonstante und T für die Temperatur. Der Einfluss der Temperatur auf die Degradation und damit auf die Lebensdauer τ wird durch die Aktivierungsenergie E_a beschrieben. Eine höhere Temperatur während der Belastung führt zu einer stärkeren Schädigung und somit auch zu größeren Verschiebungen der elektrischen Parameter. Hohe Werte für E_a führen zu einer starken Abhängigkeit der Degradation von der Temperatur.

Durch Anwendung des Arrhenius-Modells können die bei einer Temperatur ermittelten Zeitwerte bis zum Erreichen einer bestimmten Degradation umgerechnet werden auf Zeiten, die bei anderen Stresstemperaturen erreicht werden. Auf diese Weise kann für jedes Wertepaar aus Belastungsspannung und Stresszeit ein neues Wertepaar aus Belastungsspannung und für eine andere Temperatur neu berechneter Zeit ermittelt werden. In dem die Funktion der Spannungsbeschleunigung beschreibenden Diagramm (Abbildung 2.11 unten) ergibt sich durch die neuen als Punkte eingetragenen Wertepaare eine Parallelverschiebung. Für den Fall, dass die Lebensdauer für eine gegenüber den durchgeführten BT-Belastungen höheren Belastungstemperatur berechnet werden soll, liegen die neu berechneten Punkte unterhalb der bisherigen und ergeben so einen niedrigeren Wert für die Lebensdauer. Bei tieferen Temperaturen liegen die neuen Wertepaare oberhalb der bisherigen Punkte und die Lebensdauer fällt entsprechend höher aus.

In Abbildung 2.12 ist ein Beispiel gezeigt. Die Wertepaare für die Extrapolation auf Betriebsbedingungen mögen bei 125°C experimentell ermittelt worden sein. Soll nun die Lebensdauer jedoch für eine Temperatur von 150°C bestimmt werden, z.B. für ein Produkt aus dem Automobilbereich, wird durch Umrechnung der einzelnen bei den verschiedenen Belastungsspannungen erreichten Lebensdauern die neue Lebensdauer für die Betriebsspannung V_{DD} ermittelt.

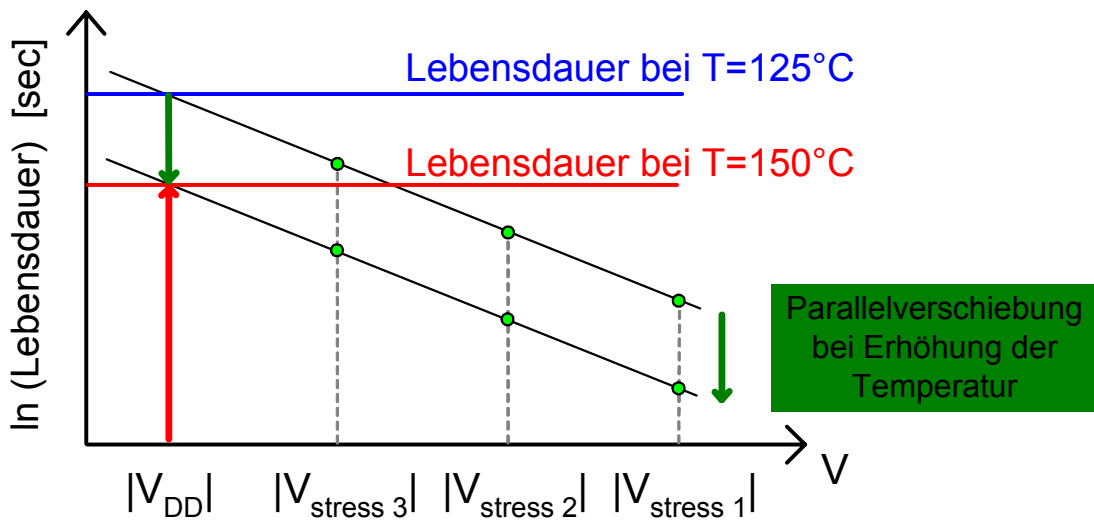


Abbildung 2.12: Schematische Darstellung des Zusammenhangs zwischen Belastungsspannung, Temperatur und Lebensdauer. Die Parallelverschiebung der Geraden berechnet sich aus der Aktivierungsenergie E_a .

Die neue Regressionsgerade durch die neu errechneten Punkte, die die Lebensdauer als Funktion der Betriebsspannung bei 150°C beschreibt, entspricht einer Parallelverschiebung der bisherigen Geraden für 125°C.

Ist die Aktivierungsenergie nicht bekannt, kann andersherum so auch der Wert für E_a bestimmt werden. Die Lebensdauer wird für mindestens zwei Temperaturen bei ansonsten identischen Belastungsbedingungen experimentell bestimmt. Mithilfe von Gleichung (2.1) kann dann E_a berechnet werden. Nach einmaliger Bestimmung der Aktivierungsenergie kann die Lebensdauer für weitere geforderte Temperaturen rechnerisch ermittelt werden.

3 Charakterisierung und Belastung

Für eine genaue Untersuchung der Auswirkungen einer BTS-Belastung eines Transistors auf seine elektrischen Parameter werden einzelne Bauelemente in speziellen Messumgebungen verschiedenen Stressbedingungen ausgesetzt. Der hierfür notwendige Laboraufbau sowie wie die benötigten Teststrukturen werden in Anhang A2 beschrieben.

In diesem Kapitel wird der strukturelle Aufbau und Ablauf der verschiedenen experimentellen Untersuchungen beschrieben. Darüber hinaus werden wichtige Kennparameter eines Transistors definiert und ihre Bedeutung diskutiert. Im Anschluss daran wird die Art und Weise der messtechnischen Bestimmung dieser Werte erläutert.

Die Degradation eines Transistors innerhalb einer Schaltung ist ein stetiger Prozess, der jedoch abhängig von der Stärke der Belastung erst nach größeren Zeiträumen eindeutig anhand von Verschiebungen seiner elektrischen Parameter (Drift) messbar wird. Um die Veränderungen eines MOS-Transistors durch BT-Belastungen in für Laborumgebungen sinnvollen Zeiträumen untersuchen zu können, werden die Degradationsprozesse bei allen hier vorgestellten Experimenten beschleunigt (siehe Kapitel 2.6.2). Die verwendeten Spannungen werden für alle Experimente jeweils angegeben, für die Belastungstemperatur wird, soweit nicht ausdrücklich anders erwähnt, $T_{stress} = 125^{\circ}\text{C}$ eingestellt.

3.1 Aufbau und Ablauf der Belastungsexperimente

In Abbildung 3.1 sind die einzelnen Schritte eines Belastungsexperimentes einschließlich Charakterisierung innerhalb eines Flussdiagramms dargestellt. In der ersten Phase werden alle zu untersuchenden Parameter des Transistors vor der eigentlichen Belastung erfasst. Die Werte, die später nach der Belastung aufgenommen werden, können dann mit diesen Referenzwerten dieser so genannten Nullstundencharakterisierung verglichen werden.

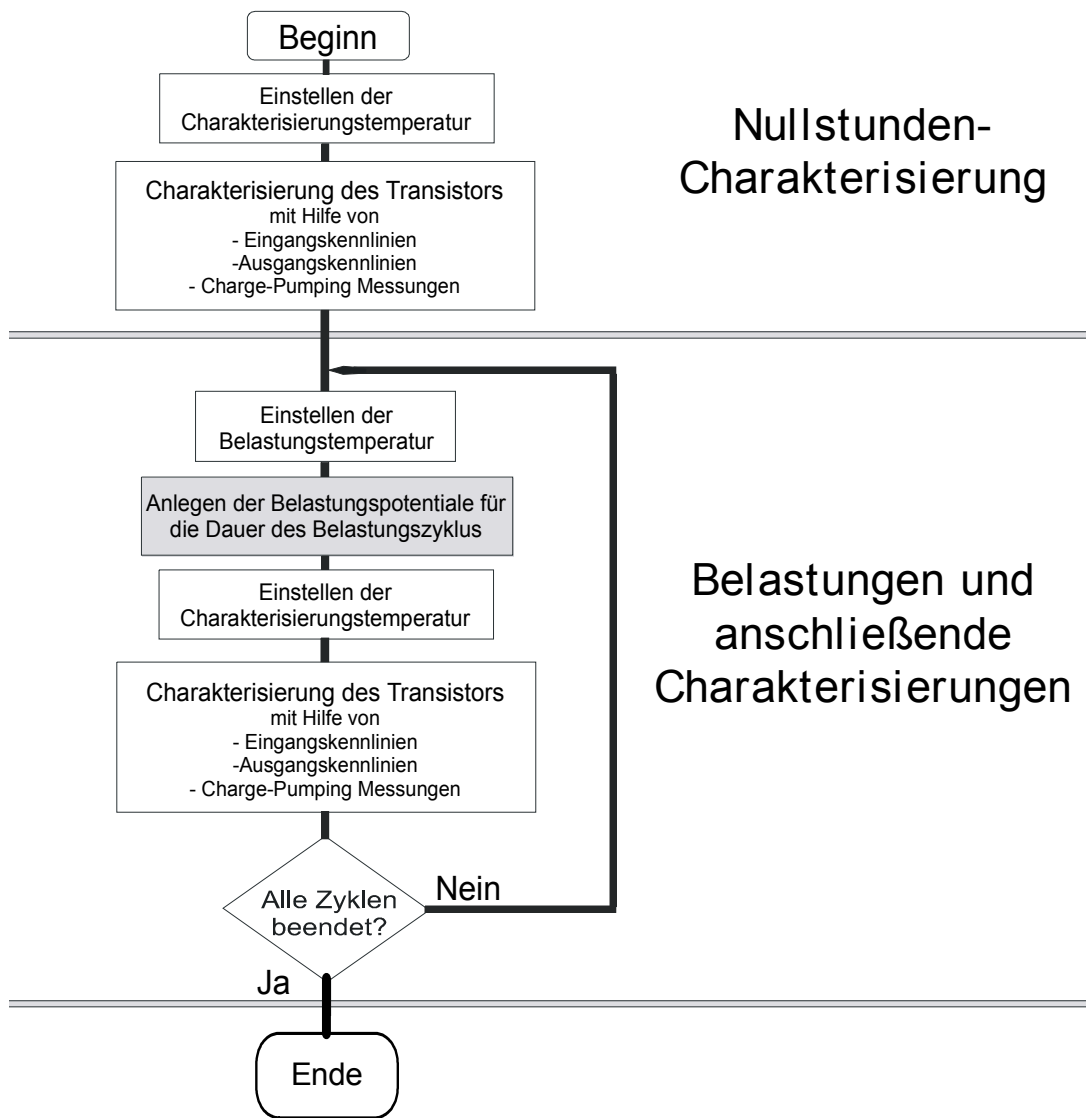


Abbildung 3.1: Flussdiagramm zur Darstellung des Ablaufs von Belastungsexperimenten

In den meisten Fällen werden hier die Werte eines bislang unbelasteten Transistors erfasst. Es können jedoch für bestimmte Untersuchungen auch bereits vorbelastete Bauelemente von Interesse sein. In der zweiten Phase wird das Bauelement Belastungen ausgesetzt. Die Parameter Zeit, Oxidfeldstärke und Temperatur bestimmen die in dieser Belastungsphase erfolgende Degradation. Nach vorbestimmten Zeitintervallen wird die Belastung unterbrochen, die Temperatur eventuell neu eingestellt und das Bauelement erneut charakterisiert. So werden die in der Belastungsphase hervorgerufenen Veränderungen durch Neuaufnahme der

Kennparameter des Transistors aufgezeichnet. Nach Abschluss der Messungen werden die Belastungsbedingungen wieder hergestellt und der Stress fortgesetzt. Auf die einzelnen während der Charakterisierung vermessenen Parameter wird genauer in den folgenden Unterkapiteln eingegangen.

Durch die sich fortlaufend wiederholende Zyklusfolge aus Belastung und Charakterisierung können nicht nur die Abweichungen der Kenngrößen des Transistors vom Ursprungswert nach einer bestimmten Zeit beurteilt werden, sondern auch der Verlauf der Änderungen der elektrischen Eigenschaften als Funktion der Zeit betrachtet werden. Da vielen der beobachteten Degradationsmechanismen Potenzgesetze oder logarithmische Zeitabhängigkeiten zugrunde liegen [Jeppson77, Ogawa95, Ogawa96, Hu85, Brox94, Thewes95, Schlünder99], wird die Unterteilung der Belastungsphasen meistens exponentiell vorgenommen. Auf diesem Weg ergeben sich bei logarithmischen Darstellungen der Zeitachsen äquidistante Schritte bei den Messwertfolgen.

3.2 Definition wichtiger elektrischer Transistorparameter

Um Transistoren eindeutig in ihren elektrischen Eigenschaften beschreiben zu können, bedarf es zunächst einer Definition der verschiedenen Kenngrößen. Darüber hinaus soll auf die Bedeutung der verschiedenen Parameter eingegangen werden.

3.2.1 Gateoxiddicke

Aus dem Quotienten der Potenzialdifferenz ober- und unterhalb des Dielektrikums des Transistors und der Dicke dieses Dielektrikums ergibt sich die elektrische Feldstärke über dem Gateoxid, die einen wesentlichen Parameter für die Belastung des Transistors im Betrieb und damit für die Alterung des Bauelementes darstellt.

Je höher die Feldstärke über dem Dielektrikum ist, desto stärker und schneller treten Schädigungen auf, die die elektrischen Parameter des Bauelementes verändern. Die Potenzialdifferenz lässt sich bei einem Belastungsexperiment einfach bestimmen. Da die Kontakte Source, Drain und Substrat während einer experimentellen BT-Belastung auf Masse gelegt werden, entspricht die an der Gateelektrode angelegte Spannung direkt der Potenzialdifferenz.

Die für die Berechnung der elektrischen Feldstärke benötigte Dicke des Dielektrikums ist durch die gewählte Technologie festgelegt. Es muss bei der Gateoxiddicke zwischen zwei verschiedenen Werteangaben mit jeweils unterschiedlichen Bedeutungen unterschieden werden. Die Schicht des Gateoxides wird durch die physikalische Dicke und durch die elektrisch wirksame (oder auch effektive Gateoxiddicke genannt) Dicke beschrieben.

3.2.1.1 Physikalische Gateoxiddicke

Der Wert der physikalischen Gateoxiddicke entspricht der tatsächlich geometrisch vorhandenen Schichtdicke des dielektrischen Materials, das während der Prozessierung aufgewachsen wird. Diese Schichtdicke kann dementsprechend geometrisch gemessen werden.

Ein sehr direkter und in seinem Prinzip sehr einfacher Weg besteht darin, das Bauelement zu zerschneiden und die Schichtdicke dort direkt zu vermessen. Um die Dickenbestimmung unter einem Mikroskop zu vereinfachen und auch um die erreichbare Genauigkeit zu erhöhen, wird das Bauteil unter einem Winkel angeschliffen. Die zu vermessende Kantenlänge wird durch den Schliffwinkel vergrößert, mittels Mikroskop vermessen und dann auf die reale Schichtdicke mittels Trigonometrie zurückgerechnet. In Abbildung 3.2 wird das Verfahren anhand einer Skizze verdeutlicht.

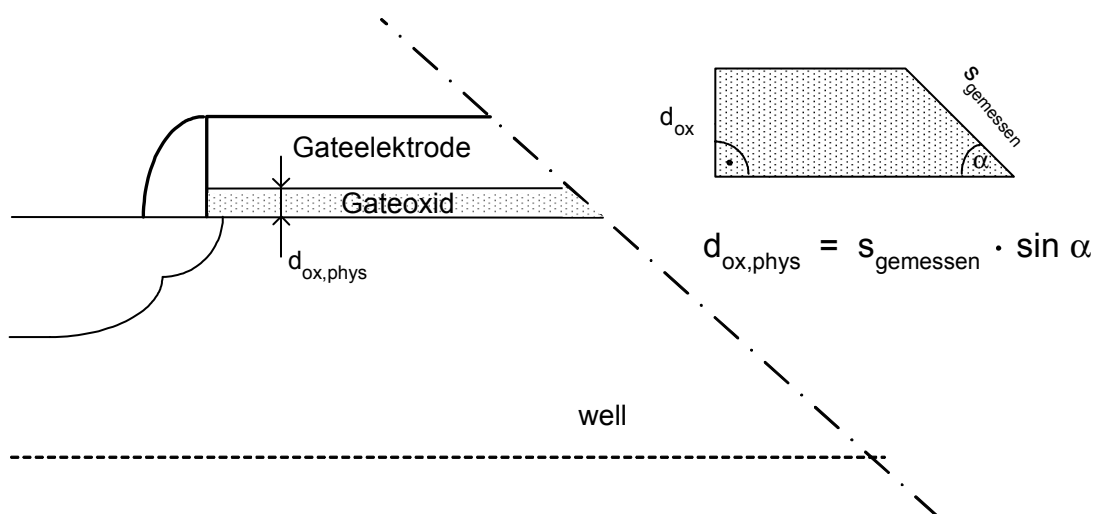


Abbildung 3.2: Bestimmung der physikalischen Gateoxiddicke $d_{ox,phys}$ unter dem Mikroskop mittels Schrägschliff-Methode

Eine weitere Möglichkeit, die Oxiddicke zu messen, besteht darin, sie mittels interferometrischer Verfahren zu bestimmen. Hierbei werden die Phänomene der konstruktiven bzw. destruktiven Interferenz von Licht ausgenutzt.

Auf das Bauelement einfallendes kohärentes Licht wird an der oberen und an der unteren Grenzfläche des Oxides reflektiert. Das an den zwei Flächen reflektierte Licht überlagert sich (siehe Abbildung 3.3). Über die entstehenden Interferenzmuster (Interferogramm) kann auf den Abstand der reflektierenden Grenzflächen zurückgeschlossen werden. Der Abstand der Interferenzwellentäler bzw. -wellenberge kann in den optischen Weg zwischen den Grenzflächen umgerechnet werden. Mithilfe des Brechungsindex des Dielektrikums und der Wellenlänge des eingesetzten Lichtstrahls kann auf die physikalische Dicke des vermessenen Materials in diesem Fall des Gateoxides geschlossen werden.

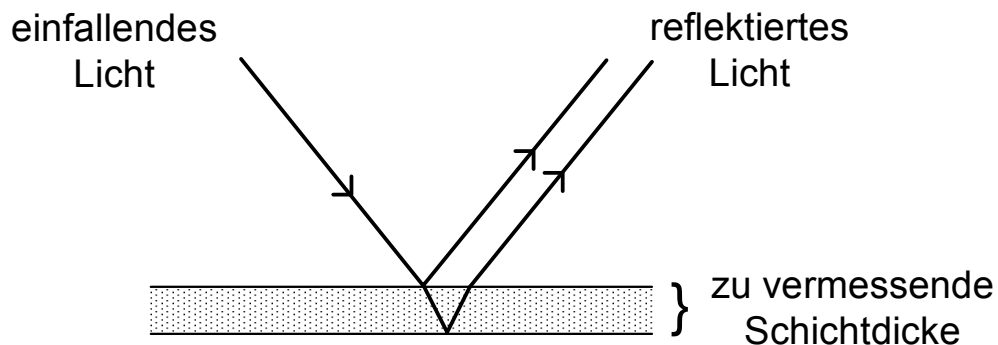


Abbildung 3.3: Zur Bestimmung der Gateoxiddicke aus der Interferenz einfallenden und reflektierten Lichts. Über das entstehende Interferogramm kann auf die Schichtdicke geschlossen werden.

Um auch bei neuesten Prozesstechnologien mit sehr dünnen Gateoxidschichten mit ausreichender Genauigkeit die Schichtdicke bestimmen zu können, kommt häufig das Ellipsometrie-Verfahren zur Anwendung [McCrackin63]. Bei dieser Messmethode wird die Änderung des Polarisationszustandes gemessen, die eintritt, wenn ein monochromatischer Strahl polarisierten Lichtes an einem Grenzflächenübergang reflektiert wird oder durch diesen durchtritt. Der Vorteil dieser Methode ist die sehr hohe Auflösung und die Tatsache, dass mehrere experimentelle Messgrößen bei einer gegebenen Wellenlänge λ bestimmt werden können.

In Abbildung 3.4 ist der Messaufbau skizziert. Mithilfe der nach der Reflexion des Lichtstrahls eintretenden Drehung der Polarisation kann auf die Schichtdicke zurückgeschlossen werden. Durch Variation des Winkels kann darüber hinaus auch die Dielektrizitätskonstante bestimmt werden [McCrackin63].

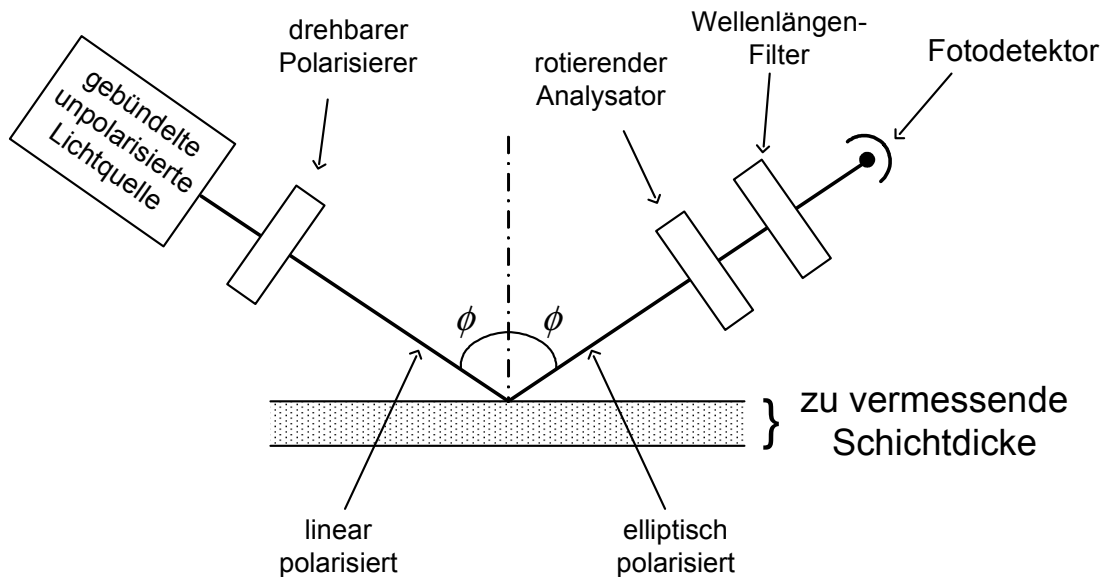


Abbildung 3.4: Verfahren zur Filmdickenbestimmung mittels Ellipsometrie. Die Winkeländerung der Polarisation des reflektierten Lichtstrahls ist proportional zur Dicke der zu vermessenen Schicht.

3.2.1.2 Elektrisch wirksame Gateoxiddicke

Die Gateoxiddicke eines Transistors kann auch auf elektrischem Weg bestimmt werden. Über die Bestimmung der MOS-Kapazität kann auf die Gateoxiddicke geschlossen werden, da sie das Dielektrikum des als Plattenkondensator interpretierten Transistors darstellt. Ist die Dielektrizitätskonstante des Materials bekannt, kann gemäß

$$d_{ox} = \frac{A}{C} \cdot \epsilon_0 \cdot \epsilon_r \quad (3.1)$$

zusammen mit der elektrisch bestimmten Kapazität des Bauelementes die Dicke d_{ox} des Dielektrikums berechnet werden. A stellt die Fläche des Kondensators dar, die durch die Gatefläche des Transistors gebildet wird. C gibt die gemessene Kapazität wieder, ϵ_0 stellt die Dielektrizitätskonstante dar und ϵ_r ist die Dielektrizitätszahl des Materials.

Vergleicht man den Wert für die Gateoxiddicke eines Transistor, den man auf elektrischem Weg bestimmt hat, mit dem Wert für die Dicke, den man über ein geometrisches Verfahren ermittelt hat, ergeben sich Abweichungen. Die elektrische Bestimmung der Oxiddicke ergibt einen höheren Wert. Die Unterschiede beruhen auf physikalischen Effekten, die vor allem bei Technologien mit dünnen Gateoxiddicken und polykristalliner Gateelektrode zum Tragen kommen. Die bei der elektrischen Bestimmung ermittelte zusätzliche Dicke lässt sich in zwei Delta-Dicken aufteilen. Es entsteht ein zusätzlicher Beitrag jeweils ober- und unterhalb des physikalisch vorhandenen Dielektrikums (siehe Abbildung 3.5).

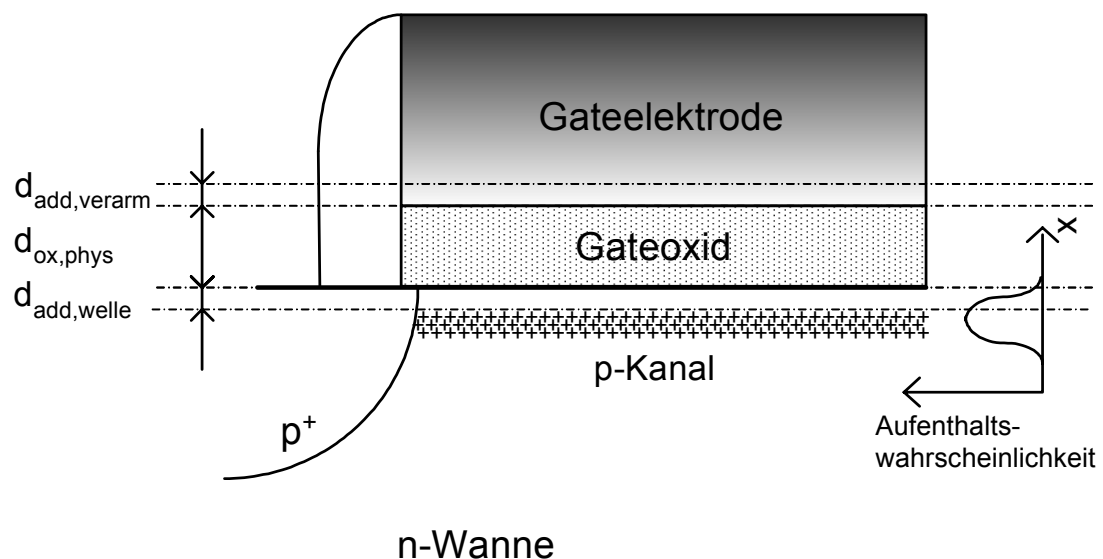


Abbildung 3.5: Schematische Darstellung zur Veranschaulichung der elektrisch wirksamen Gateoxiddicke. Zu der physikalischen Dicke ($d_{ox,phys}$) muss ein Wert oberhalb der Oxidschicht für die Verarmung in der polykristallinen Gate-Elektrode ($d_{add,verarm}$) und ein zweiter Wert unterhalb des Oxids für den Abstand der den Kanal bildenden Ladungsträger aufgrund ihrer Wellennatur ($d_{add,welle}$) addiert werden.

Der zusätzliche Beitrag unterhalb des Oxides auf der Substratseite des Bauelementes ergibt sich durch die Wellennatur der den Kanal bildenden Ladungsträger. Da sich die Ladungsträger in einem näherungsweise dreieckförmigen Potenzial befinden, nimmt die Aufenthaltswahrscheinlichkeit eines Ladungsträgers in unmittelbarer Nähe der Grenzfläche des Oxides sehr stark ab. Innerhalb der Si-Oxidschicht nimmt die Aufenthaltswahrscheinlichkeit exponentiell ab. Die Inversionsschicht des Transistors im eingeschalteten Zustand verläuft daher nicht direkt an der Oxidgrenzschicht sondern in einem bestimmten Abstand davon. Zur physikalisch vorhandenen Oxiddicke muss hier ein Wert von ca. 0,4nm addiert werden [ITRS04, Pacha04_1].

Wird für die Bildung der Gateelektrode dotiertes polykristallines Silizium eingesetzt, wie es momentan bei fast allen produktiven Halbleitertechnologien üblich ist, tritt ein Effekt auf, der in der Literatur mit ‚Polydepletion‘ beschrieben wird [Martin98, Hokari88, Shuegraf93]. Die Verteilung der Ladungsträger in der dotierten polykristallinen Gateelektrode ist nicht homogen. Im Inversionsbetrieb des Transistors verarmt die Konzentration der Ladungsträger nahe der Grenzschicht zu dem darunter liegenden Oxid. Diese Verarmungsschicht oberhalb des Oxides muss also ebenfalls bedacht werden. Sie bewirkt elektrisch gesehen ebenfalls eine Verdickung der Oxiddicke um weitere 0,3-0,5nm [ITRS04, Pacha04_1].

Prozesstechnisch kann die elektrische Dicke des Gateoxides durch die Nutzung alternativer Gateoxidmaterialien anstelle eines reinen SiO₂ und durch die Verwendung eines metallischen Materials für die Gateelektrode vermindert werden.

Für die elektrischen Eigenschaften des Transistors und auch für das Alterungsverhalten ist diese elektrisch wirksame Gateoxiddicke entscheidend, die deshalb auch oft ‚Effektive Gateoxiddicke‘ genannt wird. Das Alterungsverhalten wird durch die effektive Gateoxiddicke mitbestimmt, weil über diesen Wert im Gegensatz zur physikalischen Gateoxiddicke die tatsächlich vorhandene Feldstärke, die innerhalb des Transistors anliegt, festgelegt ist. Für eine ausreichende Beschreibung eines Transistors benötigt man daher neben dem Wert für die physikalische Gateoxiddicke auch die elektrisch wirksame Dicke.

3.2.2 Einsatzspannung

Die Einsatzspannung stellt einen elementaren Kennwert eines Transistors dar. Sie ist ein wichtiger Parameter bei der Charakterisierung des Bauelementes. Die Definition dieses Parameters erfolgt in der Praxis jedoch nicht einheitlich.

In der Theorie der Device-Physik wird die Schwellenspannung des Transistors über folgende Gleichung definiert [Sze81]:

$$V_{th} = -2\psi_B - \frac{1}{C_{OX}} [Q_{SC}(-2\psi_B) + Q_{it}(-2\psi_B) + Q_{OX}] + \frac{\phi_{MS}}{q} \quad (3.2)$$

Die Einsatzspannung hängt demnach vom Halbleitermaterial und seiner Dotierung, vom Material der Gateelektrode, vom Material des Isolators und von seiner Dicke ab. C_{OX} steht hier für die Oxidkapazität, Q_{sc} für die Ladung im Halbleiter, ϕ_{MS} für die Differenz der Austrittsarbeiten der unterschiedlichen Materialien und q für die Elementarladung. Weiterhin geht die Qualität des Isolators und der Grenzfläche über die Parameter Q_{OX} und Q_{it} ein. Q_{OX} steht für die festen Ladungen innerhalb des Isolators und Q_{it} für die Ladung in Grenzflächenzuständen. Die Einsatzspannung ist hier die an das Gate angelegte Spannung, bei der per Definition der Übergang zwischen schwacher und starker Inversion des Kanalbereiches stattfindet.

In Abbildung 3.6 werden die Zusammenhänge mithilfe des Bänderdiagramms eines p-dotierten Halbleiters weiter beschrieben. ψ_B steht für die Potenzialdifferenz zwischen dem Fermi-niveau der Kanalladungsträger und dem intrinsischen Fermi-niveau. Entspricht die bei Anlegen einer Gatespannung eintretende Bandverbiegung ψ_S des Subtrats an der Oberfläche dem Wert ψ_B , trifft das Fermi-niveau der Kanalladungsträger und das intrinsische Fermi-niveau aufeinander und es tritt schwache Inversion ein. Werden die Bänder weiter verbogen, geht beim Erreichen des Wertes $2\psi_B$ der Transistor in starke Inversion über. An diesem Punkt gilt die Einsatzspannung des Transistors als erreicht.

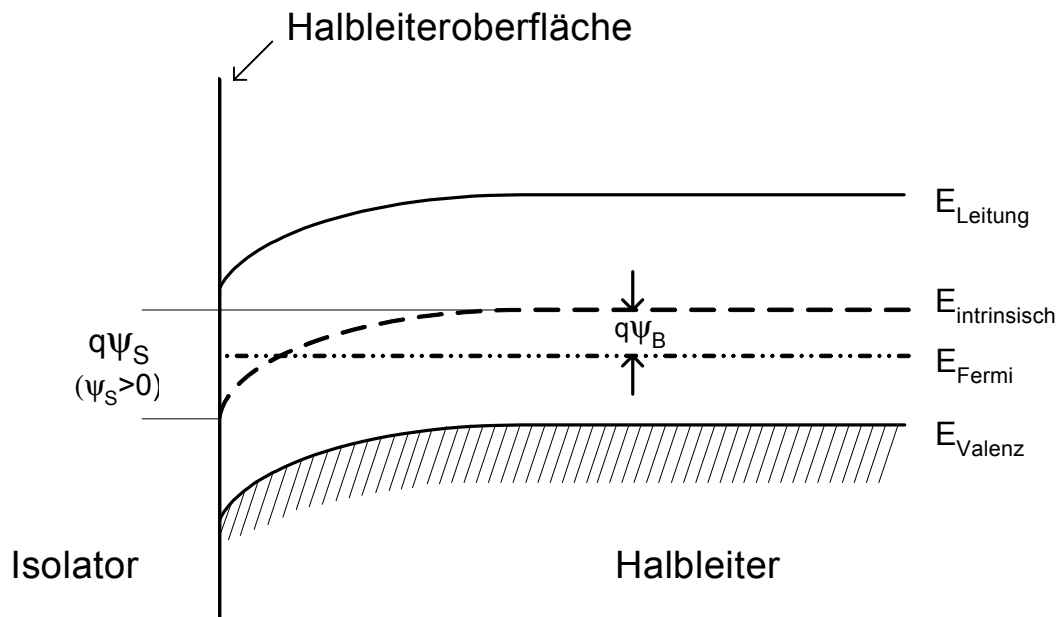


Abbildung 3.6: Bänderdiagramm der Oberfläche eines p-dotierten Halbleiters. Das Potenzial wird als gleich Null in Substrat definiert und wird gemessen gegenüber dem intrinsischen Fermienergie $E_{intrinsisch}$. Das Oberflächenpotenzial ist wie gezeigt positiv. Akkumulation gilt für $\psi_S < 0$, Verarmung für $\psi_B > \psi_S > 0$, schwache Inversion für $\psi_S > \psi_B$ und starke Inversion für $\psi_S > 2\psi_B$.

Die Bestimmung der Einsatzspannung in der Praxis erfolgt oft durch praktikable Methoden für die messtechnische Bestimmung dieses Parameters, wie im Kapitel 3.3.2 noch erläutert wird.

3.2.3 Weitere Transistorparameter

Neben dem Parameter der Einsatzspannung werden für eine genaue Charakterisierung des Bauelementes weitere Kennwerte benötigt. Je nach den an den Transistorknoten anliegenden Potenzialen leitet ein MOS-Transistor Ströme zwischen seinen Source- und Drainanschlüssen unterschiedlich gut. Die Betriebsmoden des Bauelementes durchlaufen dabei unterschiedliche Bereiche. Grob können die Arbeitsbereiche in Unterschwellenbereich, linearen Bereich und Sättigungsbereich eingeteilt werden. Je nach Applikation, in der die Bauelemente betrieben werden, erlangen die Eigenschaften des Devices in den verschiedenen Bereichen unterschiedlich hohe Bedeutung.

In Digital-Anwendungen kommt vor allem den möglichen Strömen im Sättigungsbereich und der Einsatzspannung hohe Bedeutung zu. Wie schnell ein Transistor eine am Drain angeschlossene Kapazität umladen kann, gehört hier zu den wichtigsten Eigenschaften. Hierdurch wird die mögliche Geschwindigkeit einer Schaltung bestimmt. Auch das Sperrverhalten der Transistoren erlangt eine wachsende Bedeutung, da es bei komplexen Schaltungen immer mehr den Leistungsverbrauch des Chips mitbestimmt [Pacha04_2, Armin04].

In Analog- und Mixed-Signal-Anwendungen kommt darüber hinaus den so genannten Kleinsignalparametern große Bedeutung zu. Sie bestimmen das Übertragungsverhalten von Schaltungen. Die beiden wichtigsten Kleinsignalparameter sind die (Vorwärts-) Steilheit oder auch Transkonduktanz

$$g_m = \frac{\partial I}{\partial V_G} \quad (3.3)$$

und der differenzielle Ausgangsleitwert, auch differenzieller Drain-Source-Leitwert genannt.

$$g_{DS} = \frac{\partial I}{\partial V_D} \quad (3.4)$$

3.3 Transistorcharakterisierung

Für die Charakterisierung der Transistoren werden in jedem Messzyklus unterschiedliche Parameter und Kennlinien aufgenommen. Sie lassen sich unterteilen in Eingangskennlinien, Einsatzspannungen, Ausgangskennlinien und in eine Charge-Pumping Messung. Auf die einzelnen Parameter dieser Messungen und die Art und Weise der messtechnischen Erfassung soll nun näher eingegangen werden.

3.3.1 Eingangs- und Ausgangskennlinien

Bei der Aufnahme der Eingangskennlinie wird der Drainstrom als Funktion der Gatespannung bei fester Drainspannung gemessen, wobei der Betrag der Gatespannung zwischen 0V und der für die untersuchte Technologie spezifizierten Betriebsspannung variiert wird. Bei Messungen der sog. Linearen Kennlinie wird an den Drainknoten des Transistors eine niedrige Drainspannung angelegt. Bei einer Technologie mit z.B. 1,8V spezifizierter Betriebsspannung werden typischerweise 50mV gewählt. Für betragsmäßig kleine Gatespannungen liegt der Transistorarbeitspunkt im Unterschwellenbereich, für größere Gatespannungen liegt der Arbeitspunkt im Linearen Bereich, auch ‚Triodengebiet‘ genannt.

In einer weiteren Eingangskennlinie mit betragsmäßig höherer Drainspannung werden Eigenschaften des Transistors im Sättigungsbereich bestimmt. Hiermit kann das elektrische Verhalten des Transistors in typischen Arbeitspunkten bei Analoganwendungen beurteilt werden. Während die Drainspannung in Abhängigkeit von der gewählten Technologie etwa auf dem Wert der halben regulären Betriebsspannung ($V_{DD} / 2$) liegt, wird der Drainstrom I_D als Funktion der Gatespannung aufgenommen.

Zusätzlich zu den zuvor beschriebenen Eingangskennlinien werden Ausgangskennlinien aufgenommen. Während dort die Gatespannung konstant gehalten wird, wird das Drainpotenzial variiert. Der fließende Drainstrom wird als Funktion der Drainspannung aufgetragen. Diese Kennlinie wird mit unterschiedlichen Gatespannungen aufgenommen.

Die Einsatzspannungen mehrerer von den Geometriewerten identischer Transistoren eines Wafers variieren leicht. Diese Variation der Schwellenspannung ist auf

technologiebedingte Streuung der Dotierstoffkonzentrationen innerhalb der Transistorgebiete, Abweichungen der Dicke des Gateoxids und Schwankungen weiterer Parameter zurückzuführen [Laksh86, Pelgrom89, Oehm93, Stolk98]. Um die Vergleichbarkeit der Messungen verschiedener Transistoren zu verbessern, können die Gatespannungen als Effektivwerte angegeben werden. Die Absolutwerte der Gatespannungen werden mithilfe der innerhalb der Nullstundencharakterisierung ermittelten Einsatzspannungen des Transistors berechnet: Die angelegte Spannung ergibt sich aus der Addition der Einsatzspannung und der effektiven Gatespannung. Der Einfluss variierender Einsatzspannungen auf die Kennlinien wird auf diesem Wege abgefangen.

3.3.2 Einsatzspannung

Wie in Kapitel 3.2.2 bereits beschrieben, existieren für den Parameter der Einsatzspannung unterschiedliche Methoden der messtechnischen Bestimmung [OrtizConde02]. Das für die Belastungsuntersuchungen entwickelte LabView-Programm unterstützt die Extraktion der Schwellenspannung nach mehreren Definitionen.

Es besteht ein enger Zusammenhang zwischen der Definition des Kennwertes V_{th} und der notwendigen Extraktionsmethodik. In diesem Kapitel soll näher darauf eingegangen werden.

3.3.2.1 Steilheitsmethode

Für die Methode des maximalen differentiellen Leitwertes zur Bestimmung der Einsatzspannung werden die Daten der aufgenommenen Eingangskennlinien weiter verarbeitet. Aus der aufgenommenen Kurve der Eingangskennlinie im linearen Bereich ermittelt die Messsoftware automatisch die Einsatzspannung $V_{th,lin}$ des Transistors im linearen Bereich. Dafür wird an den Punkt der Kennlinie mit maximaler Steigung eine Tangente gelegt. Deren Schnittpunkt mit der Abszisse (Gatespannungsachse) definiert die Schwellenspannung des Transistors im Linearen Bereich. In Abbildung 3.7 ist das Verfahren skizziert. Der Punkt der maximalen Steigung wird durch numerische Differentiation bestimmt. Der Einfluss von Diskretisierungsfehlern aufgrund der endlichen Anzahl von Messpunkten für die Drainstromkurve wird mittels mathematisch-numerischer Verfahren gemindert.

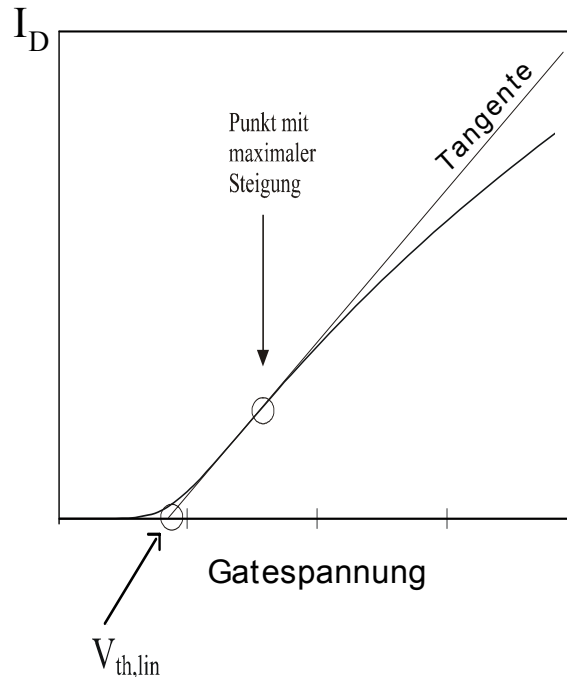


Abbildung 3.7: Numerische Bestimmung der Schwellenspannung des Transistors im Linearen Bereich aus Messung der Eingangskennlinie bei betragsmäßig kleiner Drainspannung

Aus der Eingangskennlinie bei höheren Drainspannungen wird die Schwellenspannung im Sättigungsbereich extrahiert. Das Verfahren zur Bestimmung des Wertes $V_{th,sat}$ unterscheidet sich leicht von dem bereits beschriebenen Verfahren zur $V_{th,lin}$ -Bestimmung [Thewes95, OrtizConde02].

Der erste Schritt beinhaltet die Auftragung der Quadratwurzelwerte des Drainstromes in Sättigung. Daraufhin wird wieder an den Punkt der maximalen Steigung eine Tangente gelegt. Der Schnittpunkt der Tangente mit der Gatespannungsachse bestimmt auch in diesem Fall die Einsatzspannung des Transistors. In Abbildung 3.8 wird das beschriebene Verfahren verdeutlicht.

Der Arbeitspunkt von Transistoren in Analoganwendungen wird durch den Parameter $V_{th,sat}$ weitaus besser beschrieben als durch den Wert der Schwellenspannung, der mit der Linearen Kennlinie ermittelt wird. Durch Bestimmung der Einsatzspannung sowohl im linearen als auch im Sättigungsbereich wird der Transistor genauer charakterisiert.

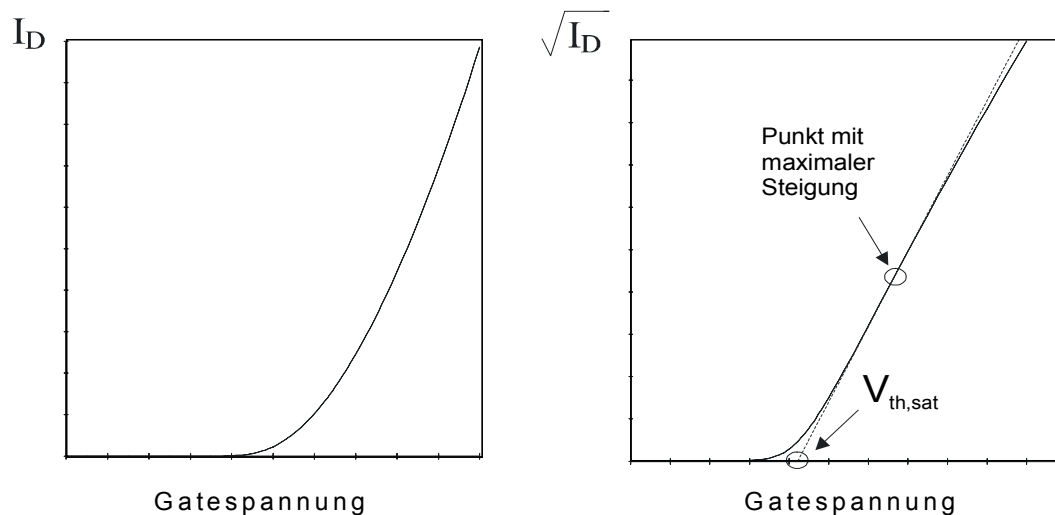


Abbildung 3.8: Bestimmung der Einsatzspannung im Sättigungsbereich

3.3.2.2 Zielstrommethode

Bei dieser Methode wird die Schwellenspannung des MOSFETs über einen zu erreichenden Stromwert pro Weite/Länge des Transistors definiert [OrtizConde02]. Die Gatespannung, die bei einer konstant anliegenden Drain-Source Spannung benötigt wird, damit das Bauelement einen festgelegten Strom leitet, wird als Einsatzspannung definiert. Da der Kanalstrom in guter Näherung direkt proportional zum Quotienten der Kanalweite und -länge ist, wird für die einheitliche Beschreibung der Einsatzspannung bei unterschiedlichen Geometrien eine Stromdichte festgeschrieben. Bei der Bestimmung der Schwellenspannung wird so die Devicegeometrie berücksichtigt.

Es wird hierbei eine Drain-Source Spannung an den Transistor angelegt und die Gatespannung so lange verändert, bis der durch die Geometrie und die geforderte Stromdichte definierte Drainstrom fließt. Eine präzise Bestimmung der korrekten Gatespannung kann über ein Newtonsches Iterationsverfahren erfolgen. Bei geringen Anforderungen an die Genauigkeit kann für Messwertextraktion die Gatespannung in einer Rampe über ein Intervall gefahren werden, in dem die Einsatzspannung zuvor abgeschätzt wurde. Der geringe Aufwand bei dieser Art der Einsatzspannungsextraktion führt speziell auch im ‚Inline-Monitoring‘, also bei der Prozesskontrolle noch innerhalb der Fertigungslinie, zu einer hohen Verbreitung.

3.3.2.3 DIBL-Effekt

Bei der Charakterisierung der Einsatzspannung der Transistoren werden wie bereits beschrieben typischerweise verschiedene Einsatzspannungswerte definiert. Die Werte werden sowohl im linearen Bereich des Transistors, als auch im Sättigungsmodus extrahiert. Die Werte $V_{th,lin}$ und $V_{th,sat}$ unterscheiden sich aufgrund des so genannten drain-induced-barrier-lowering-Effektes ('DIBL'), der bei höheren Drainspannungen auftritt [Ko89, Huang92, Liu93]. Dieser Effekt kann eine Verschiebung der Schwellenspannung um mehrere 10mV bis zu 100mV bewirken. Abbildung 3.9 zeigt den Verlauf der Einsatzspannung von Transistoren verschiedener Kanallänge als Funktion der anliegenden Drainspannung. Die Devices sind in der Technologie T11 (siehe Anhang A1) mit jeweils 10 μ m Weite gefertigt und wurden in der Länge variiert. Die Abhängigkeit der Einsatzspannungswerte von der Kanallänge, die trotz der verwendeten Stromdichte als Messkriterium hier deutlich wird, basiert auf dem Reverse-Short-Channel Effekt. Transistoren mit längeren Kanälen zeigen dabei betragsmäßig niedrigere Einsatzspannungen [Lu89, Mazuré89].

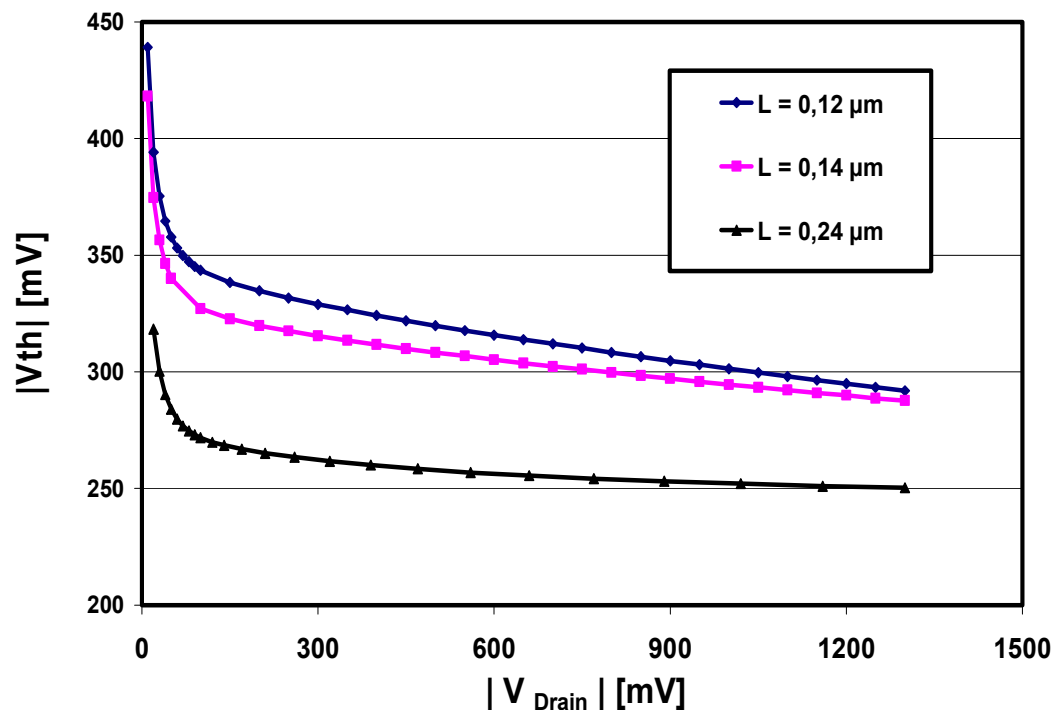


Abbildung 3.9: Abhängigkeit der Einsatzspannung von p-Kanal Transistoren von der Kanallänge und der Drainspannung. Längere Transistoren zeigen aufgrund des Reverse-Short-Channel-Effektes niedrigere Schwellenspannungen. Devices: $W/L = 10\mu\text{m} / 0,12 - 0,24\mu\text{m}$, Technologie T11.

Die Transistoren besitzen in den Bereichen oberhalb der so genannten ‚Pocket‘ oder ‚Halo‘-Implantationen an der Drain- und Source-Seite betragsmäßig höhere Einsatzspannungen als in den übrigen Bereichen. Bei längeren Kanallängen nimmt der Einfluss dieser Implantationsbereiche auf die Schwellenspannung des Transistors ab.

Um den DIBL-Effekt einzeln zu verdeutlichen, wird in der Abbildung 3.10 der Reverse-Short-Channel Effekt ausgeblendet. Hier werden nur die Abweichungen der Einsatzspannungen als Funktion der Drainspannungen aufgetragen. Die bei einer Drainspannung von 50mV ermittelte Schwellenspannung stellt dabei den Ausgangspunkt dar. Deutlich ist die starke Abhängigkeit der Einsatzspannung von der anliegenden Drain-Source Spannung ersichtlich. Mit höherer Drainspannung nimmt die Einsatzspannung ab. Der starke Anstieg der gemessenen Einsatzspannung bei sehr geringen Drainspannungen lässt sich anhand der einfachen Modellgleichung für den Kanalstrom im linearen Bereich erklären.

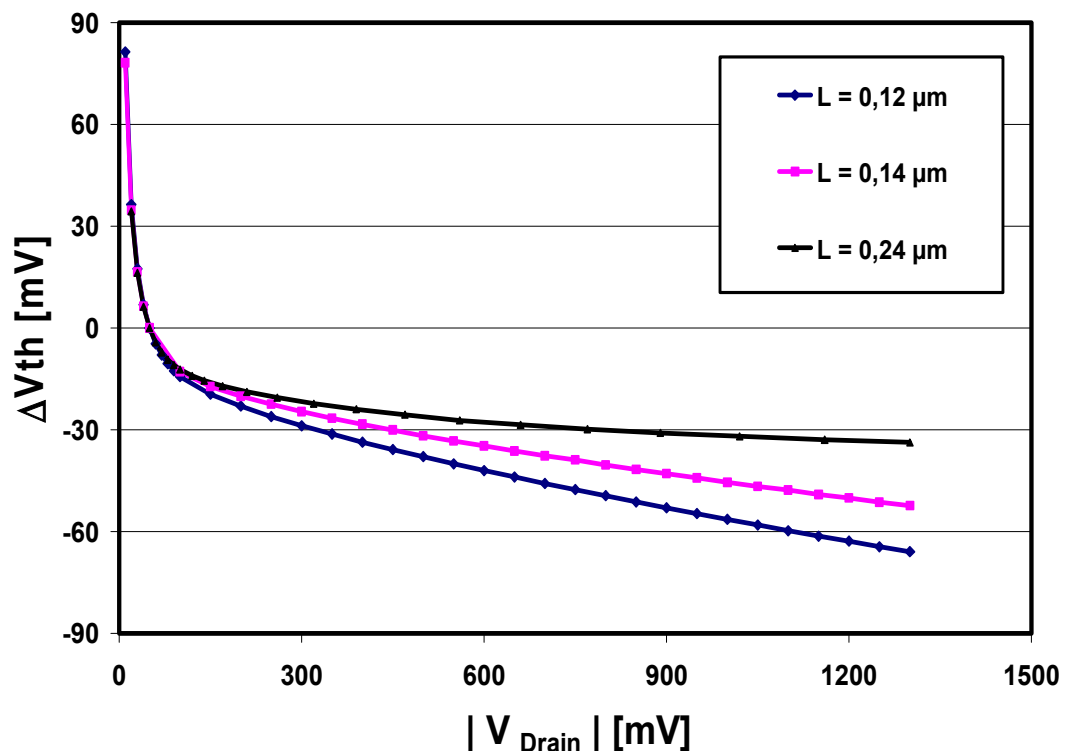


Abbildung 3.10: Abhängigkeit der Einsatzspannung von p-Kanal Transistoren von der Kanallänge und der Drainspannung normiert auf $V_{Drain} = 50\text{mV}$. Bei höheren Drainspannungen nimmt die Einsatzspannung aufgrund des DIBL-Effektes ab. Bei längeren Kanallängen tritt dieses Verhalten nur noch abgeschwächt auf. Devices: $W/L = 10\mu\text{m} / 0,12 - 0,24\mu\text{m}$, Technologie T11.

Nach [Goser90] gilt:

$$I_D = \frac{W}{L} \cdot \mu \cdot C_{OX} \cdot \left[(V_{GS} - V_{Th}) \cdot V_{DS} - \frac{1}{2} V_{DS}^2 \right] \quad (3.5)$$

Die Parameter W und L für die Geometrie, μ für die Ladungsträgerbeweglichkeit und C_{ox} für die Flächenkapazität stellen bei unserer Betrachtung Konstanten dar. Der Wert für den Kanalstrom I_D wird für die Messung entsprechend der Zielstrommethode ebenso auf einen durch Geometrie des Transistors und geforderter Stromdichte definierten Wert festgelegt.

Nach Umstellung der Gleichung ergibt sich:

$$V_{GS} - V_{Th} = \frac{1}{V_{DS}} \cdot \frac{I_D}{\frac{W}{L} \cdot \mu \cdot C_{OX}} + \frac{1}{2} V_{DS} \quad (3.6)$$

Da $\frac{I_D}{\frac{W}{L} \cdot \mu \cdot C_{OX}}$ hier konstant ist, nimmt der messtechnisch bestimmte Wert der Einsatzspannung bei kleinen Werten für die Drainspannung V_{DS} also umgekehrt proportional zu V_{DS} zu.

3.3.3 Charge-Pumping

Mithilfe von Charge-Pumping-Messungen werden Schädigungen an der Grenzfläche Si-SiO₂ des MOS-Transistors quantitativ erfasst [Heremans91, Groeseneken84]. Hier wird eine physikalische Veränderung des Transistors untersucht, aus der dann Veränderungen der elektrischen Eigenschaften des Bauelementes folgen können.

Wie bereits in Kapitel 2.2.1 beschrieben, kommt es bei der Prozessierung des isolierenden Gateoxides unumgänglich zur Bildung nicht abgesättigter Bindungen. Die Eigenschaften solcher Grenzflächenzustände und ihre Auswirkungen auf die elektrischen Parameter des Transistors wurden bereits zuvor detailliert diskutiert. Um herauszufinden, ob eine BT-Belastung zu einer Erhöhung der Zahl der Grenzflächenzustände führt und inwieweit dieser Effekt zur Degradation der elektrischen Eigenschaften der Transistoren nach BTS beiträgt, wird die Charge-Pumping-Charakterisierungsmethode eingesetzt. Die Eigenschaft der Grenzflächen-

zustände, Ladungsträger zu binden und wieder zu emittieren, wird ausgenutzt, um die Grenzflächenzustandsdichte des Transistors zu charakterisieren. Die Defektstellen werden gezielt geladen und entladen. Durch Messung des dabei fließenden Stromes können direkte Aussagen über die Anzahl der Grenzflächenzustände vor und nach der Belastung getroffen werden. Es werden akzeptor- und donatorartige Zustände parallel detektiert.

Die Konfiguration der Messung am Transistor ist in Abbildung 3.11 verdeutlicht. Die Vorgänge innerhalb des Bauelementes sollen am Beispiel eines p-Kanal Transistors erläutert werden. Bei einem n-Kanal MOSFET verlaufen die Prozesse während der Messung bis auf die Vorzeichen analog.

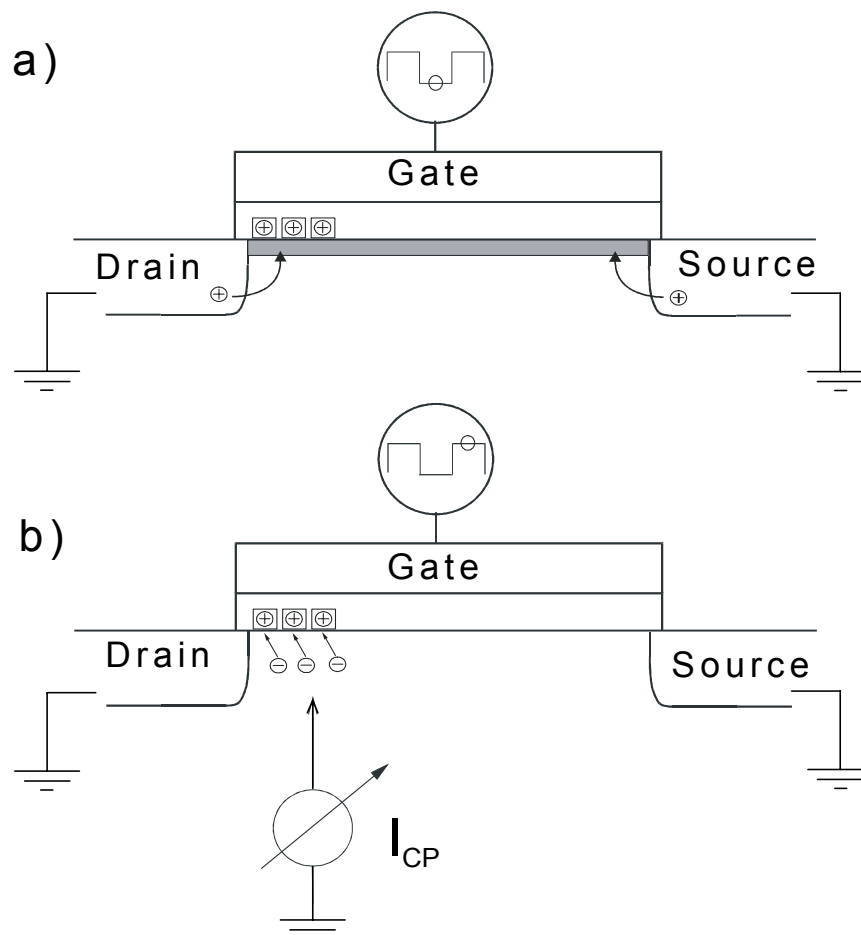


Abbildung 3.11: Messkonfiguration während der Charge-Pumping-Charakterisierungsmethode am Beispiel eines p-Kanal Transistors. An das Gate wird ein Trapezsignal angelegt. Die Störstellen werden zyklisch geladen und entladen. In dieser Darstellung wird vereinfachend nur das zyklische Laden (a) und Entladen (b) der donatorartigen Grenzflächenzustände aufgezeigt. Die akzeptorartigen Grenzflächenzustände werden jedoch parallel ebenso entladen (a) und beladen (b).

An das Gate des Bauelementes wird ein Trapezsignal V_{CP} angelegt, das den Transistor alternierend in Inversion und Akkumulation treibt. Die übrigen Knoten des Bauelementes liegen auf Massepotenzial. Der Stromfluss im Substratpfad wird mittels einer integrierenden Messung aufgezeichnet.

Ein hoher Signalpegel am Gate versetzt den Kanalbereich des zu vermessenden p-Kanal MOS-Transistors in Akkumulation ($V_{CP} > V_{FB,p}$), das durch den niedrigen Pegel hervorgerufene Feld sorgt für Inversion ($V_{CP} < V_{th,p}$). $V_{th,p}$ steht hier für die Schwellenspannung des p-Kanal Transistors, $V_{FB,p}$ für die Flachbandspannung.

Da die Zeitkonstanten für den Einfang eines Ladungsträgers in einen Grenzflächenzustand und für seine Emission aus diesem Zustand heraus wesentlich größer sind als die Zeitkonstante des Feldeffektes, ist es möglich, das Pumpen der Grenzflächenzustände vom Feldeffekt zu trennen. Die Vorgänge innerhalb des Transistors während der Messung lassen sich folgendermaßen aufteilen:

- Die Umladung des Kanalbereiches von Akkumulation zu Inversion erfolgt während der fallenden Flanke des Trapezsignals. Die benötigten Löcher für den sich aufbauenden Kanal kommen aus den p-dotierten Source- und Draingebieten.
- Während des Low-Levels am Gate ist der Kanalbereich des p-Kanal Transistors in vollständiger Inversion. In der Umgebung der Defektstellen ist die Zahl freier Löcher groß, sodass donatorartige Grenzflächenzustände positiv geladen werden (Abbildung 3.11 a)). Ladungen, die sich zu dieser Zeit in akzeptorartigen Störstellen befinden, können rekombinieren.
- Die Umladung des Kanalbereiches von Inversion zu Akkumulation erfolgt während der steigenden Flanke des Trapezsignals. Die Ladungsträger des sich abbauenden Kanals fließen zurück in die Diffusionsgebiete Drain und Source. Grenzflächenzustände, deren Zeitkonstanten größer sind als die Anstiegszeit der Flanke, bleiben zu diesem Zeitpunkt noch geladen.
- Während des High-Levels am Gate sind Elektronen im Kanalbereich die Majoritätsladungsträger. Die Elektronen rekombinieren nun mit den Ladungen der noch besetzten donatorartigen Grenzflächenzustände und besetzen akzeptorartige Grenzflächenzustände. Da die pn-Übergänge zur Source bzw. Drain gesperrt sind, kommen die benötigten Ladungsträger aus dem Substrat. Dieser Strom wird gemessen (Abbildung 3.11 b)).

Im gesamten Zyklus wird eine genau definierte Ladungsmenge von Drain und Source zu den Störstellen und dann zum Substrat transportiert bzw. aus dem Substrat zu den Störstellen, um später mit Minoritätskanalladungsträgern zu rekombinieren. Aus der Menge lässt sich direkt die Zahl der Grenzflächenzustände bestimmen, die im Laufe des Zyklus geladen und entladen werden.

Der zeitliche Mittelwert des fließenden Stromes I_{CP} ergibt sich aus der Periodendauer T_{Zyklus} des gesamten Charge-Pumping Zyklus und der Ladung Q_{it} , die sich während des Low-Pegels in den donatorartigen Grenzflächenzuständen und während des High-Pegels in den akzeptorartigen Grenzflächenzuständen befindet. Der Strom lässt sich unter Einführung der Grenzflächenzustandsdichte N_{it} , der Elementarladung q , der Frequenz $f = \frac{1}{T_{Zyklus}}$ und den Maßen des Transistorkanals

W, L ausdrücken:

$$I_{CP} = \frac{Q_{it}}{T_{Zyklus}} = N_{it} \cdot q \cdot f \cdot W \cdot L \quad (3.7)$$

Die in der Literatur übliche Größe der Grenzflächenzustandsdichte N_{it} lässt sich in die Anzahl der gepumpten Grenzflächenzustände N_{GFZ} umrechnen:

$$N_{GFZ} = N_{it} \cdot W \cdot L \quad (3.8)$$

Da die Zahl der Grenzflächenzustände, die während des Zyklus geladen und entladen werden, von den Pegeln und der Anstiegs- bzw. Abfallzeit der Flanken des angelegten Trapezsignals bedingt mitbestimmt werden, erhält man durch Variation dieser Größen während der Messung weitergehende Informationen. Eine Methode dafür beruht auf einer Änderung des Trapezpulses durch Variation des unteren und oberen Signallevels unter Beibehaltung der Pulsamplitude [Heremans91, Groeseneken84].

Der Betrag des im Substratpfad des Transistors gemessenen Stromes wird dabei als Funktion des unteren Niveaus des angelegten Trapezsignals aufgetragen. Während der Messung wird das Trapezsignal unter Beibehaltung seiner Amplitude durch Anhebung des unteren Spannungswertes über ein festgelegtes Spannungsintervall

verschoben. Die Parameter der Charge-Pumping Messung müssen auf die erlaubten Spannungswerte der zu vermessenden Transistoren angepasst werden. Innerhalb der gesamten Messung sollen nur Spannungen an die Gateelektrode gelegt werden, die den Transistor nicht außerhalb seiner Spezifikation belasten. Eine bereits durch seine Charakterisierung hervorgerufene Degradation wird auf diesem Wege ausgeschlossen.

Sollen z.B. Transistoren einer typischen $0,25\mu\text{m}$ -Technologie charakterisiert werden, die für $2,5\text{V}$ maximale Betriebsspannung spezifiziert ist, können beispielsweise folgende Parameter für die Messung gewählt werden: Das Trapezsignal besitzt eine Amplitude von $1,5\text{V}$ und beginnt mit einem unteren Level von $-2,5\text{V}$. Von dort aus wird das untere Niveau schrittweise auf 1V erhöht. Das abgedeckte Spannungsintervall ist symmetrisch. Zu Beginn der Messung alterniert das an das Gate angelegte Signal dementsprechend zwischen $-2,5\text{V}$ und -1V , bei Abschluss der Messung zwischen 1V und $2,5\text{V}$.

Über die Frequenz wird die Anzahl der Pumpzyklen innerhalb der Integrationszeit der Strommessung bestimmt. Die in dieser Arbeit durchgeführten Charge-Pumping-Charakterisierungen erfolgten soweit nicht explizit anders erwähnt bei einer Signalfrequenz von $f = 1\text{MHz}$.

In Abbildung 3.12 ist schematisch die Auftragungsart des Charge-Pumping Stromes und seine idealisierte Abhängigkeit vom Low-Level der angelegten Signale dargestellt. Während der Messung werden die drei in den Abbildungen gekennzeichneten Bereiche durchlaufen:

Bereich I:

Beim Start der Messung liegen sowohl das untere Signallevel $V_{CP,low}$ als auch das obere Level $V_{CP,high}$ betragsmäßig oberhalb der Schwellenspannung des p-Kanal Transistors ($V_{CP,low} < V_{th,p}$ und $V_{CP,high} < V_{th,p}$). Der Transistor befindet sich während des gesamten Spannungsbereiches des Signals in Inversion. Die Grenzflächenzustände werden nicht zyklisch beladen und entladen, da der Transistor keine Akkumulationsphase erreicht. Bei diesen Spannungswerten am Gate fließt im Idealfall kein Strom.

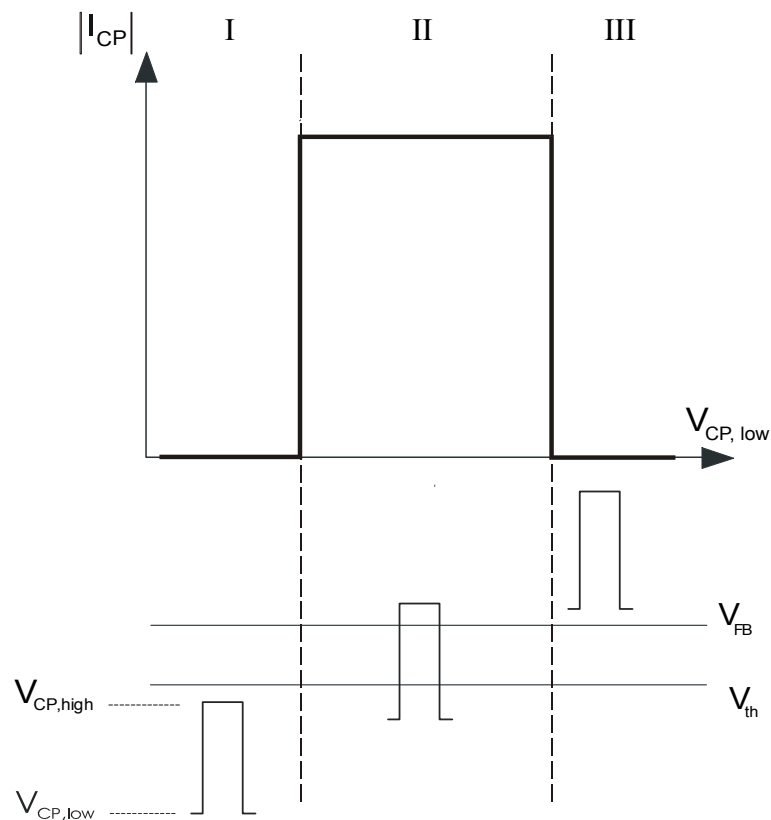


Abbildung 3.12: Idealisierter Verlauf des Charge-Pumping-Stromes in Abhängigkeit von dem am Gate angelegten Signal

Bereich II:

Werden die Signallevel im Laufe der Messung soweit erhöht, dass der obere Wert $V_{CP,high}$ die Flachbandspannung des p-Kanal Transistors überschreitet ($V_{CP,high} > V_{FB,p}$) und der untere Wert $V_{CP,low}$ weiterhin betragsmäßig oberhalb des Schwellenspannungswertes des Transistors liegt ($V_{CP,low} < V_{th,p}$), ist die Bedingung für einen Lade- und Entladezyklus der Grenzflächenzustände erfüllt. Alle erreichbaren Grenzflächenzustände werden nun durch den stetigen Wechsel von Akkumulation und Inversion zyklisch be- und entladen. Während die niedrige Spannung $V_{CP,low}$ an der Gateelektrode des Transistors liegt, befindet er sich in Inversion. Die donatorartigen Störstellen werden mit positiven Ladungsträgern aus dem aufgebauten Kanal besetzt, die Ladungen der akzeptorartigen Störstellen können rekombinieren.

Bei Erreichen des oberen Pulslevels $V_{CP,high}$ herrscht Akkumulation vor. Elektronen können mit den Ladungen der positiv besetzten Grenzflächenzustände rekombinieren, akzeptorartige Grenzflächenzustände werden besetzt. Der Charge-Pumping Strom ist im Idealfall in diesem Bereich nur noch eine Funktion der Anzahl der Grenzflächenzustände nach Gleichung (3.7). Die Kurve bildet in diesem Bereich dementsprechend ein Plateau aus.

Da jedoch Schwellen- und Flachbandspannung durch die komplexen Dotierungsprofile ortsabhängig sind und die Grenzflächenzustände nicht nur einen einzelnen diskreten Energiezustand aufweisen, sondern über der gesamten Bandlücke verteilt liegen, zeigt der Charge-Pumping Strom auch in diesem Bereich üblicherweise eine schwache Abhängigkeit von den Spannungsleveln des Trapezsignals.

Bereich III:

Nach weiterer Erhöhung der Signallevel liegen letztendlich beide Spannungslevel des Trapezsignals im Akkumulationsbereich des p-Kanal Transistors ($V_{CP,low} > V_{th,p}$ und $V_{CP,high} > V_{th,p}$). Das Bauelement befindet sich daher während des gesamten Signalverlaufes in Akkumulation. Die Grenzflächenzustände werden nicht mehr be- und entladen, da für einen Pumpvorgang die Inversionsphase und damit die nötigen Löcher fehlen. Der Charge-Pumping Strom fällt in diesem Bereich idealerweise auf Null ab.

Abbildung 3.13 zeigt die gemessene Charge-Pumping-Kurve eines Transistors vor und nach einer BT-Belastung in einer linearen und einer logarithmischen Darstellung. Der Transistor wurde in der Technologie T6 gefertigt (Anhang A1) und weist eine Geometrie von $W/L = 10\mu\text{m}/10\mu\text{m}$ auf. Das typische Plateau der Messkurve ist bei dem degradierten Transistor angehoben, d.h., die Belastung hat zur Bildung neuer Grenzflächenzustände geführt. Durch die Messung der Erhöhung des Charge-Pumping Stromes kann der erzeugte Schaden der Grenzfläche quantitativ erfasst werden.

Darüber hinaus soll anhand von Abbildung 3.13 noch auf die Abweichungen der gemessenen Kurve des undegradierten Transistors von der idealisierten Darstellung in Abbildung 3.12 eingegangen werden. Wie man sieht, erfolgt der Abfall des Stromes links und rechts des Plateaus nicht abrupt sondern graduell.

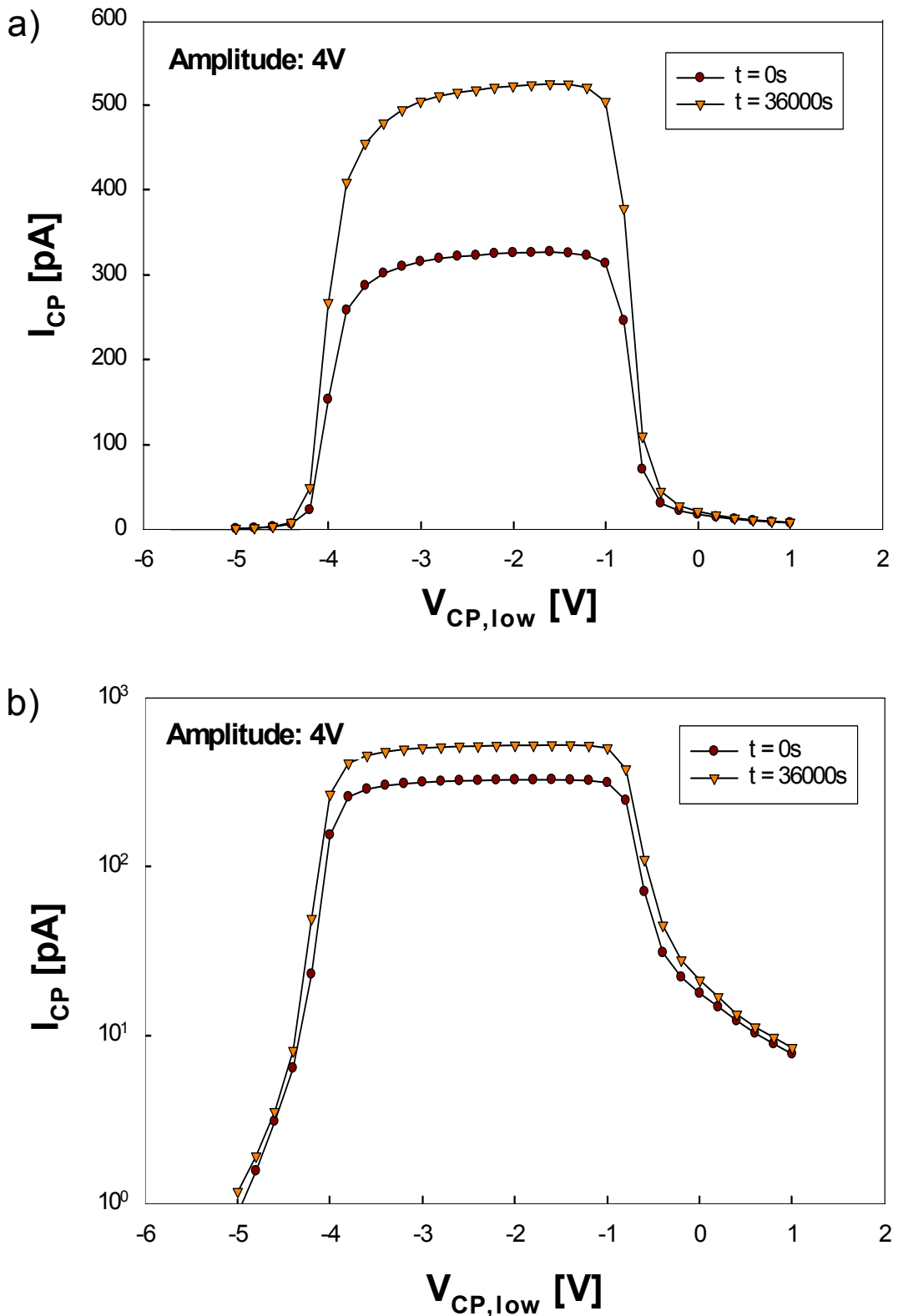


Abbildung 3.13: Charge-Pumping Strom als Funktion des unteren Signallevels $V_{CP,low}$ in linearer (a) und logarithmischer (b) Darstellung. Amplitude $V_{CP,high} - V_{CP,low} = 4V$. Die Charakterisierung erfolgt vor und nach einer BT-Belastung mit $V_{G,stress} = -8V$. Device: $W/L = 10\mu m/10\mu m$, Technologie T6.

Das liegt daran, dass die Schwellen- und Flachbandspannung von Transistoren ortsabhängig sind und, dass diese Parameter sich zu den Randbereichen ändern. Besonders bei einem LDD-Dotierungsprofil der Diffusionsgebiete, das bei den p-Kanal Transistoren der vermessenen Technologie T6 (Anhang A1) zum Einsatz kommt, ist dieser Effekt stark ausgeprägt [Heremans91, Thewes99_2].

Wird die Charge-Pumping Charakterisierungsmethode bei Transistoren mit sehr dünnen Gateoxiden durchgeführt, droht der Charge-Pumping-Strom durch signifikante Leckstromanteile überlagert zu werden. Zu der gepumpten Ladungsmenge der Grenzflächenzustände addiert sich der während der Messung durch das Gatedielektrikum fließende Leckstrom. Das Messsignal wird durch die Leckstromanteile verfälscht.

Eine Korrektur kann zum Beispiel durch Subtraktion einer Messung bei niedrigen Frequenzen (< 20 Hz) erfolgen [Sell02]. Solange das Dielektrikum nicht zu viele Grenzflächenzustände mit sehr langen Zeitkonstanten aufweist, ist der Charge-Pumping-Stromanteil bei diesen Frequenzen sehr gering.

Abbildung 3.14 zeigt Charge-Pumping-Kurven eines p-Kanal Transistors der Technologie T11 als ein Beispiel für Charge-Pumping-Charakterisierungen an Transistoren moderner Technologien. Die bereits besprochenen Abweichungen von der idealisierten Darstellung in Abbildung 3.12 treten hier noch deutlicher auf. Der Abfall des Charge-Pumping-Stromes in den Bereichen, in denen idealerweise kein zyklisches Be- und Entladen von Grenzflächenzuständen erfolgt (Bereich I und III in Abbildung 3.12), fällt geringer aus. Das Plateau der Kurve (Bereich III) bildet sich weniger deutlich aus. Die grundlegende Charakteristik der Charge-Pumping-Kurve bleibt jedoch erhalten.

Um die Aussagequalität der Kurve zu erhöhen, wird neben der regulären Charge-Pumping-Charakterisierung des Transistors mittels einer zweiten Messung der Leckstromanteil während der Charge-Pumping-Charakterisierung bestimmt. Während dieser Leckstrommessung werden zur Charge-Pumping-Messung identische Spannungswerte über dem Gateoxid angelegt und der Stromfluss durch das Oxid gemessen. Die ermittelten Werte werden von den überlagerten Charge-Pumping-Strömen subtrahiert, um leckstrombereinigte Messwerte zu erhalten. In Abbildung 3.14 werden die Kurven mit und ohne Korrekturmaßnahmen miteinander verglichen. Die korrigierte Kurve verläuft durch den Abzug der gemessenen Leckströme unterhalb der nicht korrigierten Kurve, die Unterschiede fallen jedoch gering aus, d.h., der Leckstromanteil fällt bei dem vermessenen Transistor gering aus.

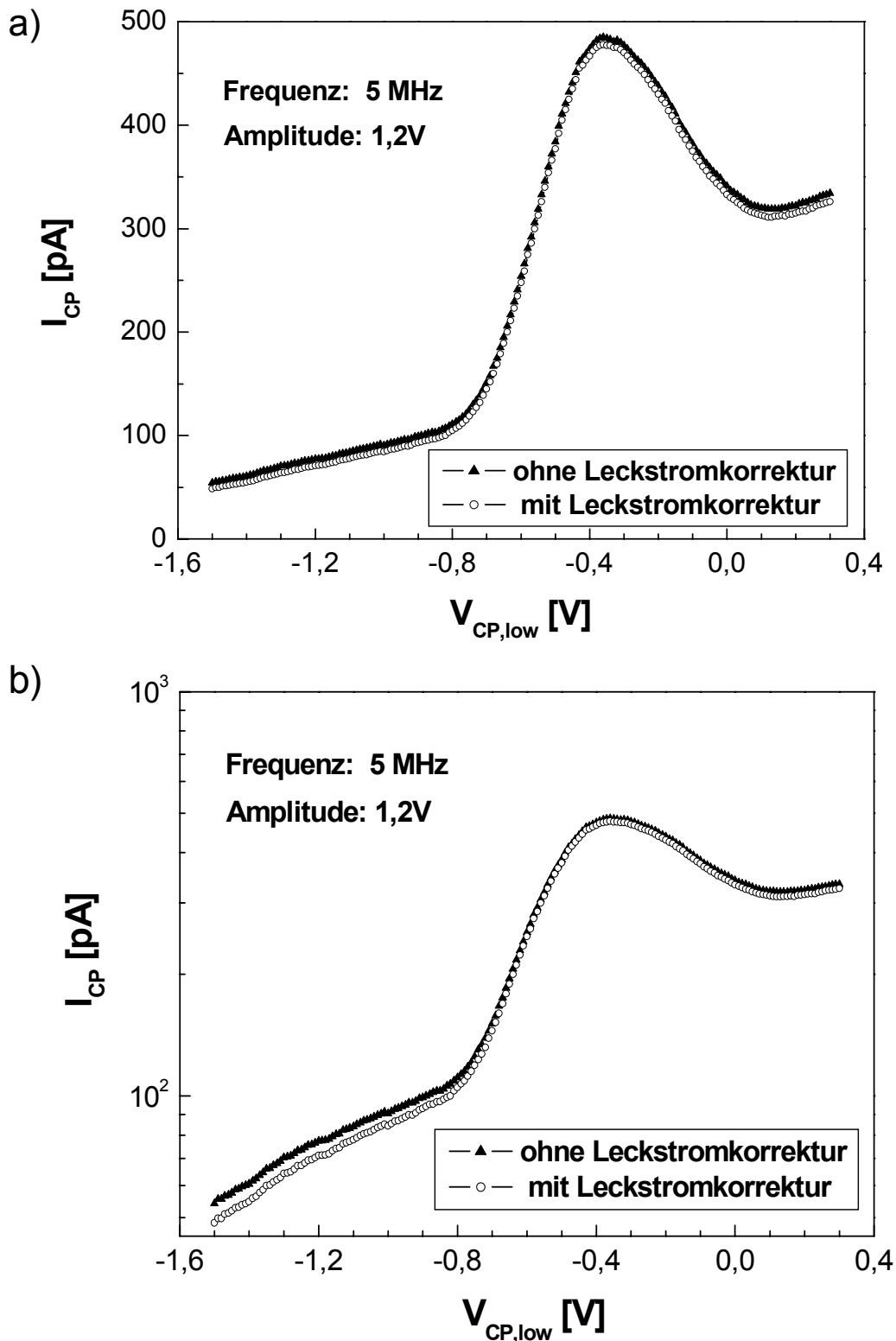


Abbildung 3.14: Charge-Pumping-Strom als Funktion des unteren Signalleivs $V_{CP,low}$ in linearer (a) und logarithmischer (b) Darstellung jeweils mit und ohne Leckstromkorrektur. Amplitude $V_{CP,high} - V_{CP,low} = 4V$. Die Charakterisierung erfolgt an einem unbelasteten Transistor mit der Geometrie $W/L = 40\mu\text{m}/0,12\mu\text{m}$ (Technologie T11). Die leckstromkorrigierte Kurve wird durch Subtraktion des mithilfe einer zweiten Messung ermittelten Gate-Leckstromanteils bestimmt.

4 Experimentelle Ergebnisse

In diesem Kapitel sollen die Ergebnisse der verschiedenen experimentellen Untersuchungen vorgestellt und diskutiert werden. Die Experimente lassen sich in zwei Gruppen aufteilen, in Untersuchungen mit homogener Belastung und Untersuchungen mit inhomogener Belastung. Die Unterscheidung bezieht sich auf das elektrische Feld, das während der Belastungssituation innerhalb des Bauelementes vorherrscht.

Das Degradationsverhalten unter diesen beiden BT-Belastungsbedingungen, die anhand von Beispielschaltungen bereits in Kapitel 2.1 diskutiert wurden, soll nun experimentell untersucht und diskutiert werden. Insbesondere werden auch Ergebnisse aus Belastungsexperimenten unter homogenen und inhomogenen Stressbedingungen miteinander verglichen.

Die Untersuchungen werden an verschiedenen Technologiegenerationen durchgeführt. Dies ermöglicht technologieunabhängige Aussagen und demonstriert darüber hinaus die Reproduzierbarkeit der Untersuchungsergebnisse.

4.1 Homogene Belastung

Die in Kapitel 2.3 beschriebenen Auswirkungen einer BT-Belastung lassen sich messtechnisch einfach nachweisen. Die Transistoren weisen nach BTS eine betragsmäßig höhere Einsatzspannung auf und leiten einen geringeren Kanalstrom. Abbildung 4.1 zeigt die Verschiebung der Einsatzspannung nach BTS als Funktion der Belastungszeit. Die Bestimmung der Einsatzspannungen erfolgt mithilfe des Zielstromkriteriums bei einer anliegenden Drainspannung von -50mV (Kapitel 3.3.2.2). Der Transistor entstammt der Technologie T10 (Anhang A1). Während der Belastungsphasen wird -2,4V an die Gateelektrode angelegt. Alle übrigen Knoten des Transistors liegen auf Masse. Der Wafer-Chuck (siehe Anhang A2) und damit der Transistor werden konstant auf $T_{stress}=125^{\circ}\text{C}$ gehalten.

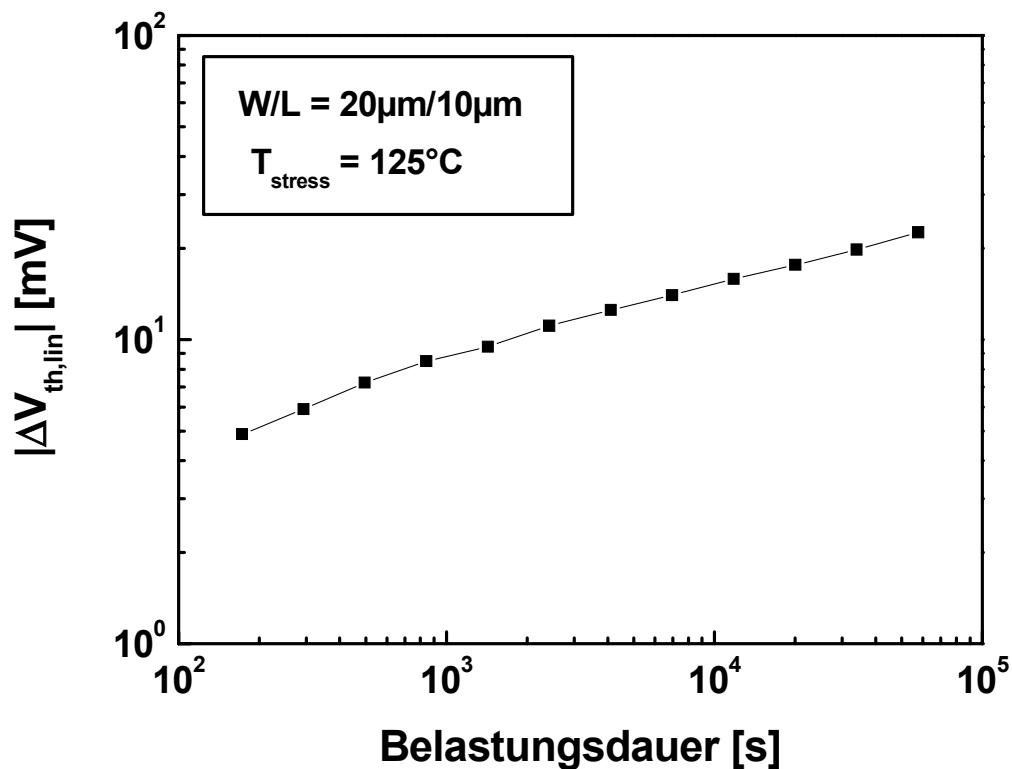


Abbildung 4.1: Verschiebung der Einsatzspannung nach BTS. Die Einsatzspannung wird mithilfe des Zielstromkriteriums bei einer Drainspannung von -50mV ermittelt. Belastung: $V_{G,stress} = -2.4V$, $t_{stress} = 16h$, $T_{stress} = 125^{\circ}C$, Technologie T10.

Die Einsatzspannung des Transistors nimmt bedingt durch die BT-Belastung betragsmäßig zu. Deutlich ist das der Parameterdegradation zugrunde liegende Potenzgesetz zu erkennen. Der Verlauf der Messwerte lässt sich durch eine Regressionsgerade annähern (vergleiche Kapitel 2.4).

Eine derartige Verschiebung der Einsatzspannung ist bei allen untersuchten Technologien zu beobachten, sie fällt jedoch bei unterschiedlichen Technologien auch bei vergleichbaren Belastungsbedingungen unterschiedlich stark aus. Abbildung 4.2 zeigt einen Vergleich der Lebensdauer von p-Kanal Transistoren der Technologien T7, T9, T10 und T11 (siehe Anhang A1) für verschiedene Belastungsspannungen. Auf der x-Achse wird jeweils der Quotient aus Belastungsspannung $V_{G,stress}$ und der für den untersuchten Transistor regulär spezifizierten Betriebsspannung V_{DD} aufgetragen. Für die Lebensdauer wird eine Verschiebung der Einsatzspannung von 100mV als Kriterium definiert.

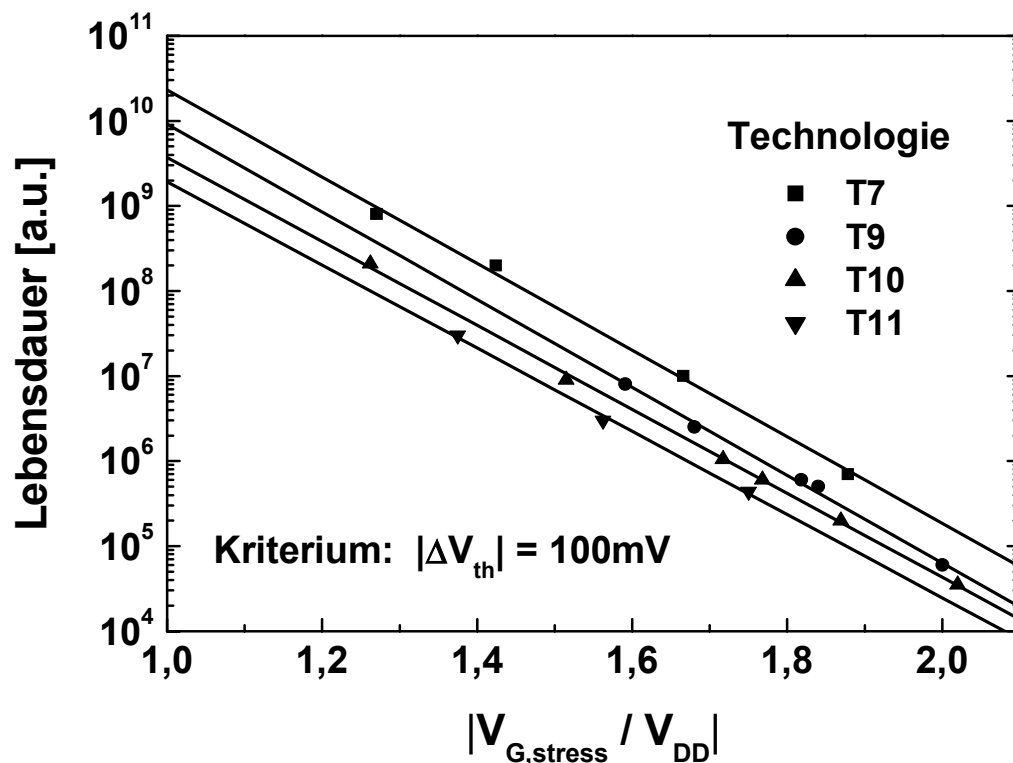


Abbildung 4.2: Lebensdauervergleich verschiedener CMOS-Technologiegenerationen als Funktion des Quotienten aus Belastungsspannung und spezifizierter Betriebsspannung V_{DD} . Als Lebensdauerkriterium wird eine Einsatzspannungsverschiebung nach BTS um 100mV definiert. Die Einsatzspannungen werden mithilfe des Zielstromkriteriums ermittelt. Belastung: $T_{stress} = 125^\circ\text{C}$, Technologie: T7, T9, T10, T11.

Durch Extrapolation wird jeweils die Lebensdauer bei nominalen Betriebsbedingungen ($|V_{G,stress} / V_{DD}| = 1$) bestimmt (vergleiche Kapitel 2.6).

Die Degradation der Einsatzspannung nach BTS und damit die Lebensdauern fallen bei den Transistoren der verschiedenen Technologien unterschiedlich stark aus. Bei den durchgeführten Experimenten fällt bei vergleichbaren Belastungsbedingungen ($V_{G,stress} / V_{DD}$) die Degradation mit jeder moderneren Technologie stärker und somit die erreichbare Lebensdauer niedriger aus.

Dies lässt sich hauptsächlich auf die bei Betriebsbedingungen vorherrschenden elektrischen Feldstärken über dem Gateoxid zurückführen. Die Feldstärken und die damit verbundene Belastung des Gateoxids nahmen in der Vergangenheit bei jeder neuen Technologiegeneration zu. Abbildung 4.3 zeigt die Feldstärke über dem Oxid bei nominellen Betriebsbedingungen für Transistoren verschiedener Technologiegenerationen auf. Auf der x-Achse wird die elektrische Gateoxidicke der Standard-

Logik-Transistoren der jeweiligen Technologiegenerationen aufgetragen (vergleiche Kapitel 3.2.1). Der am Ende der Kurve durch ein Rechteck markierte Bereich entspricht der IRTS-Vorhersage für die Feldstärke über dem Oxid bei Transistoren zukünftiger Halbleitertechnologien [ITRS04]. Die hohe elektrische Belastung der Gateoxide und die damit verbundene BTS-Problematik werden demnach auch bei zukünftigen Technologien vorherrschen und wahrscheinlich weiter zunehmen.

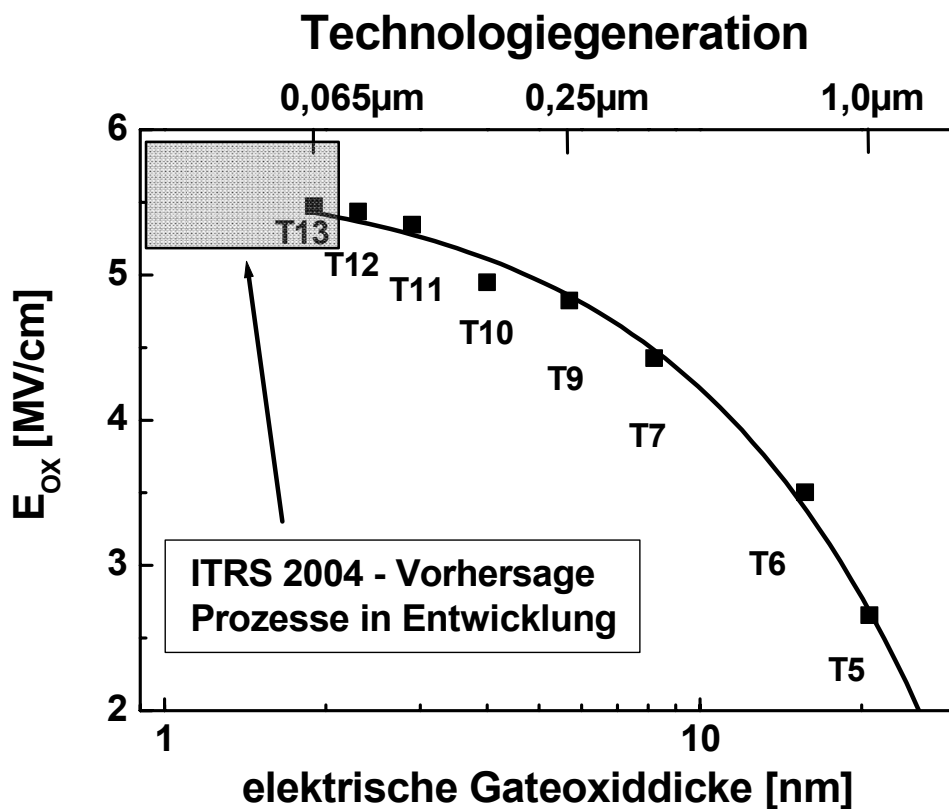


Abbildung 4.3: Oxidfeldstärke von Standard-Logik Transistoren verschiedener Technologiegenerationen bei jeweils nomineller Betriebsspannung. Der grau markierte Bereich entspricht der ITRS-Vorhersage.

Neben der Verschiebung der Einsatzspannung führt BTS auch zu einer Degradation des Kanalstroms. Der Strom, den der Transistor in einem bestimmten Arbeitspunkt leitet, fällt geringer aus. Der durch BT-Belastung eintretende Verlust an Stromtreibfähigkeit fällt jedoch nicht in allen Betriebsbereichen des Transistors gleich groß aus. Der Kanalstrom nimmt in Abhängigkeit des Arbeitspunktes des Bauelementes unterschiedlich stark ab.

Wird die Degradation des Drainstroms als Funktion der Gatespannung aufgetragen, ergibt sich das in Abbildung 4.4 gezeigte Bild. Jede einzelne Kurve steht für die prozentuale Verschiebung des Drainstroms zu einem Charakterisierungszeitpunkt in Bezug auf die Stromwerte vor der Belastung. Der eingezeichnete Pfeil kennzeichnet den zeitlichen Verlauf der Belastung. Die unterste Kurve entspricht der Kanalstromdegradation am Ende der 16 Stunden BT-Belastung.

Die Kurven zeigen eine typische Charakteristik der Drainstromdegradation nach BTS. In der Nähe der Schwellenspannung des Bauelementes fällt die Reduktion des Drainstroms besonders deutlich aus. In Arbeitspunkten mit Gatespannungen nahe der Einsatzspannung liefert der belastete Transistor nach 16 Stunden Belastung fast 30 Prozent weniger Strom. In Arbeitspunkten mit höheren Gatespannungen fällt die Reduktion des Kanalstromes schwächer aus. Bei voller Gatespannung ($V_G = V_{DD}$) liefert der Transistor nach der Stressbelastung etwa 5 Prozent weniger Strom.

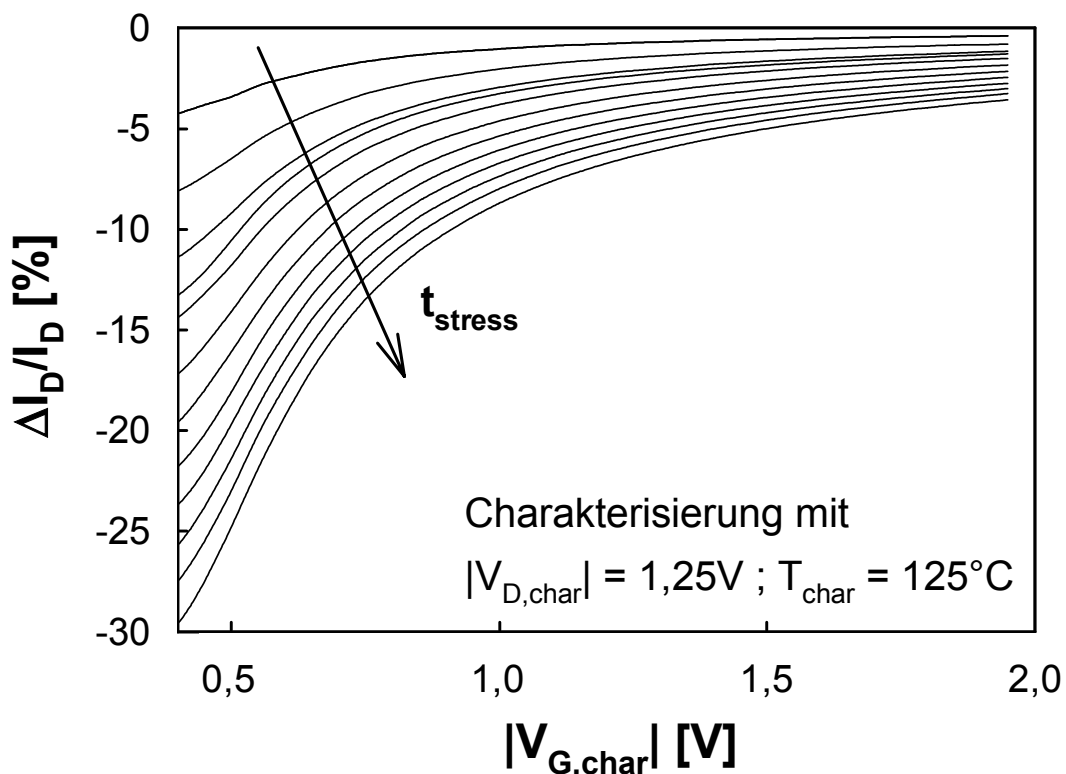


Abbildung 4.4: Degradation des Kanalstroms als Funktion der Charakterisierungs-Gatespannung. Jede Kurve repräsentiert eine Charakterisierungs-Messung nach einem Belastungsintervall.

Belastung: $V_{G,stress} = -2,4V$, $t_{stress} = 16h$, $T_{stress} = 125^{\circ}C$.

Device: $W/L = 20\mu m/10\mu m$, Technologie T10.

Das Ergebnis unterstreicht die bereits in Kapitel 2.5 getroffenen Aussagen über den Einfluss der Degradation nach BTS auf Analogapplikationen. Da hier die Transistoren typischerweise in Arbeitspunkten mit Gatespannungen von 100 – 300mV oberhalb der Schwellenspannung eingesetzt werden, fallen die Konsequenzen der Degradation für diese Schaltkreise sehr deutlich aus.

Neben dem betragsmäßigen Anstieg der Einsatzspannung und der Reduktion des Kanalstroms kann die Degradation nach BTS auch zu einer Beeinträchtigung der Beweglichkeit der Kanalladungsträger führen (siehe Kapitel 2.3). In Abbildung 4.5 wird die Steilheit g_m des bereits für Abbildung 4.4 untersuchten Transistors gezeigt. Das Maximum dieser Kurve stellt ein Maß für die Beweglichkeit dar.

Ferner ist in Abbildung 4.5 eine weitere Kurve gezeigt, die sich aus der nach Belastung ermittelten Kurve ergibt, wenn diese um den Wert der erlittenen Einsatzspannungsverschiebung verschoben wird.

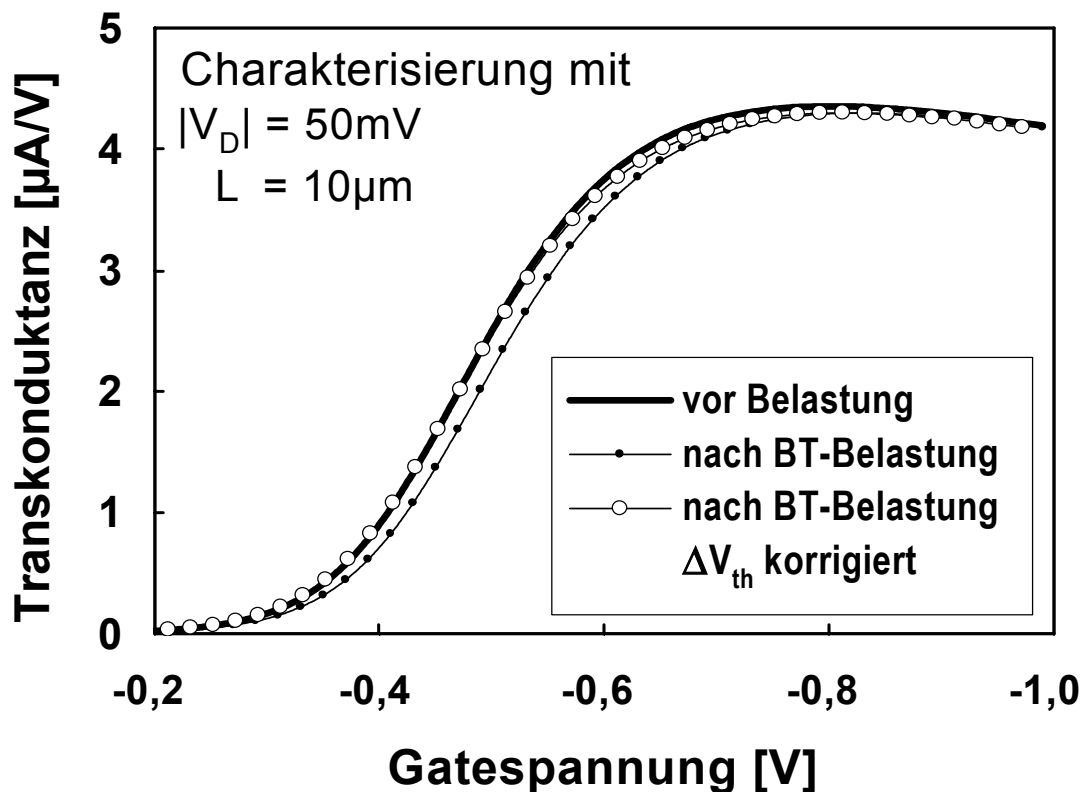


Abbildung 4.5: Differenzieller Eingangsleitwert vor und nach BTS. Die dritte Kurve (offene Symbole) wird durch Abzug der belastungsinduzierten Einsatzspannungsverschiebung gebildet. Belastungsbedingungen: $V_{G, stress} = -2,4\text{V}$, $t_{stress} = 16\text{h}$, $T_{stress} = 125^\circ\text{C}$. Device: $W/L = 20\mu\text{m}/10\mu\text{m}$, Technologie T10.

Anhand des Vergleiches dieser Kurve mit der Kurve vor Stress lässt sich eine mögliche Reduktion der Beweglichkeit ablesen. Das Maximum der korrigierten Kurve nach Stress liegt minimal tiefer als das der Kurve vor Stress. Eine Beweglichkeitsreduktion ist demnach zu erkennen, sie fällt jedoch sehr gering aus. Sie hat offensichtlich für die Drainstromdegradation nur eine geringe Bedeutung und kann für eine Modellierung daher weitestgehend vernachlässigt werden (Kapitel 5).

4.2 Inhomogene Belastung

Nach der Untersuchung des Degradationsverhaltens nach BT-Belastungen mit homogenen Belastungsbedingungen sollen nun inhomogene Belastungsbedingungen näher untersucht und diskutiert werden. Während der inhomogenen BT-Belastungen wird zusätzlich zur Belastungs-Gatespannung $V_{G,stress}$ eine Belastungs-Drainspannungen $V_{D,stress}$ an das Bauelement gelegt. Die dafür durchgeführten Vergleichsuntersuchen der verschiedenen Belastungsarten erfolgen auch hier an verschiedenen Technologien, um die Reproduzierbarkeit und Technologie-unabhängigkeit der Aussagen zu demonstrieren. Die Transistoren für die Untersuchungen entstammen den Standard CMOS-Technologien T9 und T10. Einzelheiten über die Technologien können aus Anhang A1 entnommen werden.

Für die Belastungsuntersuchungen kommen sowohl typische Transistorgeometrien für Analog- als auch für Digital- oder Hochfrequenz-Applikationen zum Einsatz. Darüber hinaus werden Transistoren mit einer Geometrie von $W/L = 20\mu\text{m}/10\mu\text{m}$ untersucht. Transistoren mit derart großen Kanallängen weisen einen vernachlässigbar kleinen Einfluss durch die Übergangsbereiche zwischen Kanal und Diffusionsgebiete auf und eignen sich daher besonders gut für die Untersuchung von Degradationsphänomenen im Kanalbereich.

In den folgenden Abbildungen werden Ergebnisse von BT-Belastungsexperimenten zur Untersuchung der Degradation der Einsatzspannung wiedergegeben. Die Experimente werden mit unterschiedlichen Drainspannungen während der Belastung durchgeführt. Als maximale Belastungs-Drainspannung wird hier $-1,2\text{V} (\leq 2/3 V_{DD})$ eingesetzt, um sicherzustellen, dass die Transistoren hauptsächlich einer BT-Belastung ausgesetzt sind. Bei höheren Belastungs-Drainspannungen können weitere zusätzliche Schädigungsmechanismen auftreten [Chaparala00].

Für Abbildung 4.6 wurden Transistoren mit der minimalen Kanallänge der gewählten Technologie ($0,18\mu\text{m}$) untersucht, wie sie typischerweise in Digitalschaltungen eingesetzt werden. Für Abbildung 4.7 werden Transistoren mit einer größeren Kanallänge ($0,5\mu\text{m}$) untersucht, wie sie üblicherweise in Analogapplikationen eingesetzt werden. Die Weite der Transistoren beträgt bei beiden Experimenten jeweils $10\mu\text{m}$. Die Einsatzspannungen werden auf Basis der Zielstrommethode ermittelt. Der Wert der Gatespannung wird als V_{th} definiert, bei der der Transistor eine Stromdichte von $70\text{nA} / (\text{W/L})$ leitet. (Vergleiche dazu Kapitel 3.3.2). Die dabei anliegende Drainspannung $V_{D,char}$ beträgt hier $-1,25\text{V}$. Bei beiden untersuchten Geometrien können wir eine schwächere Degradation bei betragsmäßig größeren Belastungs-Drainspannungen beobachten.

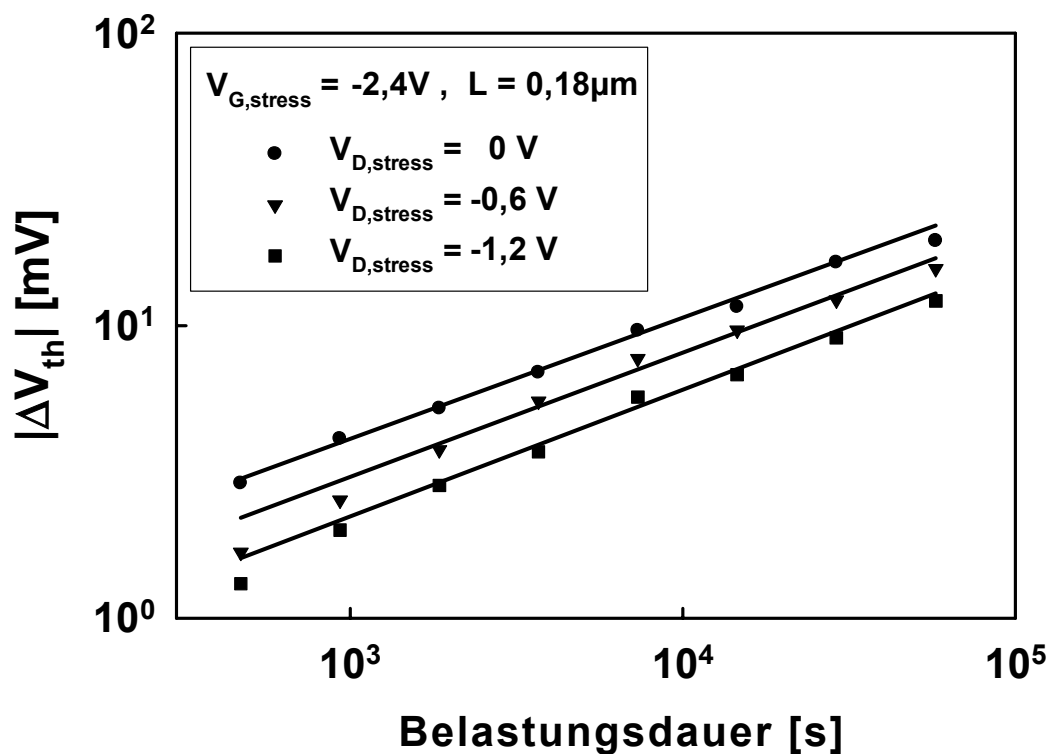


Abbildung 4.6: Degradation der Einsatzspannung als Funktion der Belastungsdauer für homogene und inhomogene BT-Belastungen. $T_{\text{stress}} = 125^\circ\text{C}$, weitere Belastungsparameter werden in der Legende angegeben. Die Einsatzspannung wird mithilfe des Zielstromkriteriums ermittelt. Device: $W/L = 10\mu\text{m}/0,18\mu\text{m}$, Technologie T10.

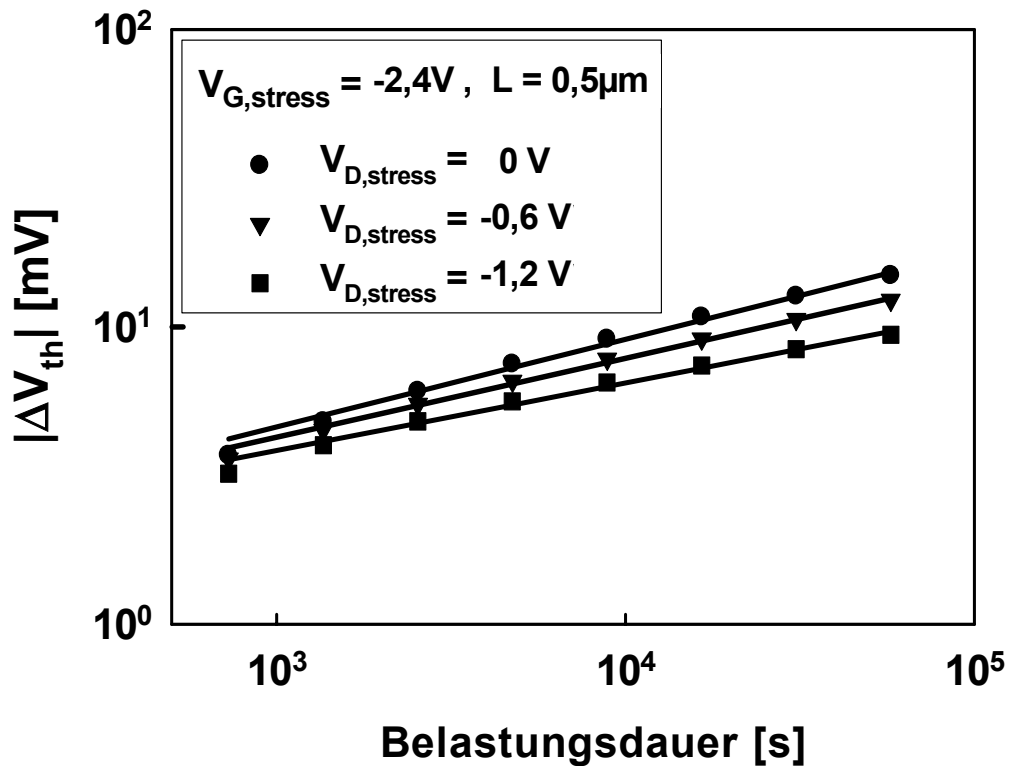


Abbildung 4.7: Degradation der Einsatzspannung als Funktion der Belastungsdauer für homogene und inhomogene BT-Belastungen.
 $T_{stress} = 125^\circ C$, weitere Belastungsparameter werden in der Legende angegeben.
 Die Einsatzspannung wird mithilfe des Zielstromkriteriums ermittelt.
 Device: $W/L = 10\mu m/0,5\mu m$, Technologie T10.

Die Wiederholung der Belastungsexperimente an Transistoren der Technologie T9 bestätigt dieses Degradationsverhalten bei Variation der Belastungs-Drainspannung. Abbildung 4.8 zeigt die Verschiebung der Einsatzspannung unter verschiedenen BT-Belastungen bei Transistoren der Technologie T9. Für die Belastung wurde hier eine Gatespannung von $V_{G, stress} = -3,3V$ gewählt. Die Belastungs-Drainspannung wurde zwischen 0 und $-1,5V$ variiert.

Auch hier kann die schwächere Degradation bei betragsmäßig höheren Drainspannungen beobachtet werden. Das unter inhomogenen Belastungsbedingungen schwächere Degradationsverhalten ist dementsprechend nicht auf die zuerst untersuchte Technologie beschränkt und kann als technologieunabhängig angenommen werden.

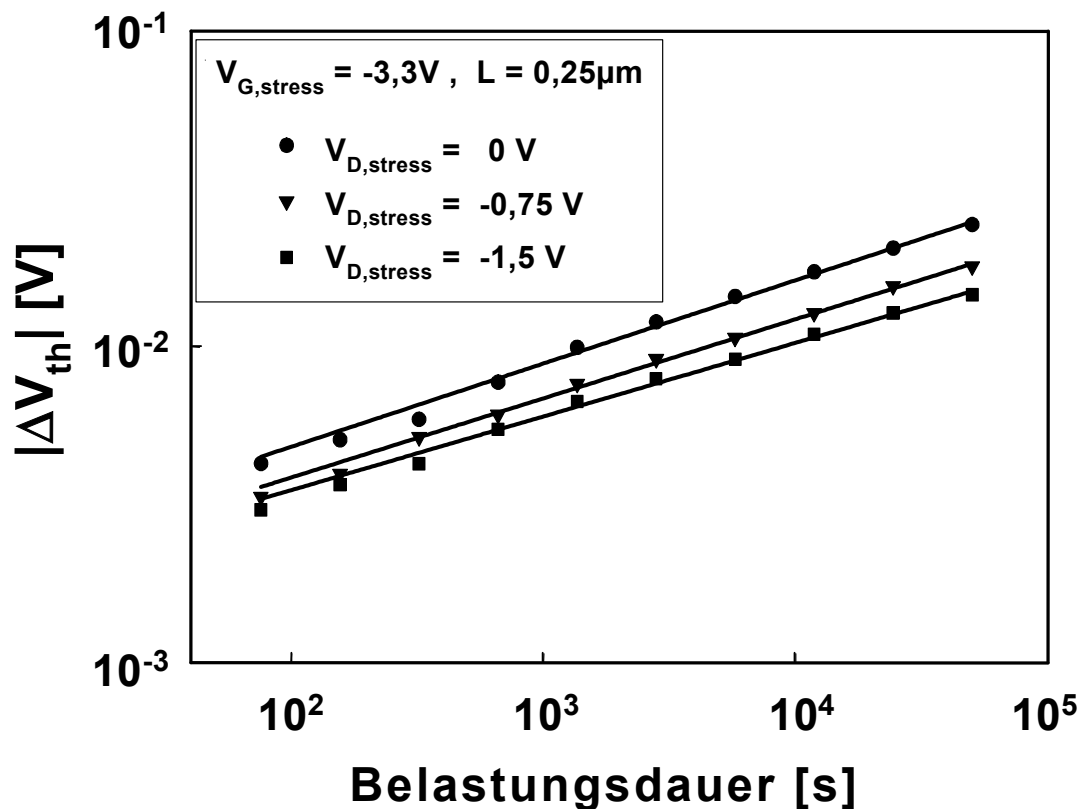


Abbildung 4.8: Degradation der Einsatzspannung als Funktion der Belastungsdauer für homogene und inhomogene BT-Belastungen.
 $T_{stress} = 125^{\circ}C$, weitere Belastungsparameter werden in der Legende angegeben.
 Die Einsatzspannung wird mithilfe des Zielstromkriteriums ermittelt.
 Device: $W/L = 10\mu m/0,25\mu m$, Technologie T9.

Die relativ gesehen schwächere Degradation der elektrischen Parameter unter inhomogenen Belastungsbedingungen kann auch anhand der Drainstromdegradation beobachtet werden. Abbildung 4.9 zeigt einen Vergleich der Drainstromdegradationen zwischen einer homogenen Belastung ohne anliegende Drainspannung ($V_{D, stress} = 0V$) (Abbildung 4.9 a)) und einer inhomogenen Belastung mit $V_{D, stress} = -1,2V$ während der Belastung (Abbildung 4.9 b)).

Die Diagramme zeigen die prozentuale Reduktion des Kanalstromes als Funktion der Charakterisierungs-Gatespannung. Jede einzelne Kurve steht für eine Messung der Drainstromdegradation nach einem Belastungsintervall (vergleiche Kapitel 3.1). Der innerhalb des Diagramms eingezeichnete Pfeil beschreibt den zeitlichen Verlauf der Degradation. Die jeweils untersten Kurven stehen für die prozentuale Verschiebung des Drainstroms nach 16 Stunden BT-Belastung mit $V_{G, stress} = -2,4V$.

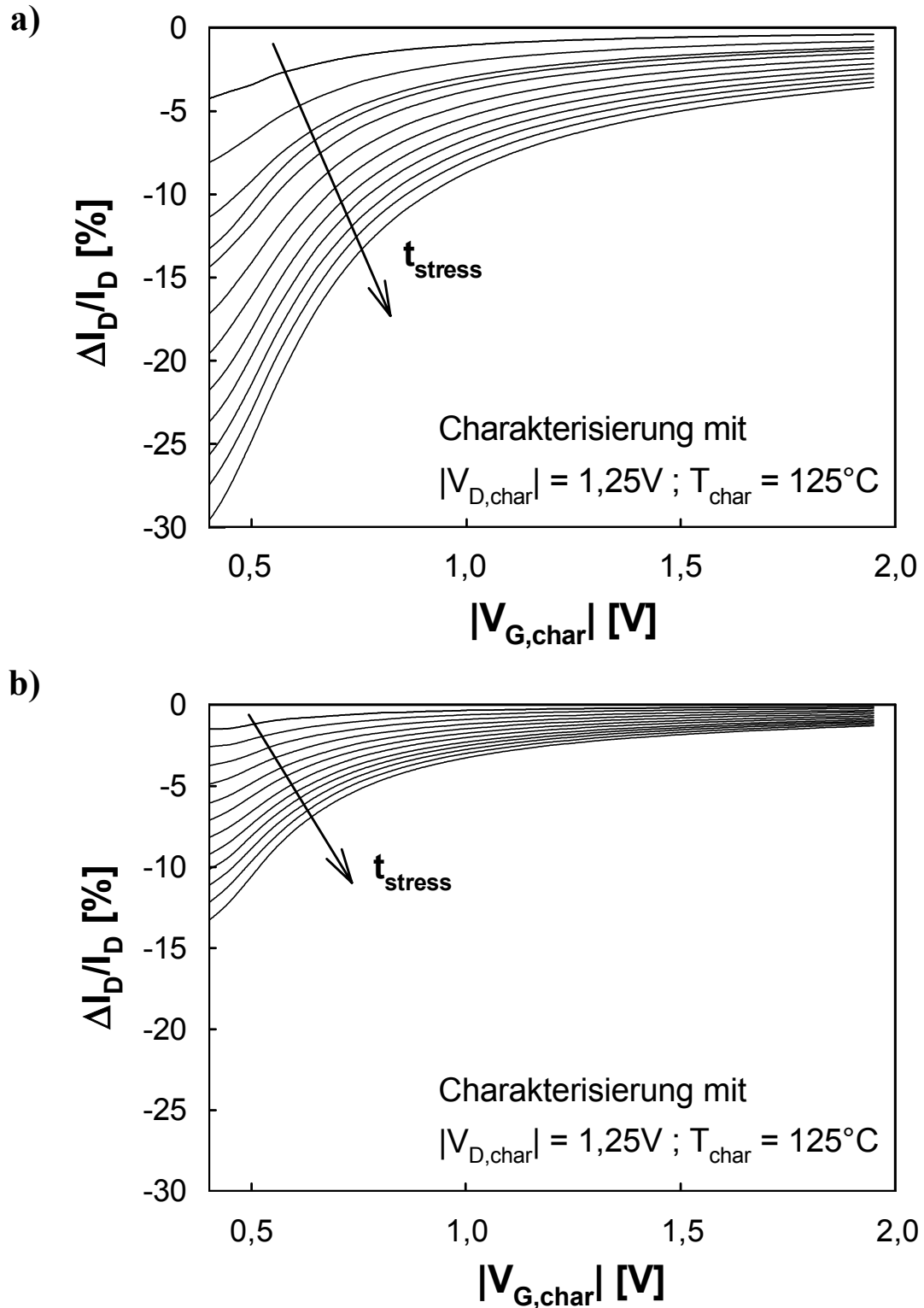


Abbildung 4.9: Degradation des Kanalstroms als Funktion der Charakterisierungs-Gatespannung und der Belastungszeit. Jede Kurve repräsentiert eine Charakterisierungs-Messung nach einem Belastungsintervall. Die unterste Kurve stellt jeweils die prozentuale Verschiebung des Stroms nach 16 Stunden BTS mit $V_{G,stress} = -2,4V$, $T_{stress} = 125^{\circ}C$ dar. a) $V_{D,stress} = 0V$, b) $V_{D,stress} = -1.2V$. Devices: $W/L=20\mu m/10\mu m$, Technologie T10.

Wir können unter beiden Belastungsbedingungen das typische Degradationsverhalten unter BTS beobachten. Die stärkste Drainstromdegradation liegt bei Arbeitspunkten mit Gatespannungen nahe der Einsatzspannung. In Arbeitspunkten mit höheren Gatespannungen fällt sie dagegen schwächer aus. Die Form der Kurven bildet sich in beiden Fällen sehr ähnlich aus, die Stärke der Drainstromdegradation fällt jedoch im inhomogenen Fall deutlich schwächer aus.

Für eine weiterreichende Charakterisierung der Transistoren wurde zusätzlich zu der Einsatzspannungs- und Drainstromcharakterisierung die Grenzflächenzustandsdichte mithilfe der in Kapitel 3.3.3 beschriebenen Charge-Pumping Technik bestimmt. Es wird das Verfahren mit konstanter Amplitude eingesetzt. [Heremans91, Groeseneken84].

Dafür wird ein Trapezsignal mit einer Amplitude von 1,5V an das Gate des Transistors gelegt. Das untere Level des Signals wird unter Beibehaltung der Amplitude von -1,9V bis auf 0,4V erhöht. Für die Frequenz des Trapezsignals wird $f = 1\text{MHz}$ eingestellt. Die Anstiegs- bzw. Abfallzeit der Flanken wird identisch mit 50ns festgelegt. Während der Charge-Pumping Charakterisierung wird der Wafer auf der bereits für die Belastungsphasen eingestellten Temperatur von $T_{stress} = 125^\circ\text{C}$ gehalten (vergleiche auch Anhang A2).

Abbildung 4.10 zeigt den Vergleich des belastungsinduzierten Anstiegs der gemessenen Grenzflächenzustandsdichten eines homogen belasteten Transistors und eines inhomogen belasteten Transistors. Für die Belastungsspannungen wurde wieder $V_{G,stress} = -2,4\text{V}$ und $V_{D,stress} = -0\text{V}$ für den homogenen Fall und $V_{G,stress} = -2,4\text{V}$ und $V_{D,stress} = -1,2\text{V}$ für den inhomogenen Fall gewählt. Das Diagramm zeigt die Messung an Transistoren mit der Geometrie $W/L = 20\mu\text{m}/10\mu\text{m}$. Aus dem Ergebnis lassen sich zwei Aussagen ableiten: Einerseits lässt sich unter inhomogenen Bedingungen ein schwächerer Anstieg der Anzahl der Grenzflächenzustände nachweisen, andererseits fällt der Anstieg in beiden Fällen relativ gering aus.

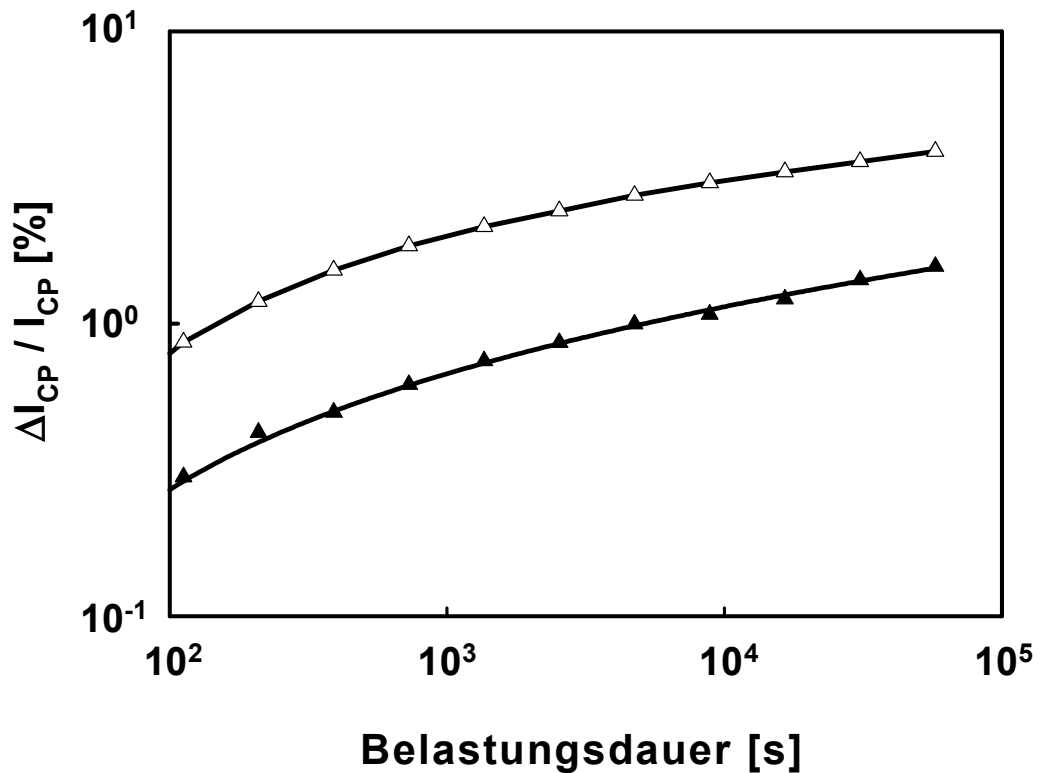


Abbildung 4.10: Prozentualer Anstieg des Charge-Pumping Stromes als Folge zusätzlicher belastungsinduzierter Grenzflächenzustände. Die Messkurve mit den offenen Symbolen steht für homogene BT-Belastung, die mit gefüllten Symbolen für inhomogene Belastung. Belastung: $V_{G, stress} = -2.4V$, $V_{D, stress} = 0V$ bzw. $V_{D, stress} = -1.2V$. Device: $W/L = 20\mu m/10\mu m$, Technologie T10.

Mithilfe der folgenden Gleichung [Heremans91, Groeseneken84] kann die durch den Anstieg der Grenzflächenzustände bedingte Einsatzspannungsverschiebung bestimmt werden:

$$\Delta V_{th} = q \cdot \Delta N_{it, donor} / C_{OX} \quad (4.1)$$

Hierbei steht q für die Elementarladung und $\Delta N_{it, donor}$ für die stressinduzierte Zunahme der donatorartigen Grenzflächenzustände. C_{OX} beschreibt die Gateoxidkapazität pro Fläche.

Mit dieser Abschätzung ergibt sich eine nur sehr schwache berechnete Verschiebung der Einsatzspannung im Bereich unterhalb von 1mV. Das lässt die Schlussfolgerung

zu, dass langsame Grenzflächenzustände, die durch ihre langen Zeitkonstanten bei der Charge-Pumping Charakterisierung mit 1MHz nicht erfasst werden, für die beobachtete Degradation verantwortlich sind. Eine weitere Erklärung für die Einsatzspannungsverschiebung können unter der BT-Belastung neu eingebaute feste positive Ladungen darstellen (vergleiche Kapitel 2.2.2), die jedoch üblicherweise bei dünnen Oxiden nur eine untergeordnete Rolle spielen [LaRosa97, Schlünder99].

5 Modellierung

Der Schädigungsverlauf und die daraus resultierenden Veränderungen des elektrischen Verhaltens des Bauelementes unter inhomogenen Bedingungen können durch die bisherigen klassischen Modelle nicht hinreichend genau beschrieben werden. Für die Drift des Kanalstromes und der Schwellenspannung ergeben sich unter inhomogenen Belastungsbedingungen geringere Werte als bei homogener Belastung. Eine Bewertung der Zuverlässigkeit des Transistors allein unter der klassischen, in Kapitel 2.6.1 beschriebenen ‚worst-case‘-Bedingung $V_G = V_{G, stress}$ und $V_{DS} = 0V$ kann folglich zu einer zu kritisch errechneten Lebensdauer führen.

Bei modernen CMOS-Technologien nimmt die Parameterdegradation teilweise so große Werte an, dass in vielen Fällen der ‚worst-case‘-Ansatz nicht mehr sinnvoll eingesetzt werden kann. Es kann der Fall eintreten, dass die untersuchte Technologie unter den bei der ‚worst-case‘-Betrachtung zu hoch angenommenen Belastungen das Lebensdauerziel nicht erfüllt, unter realen Anforderungen innerhalb des Produktes jedoch sicher die geforderte Lebensdauer erreicht. Die Forderung nach der Erfüllung der Lebensdauerkriterien unter ‚worst-case‘-Bedingungen würde in diesen Fällen unnötig hohe Aufwände und Kosten bei der Technologie- und Produktentwicklung generieren (‚over-engineering‘).

In diesem Kapitel wird ein neues analytisches Modell entwickelt, das das Degradationsverhalten auch in Abhängigkeit inhomogener Belastungsbedingungen beschreiben kann und somit nicht auf den ‚worst-case‘ homogener Belastungen beschränkt ist. Nach der Herleitung wird das Modell anhand gemessener Daten verifiziert und die Vorhersagequalität für die Drainstromdegradation diskutiert. Darüber hinaus wird eine Möglichkeit vorgestellt, die Komplexität der Modellgleichung zu reduzieren. Für den wichtigen analogen Fall müssen dabei nur geringe Einbußen bei der Genauigkeit in Kauf genommen werden. Abschließend werden Konsequenzen für die Lebensdauer einer Schaltung diskutiert und Richtlinien für die Lebensdauerextrapolation gegeben [Schlünder03, Schlünder05_1].

5.1 Herleitung der Modellgleichungen

Als ersten Schritt berechnen wir die belastungsbedingte Degradation der Einsatzspannung als Funktion des Ortes entlang des Kanals innerhalb des Transistors. Dann wird der Einfluss dieser lokalen Degradation auf die Strom-Spannungs-Charakteristik des Bauelementes bestimmt. Da die Einsatzspannung keinen adäquaten Parameter darstellt, um physikalisch die inhomogenen Degradationseffekte abzubilden, nutzen wir den Drainstrom als Charakterisierungsparameter. Wir betrachten typische analoge und Hochfrequenz-Arbeitspunktbedingungen, in denen die Transistoren im Sättigungsbereich arbeiten. Bei analogen Arbeitspunkten werden relativ geringe Spannungen an die Gateelektrode angelegt, während bei typischen Hochfrequenz-Applikationen auch mittlere und hohe Gatespannungen vorkommen. Der Einfachheit halber werden in diesem Kapitel alle Parameter jeweils als absolute Beträge angegeben.

Aufbauend auf der exponentiellen Abhängigkeit der Degradation von dem anliegenden Potenzial zwischen Gateelektrode und Kanal [Jeppson77, Ogawa95, LaRosa97, Schlünder99] kann die Verschiebung der Einsatzspannung $\Delta V_{th}(y)$ folgendermaßen beschrieben werden:

$$\Delta V_{th}(y) = V_0(t_{stress}, T_{stress}) \cdot e^{[C \cdot (V_{G, stress} - V_{Ch, stress}(y))]} \quad (5.1)$$

Dabei stellt $V_0(t_{stress}, T_{stress})$ eine von der Stresszeit und Stresstemperatur abhängige Konstante dar, C eine technologieabhängige Konstante und $V_{ch, stress}(y)$ das Kanalpotenzial während des Stresses an der Position y entlang des Kanals. Für die Source-Seite wird $y = 0$ und für die Drain-Seite $y = L$ festgelegt.

Nehmen wir für das Kanalpotenzial eine lineare Approximation gemäß

$$V_{ch, stress}(y) = V_{D, stress} \cdot y / L \quad (5.2)$$

an, so erhalten wir schließlich für das Kanalpotenzial unter Belastung folgende Gleichung:

$$\Delta V_{th}(y) = V_0(t_{stress}, T_{stress}) \cdot e^{[C \cdot (V_{G, stress} - V_{D, stress} \cdot y / L)]} \quad (5.3)$$

Nach Abschätzung der lokalen Einsatzspannungsdegradation soll nun der Einfluss auf den Drainstrom bestimmt werden. An jeder Stelle innerhalb des Kanals, an der die Inversionsbedingungen erfüllt ist, gilt:

$$I + \Delta I = W \cdot \mu(y) \cdot C_{OX} \cdot [V_G - (V_{th} + \Delta V_{th}) - V_{ch}(y)] \partial V_{ch}(y) / \partial y \quad (5.4)$$

Hier steht I für den Kanalstrom, ΔI für die stressinduzierte Änderung des Kanalstromes und $V_{ch}(y)$ für das Kanalpotenzial unter Charakterisierungsbedingungen.

Die Beweglichkeit der Ladungsträger $\mu(y)$ entlang des Kanals an der Position y hängt von der vertikalen elektrischen Feldstärke über dem Oxid ab. Zur Beschreibung dieser Abhängigkeit nutzen wir folgende Näherung [Ko89]:

$$\mu(y) = \mu_0 / \{1 + \Theta \cdot [V_G - V_{th} - V_{ch}(y)]\} \quad (5.5)$$

Hier stellt μ_0 die maximale Beweglichkeit bei schwachen vertikalen Feldern dar. Der Parameter θ ist ein Maß für die Beweglichkeitsreduktion mit steigender lokaler Gate-zu-Kanalspannung. Exakt würde hier auch die belastungsinduzierte Verschiebung der Schwellenspannung $\Delta V_{th}(y)$ Einfluss auf die Beweglichkeit nehmen. Der Einfluss ist in den für die Belastungsrechnung relevanten Arbeitspunkten jedoch relativ klein und wird daher hier vernachlässigt.

Wir nutzen eine lineare Näherung für das Kanalpotenzial unter Charakterisierungsbedingungen (vergleiche Gleichung (5.2)) und nehmen für das Kanalpotenzial am Abschnürpunkt des Kanals $V_G - V_{th}$ an, sodass

$$V_{ch(y)} = (V_G - V_{th}) \cdot \frac{y}{(L - L_{sat})} \quad (5.6)$$

gilt. L_{sat} steht hierbei für die Länge des Kanals zwischen Abschnürpunkt und Drain. Substituiert man Gleichung (5.4) mit Hilfe von Gleichung (5.6), so erhält man einen Ausdruck, der über die gesamte invertierte Länge des Kanals integriert werden kann. Die Integralgrenzen liegen dann bei $y=0$ für die Source-Seite und $L-L_{sat}$ für das andere Ende des Kanals am Abschnürpunkt.

Wenn man die Grenze $L-L_{sat}$ unter Vernachlässigung der Abschnürlänge L_{sat} durch L ersetzt, erhalten wir:

$$(I + \Delta I) \cdot L = \int_0^L (I + \Delta I) dy \quad (5.7)$$

$$= W\mu_0 C_{OX} \cdot \frac{V_G - V_{th}}{L} \int_0^L \frac{(V_G - V_{th}) \left(1 - \frac{y}{L}\right) - V_0 \cdot e^{\left[C \cdot \left(V_{G, stress} - V_{D, stress} \cdot \frac{y}{L} \right) \right]}}{1 + \Theta \cdot \left\{ (V_G - V_{th}) \cdot \left(1 - \frac{y}{L}\right) \right\}} dy$$

Dieses Integral kann auf analytischem Weg gelöst werden und führt zu einer analytischen Gleichung, die den Drainstrom vor und nach BT-Belastung beschreibt. Die sich ergebenden Ausdrücke sind zu komplex, um hier im Detail beschrieben zu werden. Im Anhang A3 werden jedoch die softwaregestützt ermittelte analytische Lösung und die wichtigsten Zwischenschritte der Berechnungen aufgelistet.

Im folgenden Unterkapitel wird diese Modellgleichung benutzt, um die Degradation von p-Kanal Transistoren unter verschiedenen Belastungsbedingungen und für verschiedene Arbeitspunkte im Betrieb vorherzusagen. Die berechneten Daten werden mit tatsächlichen experimentell ermittelten Werten verglichen.

5.2 Verifikation

In Abbildung 5.1 ist die Drainstromdegradation als Funktion der während der Belastung anliegenden Drainspannung aufgetragen. Es sind Werte für vier verschiedene Charakterisierungs-Gatespannungen abgebildet. Die Punkte entsprechen den gemessenen Werten, die durchgezogenen Kurven spiegeln die modellierten Daten wider. Punkte mit identischer Belastungs-Drainspannung stehen für die bei jeweils einem Belastungsexperiment ermittelten Daten. Für die Untersuchung wurden jeweils Langkanaltransistoren der Technologie T10 (Anhang A1) einer BT-Belastung ausgesetzt. Für die Belastungs-Gatespannung wurde $V_{G, stress} = -2,4V$ gewählt, die Belastungs-Drainspannung $V_{D, stress}$ wurde zwischen 0V und -1,25V variiert.

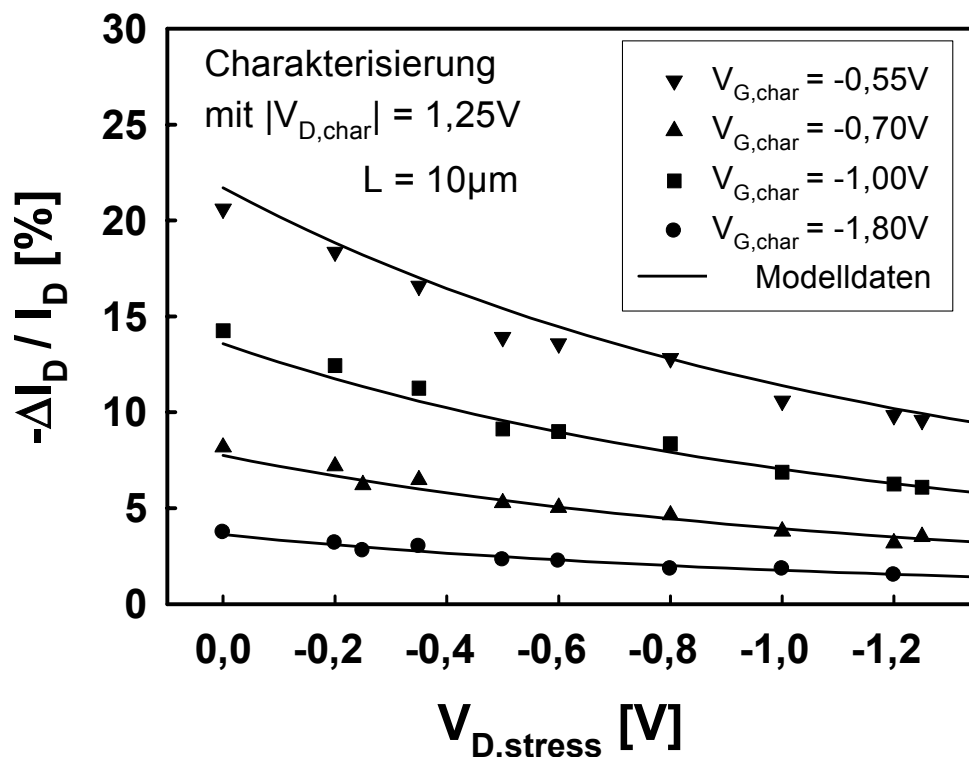


Abbildung 5.1: Degradation des Drainstromes als Funktion der Belastungs-Drainspannung für verschiedene Charakterisierungs-Gatespannungen. Die durchgezogenen Linien entsprechen den Daten der Modellvorhersage auf Basis der Gleichung (5.8). Belastung: $V_{G, stress} = -2,4V$, $V_{D, stress} = 0...-1,25V$, $T_{stress} = 125^{\circ}C$, $t_{stress} = 16h$. Device: $W/L = 20\mu m/10\mu m$, Technologie T10.

Für die rechnerisch ermittelten Daten wurden für die Parameter der Modellgleichung folgende Werte eingesetzt: $V_0 = 3,3\text{mV}$, $C = 1,4\text{V}^{-1}$ und $\theta = 1,0\text{V}^{-1}$. C und θ stellen dabei belastungsunabhängige Parameter dar. Sie hängen nicht von der Belastungsspannung, der Belastungsdauer oder der Temperatur ab. V_0 dagegen ändert sich in Abhängigkeit von den Belastungsparametern entsprechend der Gleichung (5.1).

Zwischen den modellierten Daten für die Drainstromdegradation und den tatsächlich ermittelten Daten kann eine sehr gute Übereinstimmung beobachtet werden.

In Abbildung 5.2 werden noch einmal dieselben Daten wie in Abbildung 5.1 aufgetragen, in diesem Fall jedoch normiert auf den Wert der Drainstromdegradation unter homogenen Belastungsbedingungen bei $V_{D,stress} = 0\text{V}$. Bei dieser Auftragungsart lässt sich der Einfluss der verschiedenen Gatespannungen bei der Charakterisierung gut erkennen.

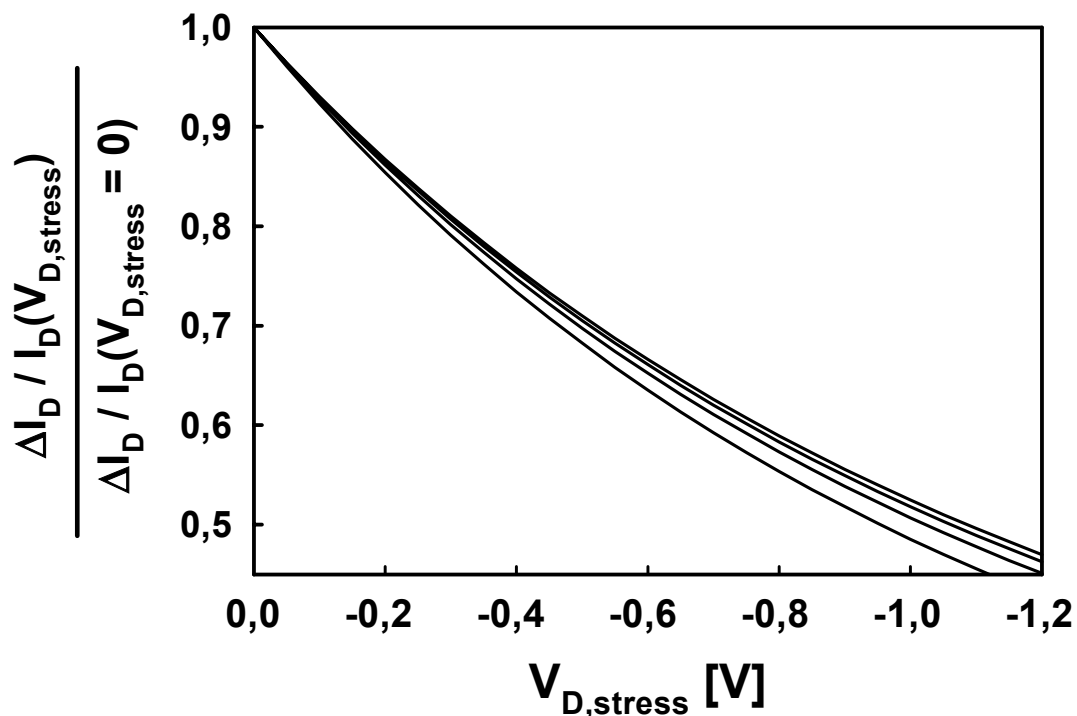


Abbildung 5.2: Dieselben Modelldaten für die Drainstromdegradation wie in Abbildung 5.1 jedoch normiert auf die erreichte Drainstromdegradation im homogenen Fall ($\Delta I_D / I_D(V_{D,stress} = 0) = 1$). Die unterschiedlichen Kurvenformen spiegeln die Abhängigkeit der Beweglichkeit von der Charakterisierungsgatespannung $V_{G,Char}$ wider.

Die jeweils leicht unterschiedlichen Kurvenformen entstehen aus der modellierten Abhängigkeit der Ladungsträgerbeweglichkeit von der Gatespannung unter Charakterisierungsbedingungen.

Abbildung 5.3 und Abbildung 5.4 zeigen einen weiteren Vergleich zwischen modellierten und gemessenen Daten. Die Werte werden wie in Abbildung 5.1 aufgetragen, für die belasteten Transistoren werden jedoch andere Transistorgeometrien gewählt, die typischerweise in Analog- oder Hochfrequenz-Applikationen eingesetzt werden. Für Abbildung 5.3 wurden Transistoren mit der Geometrie $W/L = 10\mu\text{m}/0,5\mu\text{m}$ belastet, für Abbildung 5.4 Transistoren mit einem Geometrieverhältnis von $W/L = 10\mu\text{m}/0,18\mu\text{m}$. Für die Modellparameter werden dieselben Werte wie für Abbildung 5.1 benutzt.

Auch für diese Transistorgeometrien kann eine gute Übereinstimmung zwischen den Modellvorhersagen und der tatsächlich gemessenen Drainstromdegradation beobachtet werden.

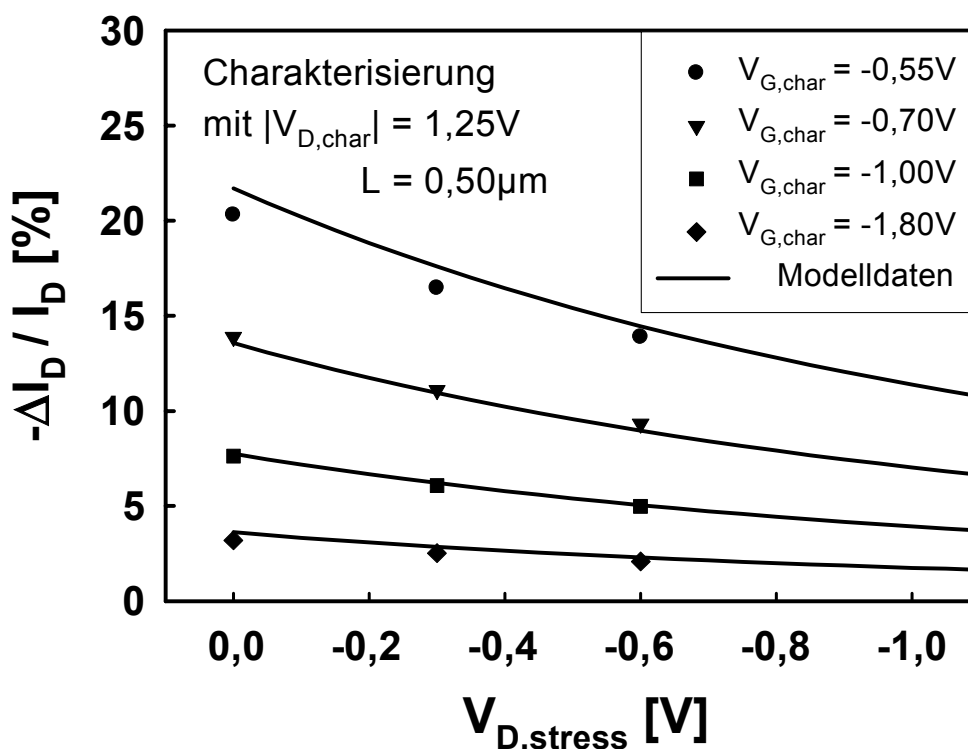


Abbildung 5.3: Gemessene und modellierte Drainstromdegradation als Funktion der Belastungs-Drainspannung für verschiedene Charakterisierungsbedingungen ähnlich Abbildung 5.1 jedoch mit einer für Analog-Applikationen typischen Transistorgeometrie. Belastung: $V_{G, \text{stress}} = -2,4\text{V}$, $V_{D, \text{stress}} = 0 \dots -0,6\text{V}$, $T_{\text{stress}} = 125^\circ\text{C}$, $t_{\text{stress}} = 16\text{h}$. Device: $W/L = 10\mu\text{m}/0,5\mu\text{m}$, Technologie T10.

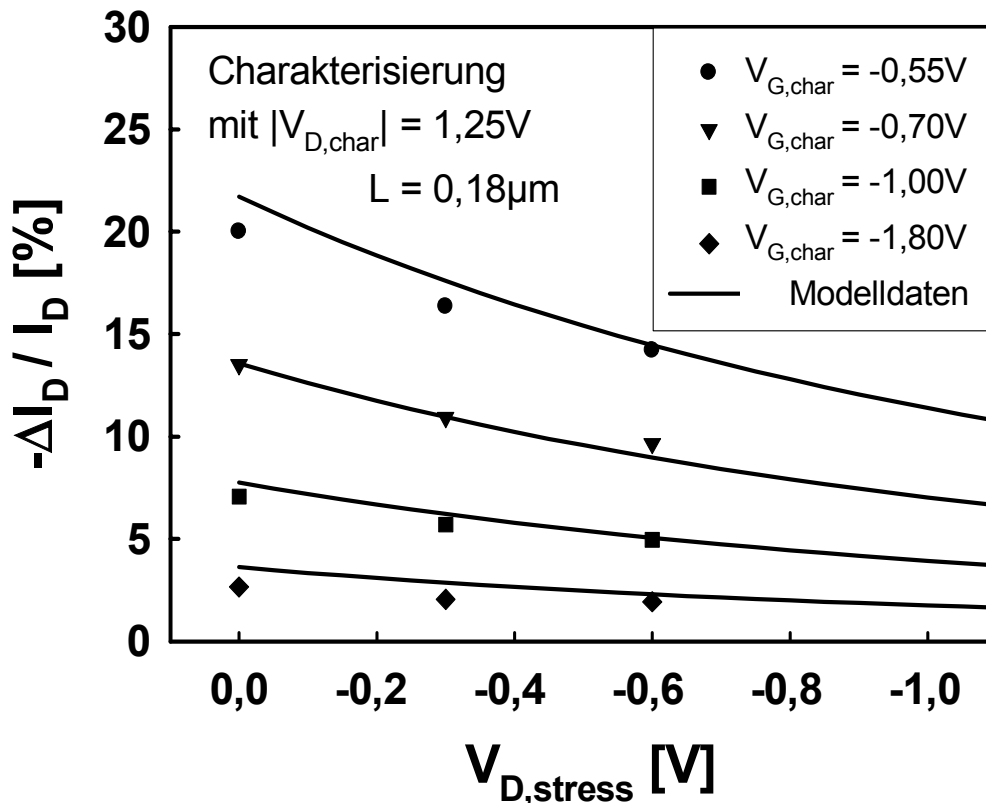


Abbildung 5.4: Gemessene und modellierte Drainstromdegradation als Funktion der Belastungs-Drainspannung für verschiedene Charakterisierungsbedingungen ähnlich Abbildung 5.1 jedoch mit minimaler Kanallänge. Belastung: $V_{G,stress} = -2,4V$, $V_{D,stress} = 0 \dots -0,6V$, $T_{stress} = 125^\circ C$, $t_{stress} = 16h$. Device: $W/L = 10\mu m/0,18\mu m$, Technologie T10.

Für die Verifikation der Modellgleichung an einer weiteren Technologiegeneration wurden Transistoren der Technologie T9 gewählt (siehe Anhang A1). Für die BT-Belastung wurde hier $t_{stress} = 16h$ und $V_{G,stress} = -3,3V$ gewählt. Die Belastungs-Drainspannung $V_{D,stress}$ wurde für die verschiedenen Experimente zwischen $0V$ und $-1,75V$ variiert.

Auch hier kann eine gute Vorhersagequalität der modellierten Daten im Vergleich zu den Messdaten festgestellt werden. Das Degradationsverhalten unter inhomogenen Belastungsbedingungen und die ausgearbeitete Modellierung können daher in guter Näherung als technologieunabhängig betrachtet werden.

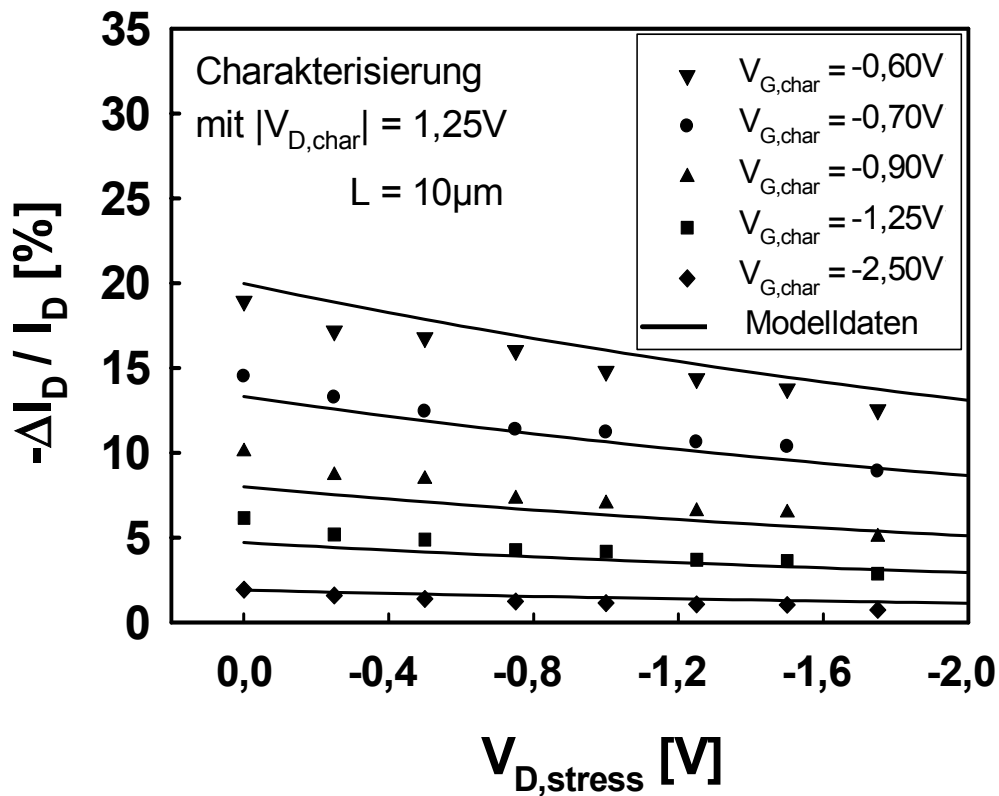


Abbildung 5.5: Gemessene und modellierte Drainstromdegradation als Funktion der Belastungs-Drainspannung für verschiedene Charakterisierungsbedingungen ähnlich Abbildung 5.1 jedoch für die Technologie T9.
Belastung: $V_{G,stress} = -3,3V$, $V_{D,stress} = 0 \dots -1,75V$, $T_{stress} = 125^\circ C$, $t_{stress} = 16h$.
Device: $W/L = 10\mu m/20\mu m$.

5.3 Vereinfachte Modellierung

Mit leichten Einbußen bei der Genauigkeit kann die Gleichung für die Drainstromdegradation deutlich vereinfacht werden. Wird die Abhängigkeit der Ladungsträgerbeweglichkeit von der anliegenden Gatespannung vernachlässigt, wird der Nenner innerhalb des Integrals auf der rechten Seite der Gleichung (5.7) gleich eins. Unter diesen Umständen verliert die Gleichung deutlich an Komplexität. Nach Integration und einigen Umformungen erhalten wir:

$$\Delta I_D / I_D = -2 \cdot \frac{V_0}{V_G - V_{th}} \cdot e^{(C \cdot V_{G, stress})} \cdot \frac{1 - e^{(-C \cdot V_{D, stress})}}{C \cdot V_{D, stress}} \quad (5.8)$$

Diese vereinfachte Modellierung stellt für viele Zwecke eine ausreichende Approximation für die Degradation des Stromes in Charakterisierungs-Arbeitspunkten mit geringen Gatespannungen dar, wie sie typisch für viele Analogapplikationen sind.

Wird die Drainstromdegradation für Arbeitspunkte mit höheren Gatespannungen mit dieser vereinfachten Gleichung (5.8) berechnet, weist die Modellierung Daten mit leicht erhöhten Abweichungen gegenüber gemessenen Daten auf. Unter Beibehaltung der Werte für die Parameter V_0 und C führt die Berechnung zu einer Überschätzung der Degradation, die gemessene Degradation fällt geringer aus. Die durch die höheren Charakterisierungs-Gatespannungen eintretende Beweglichkeitsreduktion der Ladungsträger wird nicht mehr berücksichtigt. Bei einer normierten Auftragung der Modellvorhersage der vereinfachten Gleichung (5.8) wie in Abbildung 5.2 liegen die Kurven übereinander. Die Krümmung der Kurven fällt identisch aus.

5.4 Richtlinien für die Lebensdauervorhersage

In Abbildung 5.6 und 5.7 sind Diagramme für die Vorhersage der Lebensdauer abgebildet (vergleiche Kapitel 2.6). Es werden Daten sowohl für ein $10\mu\text{m}$ langes Bauelement als auch für einen minimal langen Transistor bei homogener und inhomogener Belastung gezeigt. Für die Bestimmung der Lebensdauer werden verschiedene Kriterien eingesetzt. In Abbildung 5.6 wird die Degradation des Drainstroms als lebensdauerbegrenzendes Kriterium herangezogen. Bei einem typischen Wert von 10% für die erreichte Drainstromdegradation wird hier die Lebensdauer als erreicht definiert.

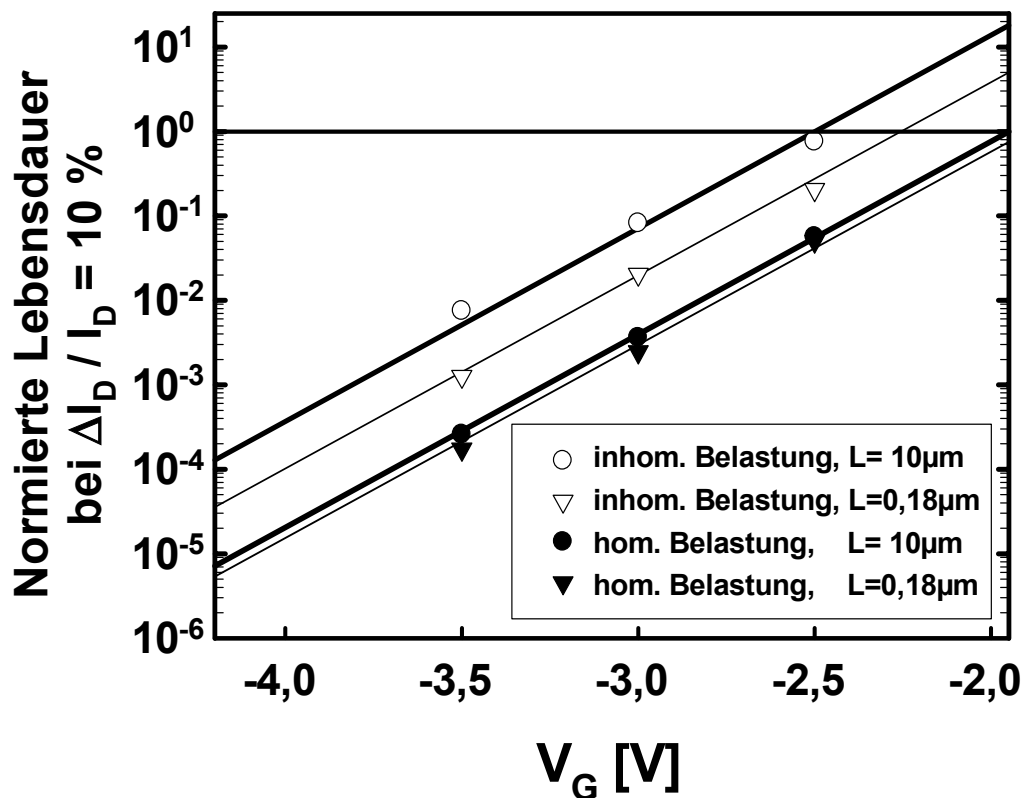


Abbildung 5.6: Lebensdauerextrapolation homogen und inhomogen belasteter Transistoren unterschiedlicher Kanallängen auf die bei normalen Betriebsbedingungen anliegende Gatespannung ($V_{DD} + 5\% = 1,9\text{V}$). Bei einer belastungsinduzierten Drainstromdegradation von 10% wird die Lebensdauer als erreicht definiert. Die Lebensdauern werden normiert dargestellt bezogen auf die Lebensdauer des Langkanaltransistors unter homogenen Belastungsbedingungen. Der Drainstrom wird im Arbeitspunkt $V_{G,char} = V_{D,char} = -1,8\text{V}$ aufgenommen. Belastung: $V_{D,stress} = 0\text{V}$ (homogen) und $-1,2\text{V}$ (inhomogen), $T_{stress} = 125^\circ\text{C}$. Devices: $W/L = 20\mu\text{m}/10\mu\text{m}$ und $10\mu\text{m}/0,18\mu\text{m}$, Technologie T10.

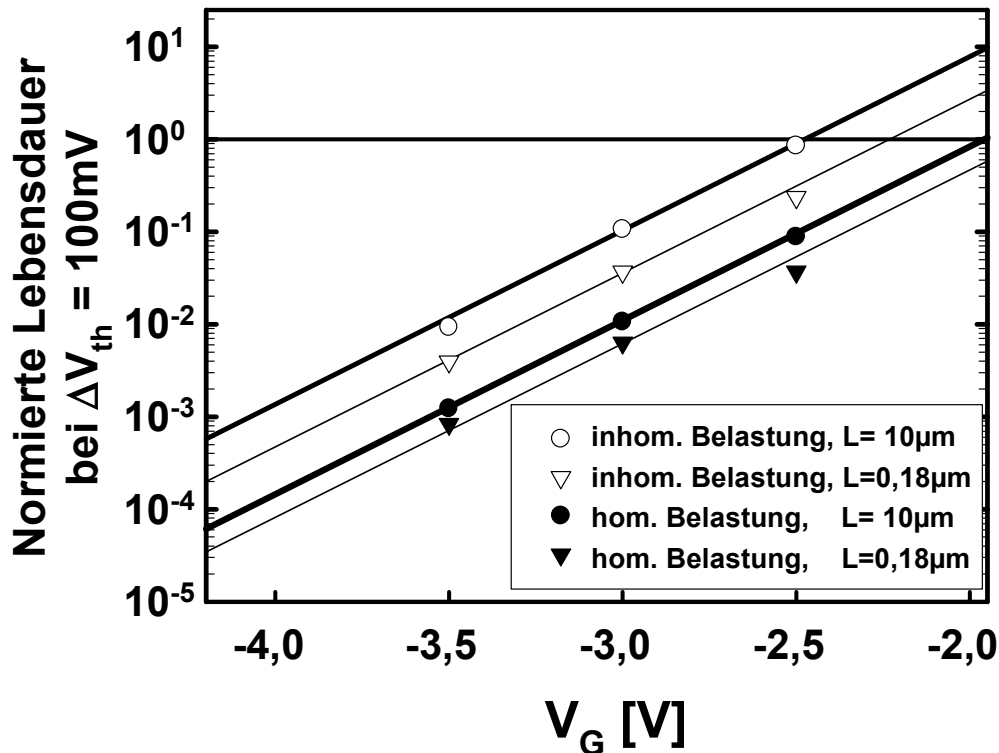


Abbildung 5.7: Lebensdauerextrapolation homogener und inhomogener belasteter Transistoren unterschiedlicher Kanallängen auf die bei normalen Betriebsbedingungen anliegende Gatespannung ($V_{DD} + 5\% = 1,9\text{V}$). Bei einer belastungsinduzierten Einsatzspannungsverschiebung von 100mV wird die Lebensdauer als erreicht definiert. Die Lebensdauern werden normiert dargestellt bezogen auf die Lebensdauer des Langkanaltransistors unter homogenen Belastungsbedingungen. Die Einsatzspannung wird mithilfe des Zielstromkriteriums ermittelt ($V_{th} = V_G @ 70\text{nA pro } W/L$). Belastung: $V_{D,stress} = 0\text{V}$ (homogen) und $-1,2\text{V}$ (inhomogen), $T_{stress} = 125^\circ\text{C}$. Devices: $W/L = 20\mu\text{m}/10\mu\text{m}$ und $10\mu\text{m}/0,18\mu\text{m}$, Technologie T10.

Darüber hinaus wird in Abbildung 5.7 die Verschiebung der Schwellenspannung des Transistors als Kriterium eingesetzt. Die belastungsinduzierte Verschiebung der Einsatzspannung ist für homogene BT-Belastungen ein gebräuchlicher Parameter für die Ermittlung der Lebensdauer und soll von daher neben dem Drainstromkriterium hier ebenfalls herangezogen werden.

Die Punkte stehen für die gemessenen Daten, die durchgehenden Linien ergeben die Extrapolation der Lebensdauer auf die Betriebsspannung für das gegebene Lebensdauerkriterium. Alle Werte sind normiert bezogen auf die Lebensdauer des Langkanaltransistors unter homogenen Belastungsbedingungen.

Die bei homogenen Belastungsbedingungen erreichte Lebensdauer dieses Transistors bei der maximal geforderten nominellen Betriebsspannung plus 5% Sicherheitsreserve (1.9 V) entspricht hier dem Wert eins.

Für beide Lebensdauerkriterien und für beide Kanallängen können wir einen mehr als achtfachen Zuwachs an Lebensdauer für den inhomogenen Fall beobachten. Die Lebensdauer des minimal langen Transistors ist im Vergleich mit dem langen Transistor für homogene und inhomogene Belastung jeweils geringer. Eine Erklärung dafür kann in degradationsverstärkenden Effekten im Übergangsbereich des Kanalbereiches zu den Diffusionsgebieten und unterhalb des ‚Spacer-Oxids‘ liegen.

Aus diesen Ergebnissen folgen zwei Konsequenzen für die Praxis der Lebensdauer vorhersage für Schaltungen:

1. Für Standard-Applikationen sind die Anforderungen an die Zuverlässigkeit der Transistoren bezüglich inhomogener BT-Belastung automatisch abgedeckt, solange die Zuverlässigkeit bei homogener BT-Belastung gewährleistet ist.
2. In speziellen Schaltungsapplikation, in denen inhomogene BT-Belastungszustände unter erhöhten Versorgungsspannungen auftreten oder verschärfte Anforderungen eine Reduktion der Bereiche der erlaubten Parameterdegradation erfordern, führt der Lebensdauererfolg unter inhomogener Belastung zu einer Entspannung in Vergleich zu den aus homogenen Belastungsexperimenten gewonnenen Lebensdauer vorhersagen.

6 Schaltungssimulation

Möchte man die Auswirkungen von Alterungsphänomenen von MOS-Feldeffekttransistoren auf die Funktion einer Schaltung oder eines vollständigen Produktes präziser als mit dem in Kapitel 2.6.1 beschriebenen ‘Worst-Case’ und ‘Duty-Cycle’ Ansatz bewerten, darf man sich nicht auf die isolierte Betrachtung eines einzelnen Bauelementes beschränken. Im Prinzip muss jeder eingesetzte Transistor in seinem Schaltungsverbund betrachtet werden.

Diese Sichtweise ist sowohl für die exakte Bestimmung der Belastungsbedingungen der einzelnen Transistoren notwendig, als auch für die korrekte Erkennung und Bewertung der Auswirkungen der Parameterdegradation des Bauelementes auf die Funktion der gesamten Schaltung. Diese beiden Funktionen werden idealerweise von einem Schaltungssimulator übernommen [Schlünder05_2].

In diesem Kapitel soll auf die Möglichkeiten und die Arbeitsweise eines so genannten ‘Aging-Simulators’ eingegangen werden. Im Rahmen einer Zusammenarbeit mit der Firma ‘Cadence’ wurde hierfür speziell das Software-Tool ‘RelExpert’ eingesetzt und zusammen mit dem Anbieter an die Technologie T11 (Anhang A1) angepasst. Anhand eines ausgewählten Schaltkreises werden die verschiedenen Schritte der Simulation einer belasteten Schaltung erläutert.

6.1 Methodik der Simulation geschädigter Schaltungen

Bislang beschränken sich die meisten der eingesetzten Simulatoren auf die Simulation von Schaltungen mit konstanten Transistorparametern. Eine belastungsinduzierte Verschiebung der elektrischen Parameter der Bauelemente und somit eine Veränderung des Verhaltens der gesamten Schaltung wird bei diesem Ansatz nicht berücksichtigt. Soll ein Simulations-Softwarepaket auch diese Funktionalität zur Verfügung stellen, müssen zusätzliche Bewertungs- und Kalkulationsschritte in den Simulationsablauf implementiert werden.

6.2 Simulationsablauf für eine belastete Schaltung

Die notwendigen Schritte für eine derartige erweiterte Simulation lassen sich in vier Punkte gliedern:

- 1) Angleichung der Transistor-Modellparameter und Degradations-Modellparameter an die betrachtete Halbleitertechnologie;
- 2) Ermittlung der auftretenden Belastungen jedes einzelnen Transistors der zu simulierenden Schaltung bei Beschaltung mit Eingangssignalmustern, so genannten ‚Stimuli‘;
- 3) Berechnung der degradationsbedingt zu erwartenden Änderungen der Transistormodellparameter für jedes Bauelement;
- 4) Neusimulation des Schaltungsverhaltens nach einer vorgegebenen Belastungszeit mithilfe der neu ermittelten Modellparameter.

Auf diese Komponenten soll nun in den folgenden Unterkapiteln genauer eingegangen werden.

6.2.1 Bestimmung der in der Modellgleichung verwendeten Parameter

Für die Simulation von Schaltungen muss ein Simulator auf jede betrachtete Technologie abgeglichen werden. Für die verschiedenen Transistortypen steht dem Simulator eine Liste mit deren Modellparametern (‚model-card‘) zur Verfügung, die das elektrische Verhalten der Transistoren dieser Technologie beschreibt und daher für jede Technologie angepasst werden muss. Mithilfe von zahlreichen Messungen werden die dafür notwendigen elektrischen Parameter einer CMOS-Technologie extrahiert und die Modellparameter des Simulators angepasst. Die ermittelten Modellparameter beschreiben das Transistorverhalten im unbelasteten Zustand, der Satz dieser Parameter wird dementsprechend auch ‚fresh-model-card‘ genannt.

Zusätzlich zu der Bestimmung der in den Modellgleichungen verwendeten Parameter für unbelastete Transistoren werden für die Simulation von belasteten Schaltungen weitere Parameter für Degradationsmodelle benötigt. Mithilfe dieser

zusätzlichen Parameter kann das elektrische Verhalten der Transistoren und damit einer untersuchten Schaltung nach einer simulierten Belastung berechnet werden. Die Degradation der Transistoren während des Betriebes einer Schaltung zeigt eine deutliche Abhängigkeit von der eingesetzten Technologie. Unterschiedliche Prozessschritte innerhalb des Herstellungsverfahrens zeigen einen starken Einfluss auf das Degradationsverhalten der Transistoren (siehe dazu auch Kapitel 7.1.1). Der für die ‚fresh-model-card‘ beschriebene Abgleichprozess der Modellparameter muss von daher für jede Technologie, für die der Simulator eingesetzt werden soll, auch für die Degradationsmodelle neu erfolgen. Dafür wird das elektrische Verhalten der Transistoren der zu simulierenden Technologie nicht nur wie bereits beschrieben durch Messungen an unbelasteten Transistoren untersucht, sondern auch nachdem die Transistoren bestimmten Belastungen ausgesetzt wurden. Hierfür werden Experimente mit vielen unterschiedlichen Belastungsbedingungen durchgeführt und die Transistoren jeweils vor und nach Ende der Belastungen charakterisiert. Die Degradationsmodellparameter des Simulators werden mithilfe der extrahierten Messwerte solange angepasst, bis die Simulationsaussagen möglichst genau dem realen elektrischen Verhalten der Transistoren nach Belastungen entsprechen.

6.2.2 Bestimmung der individuellen Belastungen der Transistoren

Für die Simulation eines Schaltkreises nach einer vorgegebenen Betriebszeit muss an erster Stelle eine genaue Ermittlung der auftretenden Belastungen aller in dem untersuchten Schaltungsteil eingesetzten Transistoren stehen. Die Belastung der Transistoren einer Schaltung fällt im Betrieb unterschiedlich stark aus. Dementsprechend verläuft die Degradation der einzelnen Transistoren ebenfalls unterschiedlich stark. Die durchlaufenen Arbeitspunkte, die Dauer, wie lange sich die Transistoren in diesen Arbeitspunkten befinden, und die vorherrschende Temperatur bestimmen die Stärke der Belastung und damit den Wert der daraus resultierenden Parameterverschiebung.

Der Simulator berücksichtigt das unterschiedliche Degradationsverhalten der Transistoren der belasteten Schaltung mithilfe eines für jeden Transistor ermittelten Parameters, des so genannten ‚Age-Wertes‘. Dieser Wert stellt ein Maß für die Stärke der während der Belastung erlittenen Parameterdegradation jedes einzelnen Transistors dar. Im Folgenden soll der Ablauf der Simulation einer belasteten Schaltung anhand einer Beispielschaltung beschrieben werden. Die mögliche

Anzahl der Transistoren innerhalb eines zu simulierenden Schaltungsteils ist typischerweise durch die Rechenleistung und die damit verbundene Wartezeit begrenzt.

In Abbildung 6.1 ist eine Wortleitungs-Treiberschaltung („segmented word-line driver“) abgebildet, wie sie oft innerhalb moderner Speicherschaltungen mit hierarchischer Wortleitungsarchitektur eingesetzt wird [Kiriata98, Wuensche02]. Für ein korrektes Zeitverhalten innerhalb des Speichermoduls, in dem dieser Schaltungsblock eingesetzt wird, darf sich das Ausgangssignal (WLDV) des Treibers im Laufe der Betriebsdauer der Schaltung belastungsbedingt nicht zu stark verändern. Eine zu große Parameterdegradation der Transistoren in diesem Schaltungsblock kann daher die Funktion des gesamten Speichermoduls beeinträchtigen und somit die Lebensdauer limitieren.

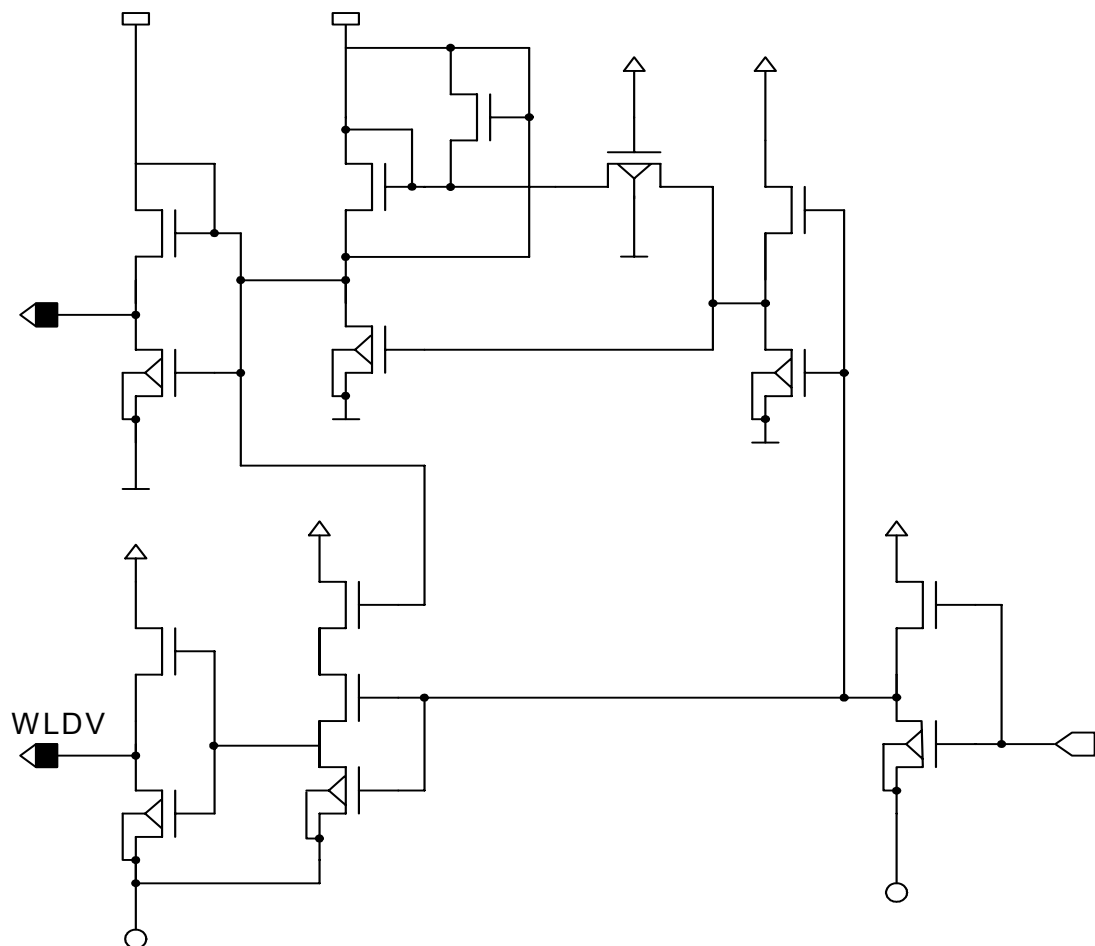


Abbildung 6.1: Typische Wortleitungs-Treiberstufe einer hierarchischen Wortleitungs-Treiberarchitektur für Speicherschaltungen („segmented wordline driver“). Zu starke Parameterdegradationen der Transistoren dieser Schaltung können zu Fehlern im Zeitverhalten der Speicherschaltung führen.

Die auftretende Belastung wird für jeden Transistor der Schaltung mithilfe der vorgegebenen Eingangssignalmuster („stimuli“) integrierend erfasst und mithilfe des Age-Wertes gespeichert. Befinden sich einzelne Transistoren der Schaltungen nur selten in belastungskritischen Arbeitspunkten, wird sich bei diesen Transistoren im Laufe der Betriebsdauer nur eine sehr geringe Verschiebung der elektrischen Parameter einstellen. Für diese Transistoren werden geringe Age-Werte erfasst. Für andere Transistoren, die lange in belastungskritischen Arbeitspunkten betrieben werden und somit stärkere Degradation erleiden, werden größere Age-Werte registriert. Insbesondere derartige uneinheitliche Belastungen und daraus resultierende Parameterverschiebungen können zu Beeinträchtigungen der Schaltungsfunktion führen (vergleiche Kapitel 2.5 und [Schlünder05_2]).

6.2.3 Berechnung der zu erwartenden Änderung der Modellparameter

Der Anwender kann nun für die Simulation der belasteten Schaltung neben den beschriebenen Eingangssignalmustern und der Betriebsspannung eine gewünschte Betriebszeit und eine Temperatur vorgeben. Die Belastungszeit entspricht dabei einem Vielfachen des durch das eingesetzte Eingangssignalmuster gebildeten Zeitintervalls. Mithilfe dieser Parameter und der ermittelten Age-Werte aus dem zuvor beschriebenen ersten Schritt berechnet die Simulationssoftware nun individuell für jedes Bauelement die auftretenden Belastungen. Im zweiten Schritt wird dann für die erwartete Degradation die Modellgleichung jedes Transistors angepasst. Auf Basis der hinterlegten Degradationsmodelle werden belastungsinduziert degradierte Transistor-Modellparameter neu berechnet und für jeden eingesetzten Transistor gespeichert.

Nach Abschluss der Berechnungen steht eine neue, zweite Liste mit Modellparametern für die Transistoren der Schaltung zur Verfügung, die so genannte ‚aged model-card‘. Diese Liste enthält dieselben Parameter wie die fresh-model-card, die dem Simulator vor der simulierten Belastung zur Verfügung steht, die Werte für die Parameter beschreiben nun jedoch das elektrische Verhalten der Transistoren nach der simulierten belastungsbedingten Degradation. Die aged-model-card beschreibt im Gegensatz zur fresh-model-card nicht nur jeden verschiedenen Transistortyp der Schaltung, sondern die Modellparameter werden für jeden einzelnen Transistor der Schaltung individuell aufgeführt.

6.2.4 Neusimulation des Schaltungsverhaltens

Nach der Berechnung der belastungsbedingt veränderten Modellparameter für jedes Device kann die Schaltung neu simuliert werden. Die Simulation beschreibt nun das elektrische Verhalten der Schaltung nach der vorgegebenen Betriebsdauer. Die Auswirkung der Parameterverschiebung der einzelnen Bauelemente auf die Funktion der Schaltung kann betrachtet und bewertet werden. Das Verhalten der simulierten Schaltung kann auf diesem Wege im unbelasteten Zustand mit dem nach einer festgelegten Betriebszeit verglichen werden. Für die Simulation des Verhaltens der unbelasteten Schaltung wird die fresh-model-card herangezogen, für die Simulation der belasteten Schaltung die aged-model-card.

Abbildung 6.2 zeigt den berechneten Verlauf der Ausgangsspannung WLDV des Wortleitungstreibers mit dem eingesetzten Eingangssignalmuster vor und nach der gewählten Betriebsdauer. Abbildung 6.3 zeigt die erste abfallende Flanke des Ausgangssignals in vergrößerter Darstellung.

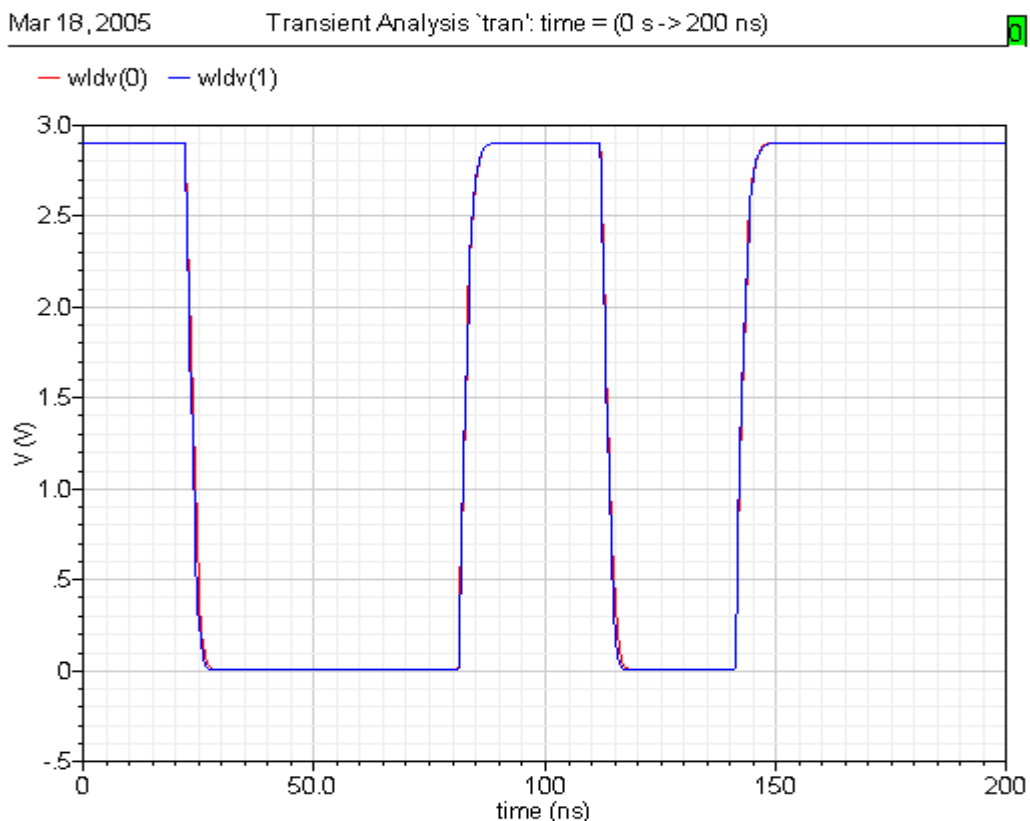


Abbildung 6.2: Simulierte Ausgangsspannung WLDV des Wortleitungstreibers vor und nach einer simulierten Betriebsdauer.

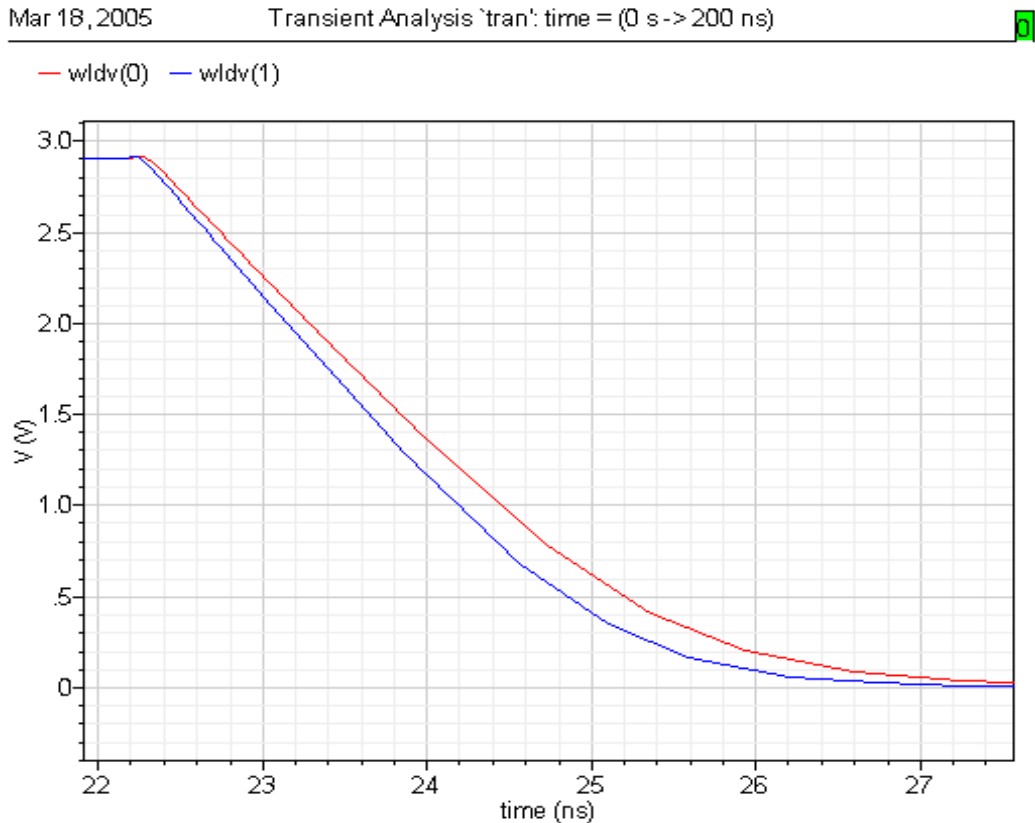


Abbildung 6.3: Vergrößerter Ausschnitt der ersten fallenden Flanke der Ausgangsspannung aus Abbildung 6.2. Der unterschiedliche zeitliche Verlauf ergibt sich aus der Parameterdegradation der Transistoren innerhalb der Treiberschaltung.

Diese Möglichkeit der Analyse bietet zahlreiche Vorteile für die Schaltungsentwicklung. Die Lebensdauer der Schaltung kann so präziser ermittelt werden, da die Auswirkungen der Parameterdegradation der Transistoren auf die Funktion der Schaltung direkt untersucht werden können. So genannte Margins für den sicheren Betrieb einer Schaltung (Kapitel 2.5) können auf das notwendige Maß reduziert werden und die maximal mögliche Performance der Halbleitertechnologie kann besser genutzt werden. Darüber hinaus können mithilfe des Simulators belastungskritische Transistoren (so genannte ‚Hot Spots‘) identifiziert werden, deren mögliche Degradation die Funktion der Schaltung stark beeinträchtigen. Schaltungstechnische Maßnahmen mit dem Ziel die zuverlässige Funktion der Schaltung zu sichern („Reliability by Design“), können auf diese Schaltungsteile konzentriert werden. Der Aufwand für das Design dieser Maßnahmen und der eventuell zusätzliche Flächenbedarf kann so in Grenzen gehalten werden [Schlünder05_2]. Im folgenden Kapitel soll neben prozesstechnischen Maßnahmen auch auf derartige Schaltungsmaßnahmen zur Erhöhung der Zuverlässigkeit eingegangen werden.

7 Maßnahmen zur Verminderung der Degradation

Um die mögliche Betriebslebensdauer einer Schaltung zu verlängern, können unterschiedliche Maßnahmen ergriffen werden [Schlünder05_2]. In diesem Kapitel sollen dafür mehrere Wege beschrieben werden. Es wird hierbei zwischen zwei unterschiedlichen Ansätzen unterschieden. Zum einen werden prozesstechnische Maßnahmen erläutert, die die BTS-Festigkeit der Transistoren verbessern. Zum anderen wird mit schaltungstechnischen Maßnahmen ein gänzlich anderer Ansatz verfolgt. Hier sollen Belastungssituationen vermieden oder deutlich reduziert werden, die zur Schädigung der Transistoren und somit zur Verschiebung der elektrischen Parameter führen.

7.1 Prozesstechnische Maßnahmen

Eine Grundlage für eine Optimierung der BTS-Festigkeit einer Halbleitertechnologie liegt im Verständnis dafür, welche einzelnen Prozessschritte innerhalb des Herstellungsverfahrens einen starken Einfluss auf das Degradationsverhalten zeigen. Aus den physikalischen Grundlagen der bei BTS einsetzenden Schädigungsmechanismen und aus Vergleichsuntersuchungen an Prozessvariationen lassen sich Einflussfaktoren ableiten und identifizieren, die das Degradationsverhalten der Transistoren einer Technologie mitbestimmen. Darauf aufbauend lassen sich gezielt Veränderungen bei der Prozessierung einführen, um die Lebensdauer der Transistoren und Schaltungen zu verbessern.

7.1.1 Einfluss verschiedener Prozessierungsschritte

Während der Belastungssituation werden Wasserstoffbindungen an der Grenzfläche zwischen Silizium und Gateoxid aufgebrochen und der Wasserstoff diffundiert von der Bruchstelle weg (Kapitel 2). Die Beschaffenheit der Wasserstoffbindungen, das Diffusionsverhalten des Wasserstoffes im Oxid und letztendlich auch die Wahrscheinlichkeit einer erneuten Bindung mit einem elektrisch neutralen Element zeigen von daher einen großen Einfluss auf das Degradationsverhalten bei BTS. Dementsprechend beeinflussen alle Prozessschritte, die den Wasserstoffgehalt der Gateoxide mitbestimmen, mehr oder weniger deutlich auch das Degradationsverhalten des Oxids bei BT-Belastungen. Darüber hinaus entscheidet ebenso das thermische Budget aller dem Aufwachsen des Gateoxides folgenden Prozessschritte über die Stabilität gegenüber BTS. Sehr deutlich wird dies bei Variation verschiedener Temperatur-Ausheilschritte, die häufig in einer mit Wasserstoff angereicherten Atmosphäre stattfinden.

Beachtet werden muss in diesem Zusammenhang die sehr stark ausgeprägte Eigenschaft der Diffusion des Wasserstoffs. Die Art und Weise der Passivierung am Ende eines Halbleiterprozesses zeigt z.B. erfahrungsgemäß einen sehr ausgeprägten Einfluss auf die Zuverlässigkeit der Gateoxide. Obwohl bei modernen Halbleitertechnologien eine Vielzahl von Metalllagen die Passivierungsprozessschritte räumlich von den Gateoxiden trennen, lässt sich ein Einfluss auf das Degradationsverhalten bei BTS deutlich nachweisen. Speziell der Einsatz von Polyimid bei der Passivierung, das häufig für eine Optimierung der Gehäuseeinbau-Prozesse benutzt wird, zeigt einen negativen Einfluss auf die BTS-Festigkeit. Empirische Untersuchungen zeigen, dass derartige Polyimidschichten das Gateoxid durch ihre großen Vorräte an Wasserstoff, der durch Diffusion bis hinunter zum Gateoxid gelangen kann, stark beeinflussen können.

Wird für die Metallisierung der Halbleitertechnologie Kupfer eingesetzt, werden solche Einflüsse abgeschwächt. Die innerhalb der Metallisierungsebenen notwendigen Diffusionsbarrieren für die Kupferleitungen reduzieren den Einfluss der Passivierungsschritte auf die Qualität des Gateoxids. In den durch die Barrieren abgegrenzten Bereichen bleibt der Wasserstoffvorrat weitestgehend konstant. Prozessschritte oberhalb der Diffusionsbarrieren zum Gateoxid zeigen daher geringere Einflüsse auf das Oxidmaterial, als bei Technologien mit Prozessführung ohne Diffusionsbarrieren.

7.1.2 Verwendung von Deuterium als Wasserstoffersatz

Die große Bedeutung des Wasserstoffs für das Degradationsverhalten des Gateoxids in Belastungssituationen führte nach dem Anstieg von Zuverlässigkeitsproblematiken bei Halbleitertechnologien zu ausführlichen Versuchen mit Wasserstoffisotopen. Speziell für die Verbesserung des Degradationsverhaltens bei Belastung mit heißen Ladungsträgern wurde mit dem Wasserstoffisotop Deuterium experimentiert. Werden die der Prozessierung des Gateoxids folgenden Ausheilschritte in einer mit Deuterium angereicherten Atmosphäre durchgeführt, führt dies zu einer deutlichen Reduktion des Degradationsprozesses bei Belastung mit heißen Ladungsträgern und somit zu erhöhten Lebensdauern [Kizilyalli97, Clark01].

Dieser positive Einfluss lässt sich auch bei BT-Belastungen beobachten. Abbildung 7.1 zeigt eine Vergleichsuntersuchung an einer 0,25 μm -Technologie mit und ohne Deuterium Anneal-Schritten (Technologie T9 und T9D2).

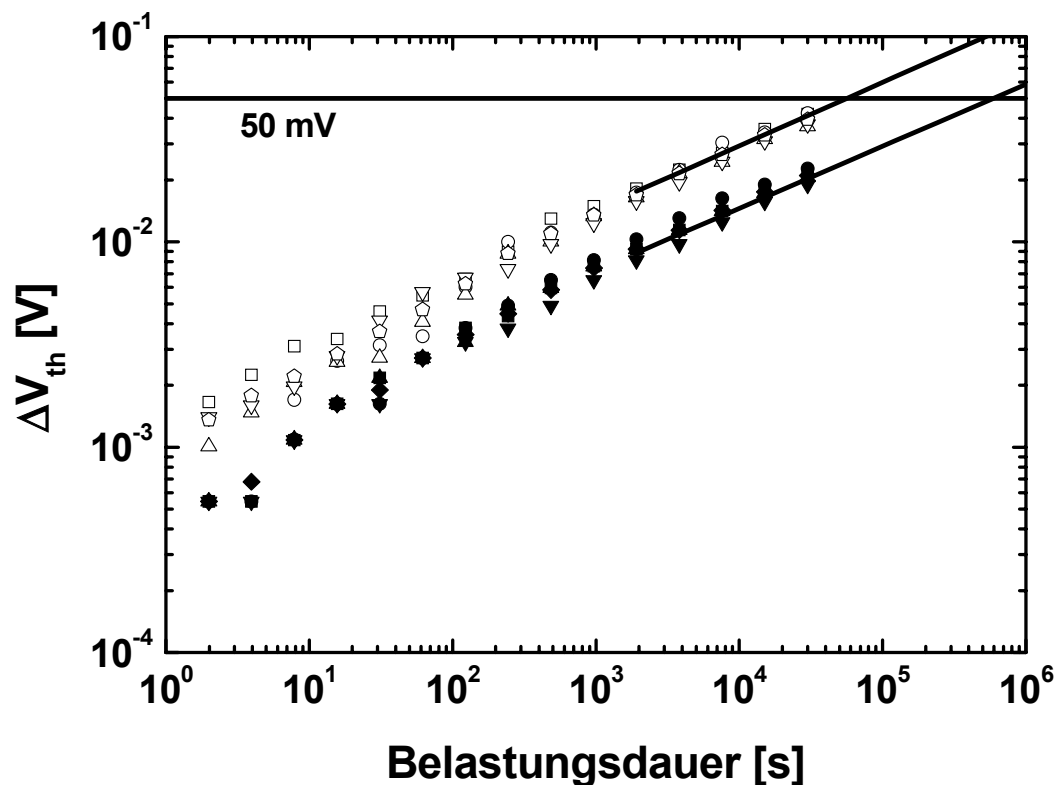


Abbildung 7.1: Vergleich der belastungsinduzierten Verschiebung der Einsatzspannung einer Technologie mit (ausgefüllte Symbole) und ohne Deuterium-Ausheilschritte (offene Symbole). Die Einsatzspannung wird mithilfe des Stromkriteriums ermittelt. Belastung: $V_{G, stress} = -4,0\text{V}$, $T_{stress} = 125^\circ\text{C}$. Devices: $W/L = 10/0,25\mu\text{m}$, Technologie T9 und T9D2.

P-Kanal Transistoren beider Technologievariationen werden einer identischen BT-Belastung ausgesetzt. An die Transistoren wird 16 Stunden lang bei $T_{stress} = 125^{\circ}\text{C}$ eine Gatespannung von $V_{G, stress} = -4,0\text{V}$ angelegt. Die Einsatzspannung wurde nach dem Zielstromkriterium (siehe Kapitel 3.3.2.2) vor und nach bestimmten Belastungsintervallen aufgenommen. Im Fall der Transistoren mit Deuterium-Ausheilschritten kann eine deutliche Abschwächung der Parameterdegradation beobachtet werden.

Die geringere Degradation der Einsatzspannung lässt sich anhand des zugrunde liegenden Schädigungsmechanismus erklären (siehe Kapitel 2.2). Viele der im Übergangsbereich zwischen Silizium und Gateoxid zunächst offenen Bindungen („Dangling Bonds“) werden während der Deuterium-Ausheilschritte mit Deuterium anstatt mit Wasserstoff abgesättigt. Unter BT-Belastung werden diese Bindungen teilweise wieder aufgebrochen. Der Vorteil der Deuterium-Bindung liegt hier jedoch nicht in einer höheren Bindungsenergie. Die zum Aufbrechen der Bindung notwendige Energie ist weitestgehend identisch und erklärt daher nicht das abgeschwächte Schädigungsverhalten. Das Isotop ist aber im Vergleich zu Wasserstoff größer und zeigt eine geringere Diffusionsgeschwindigkeit. Die Wahrscheinlichkeit, dass eine aufgebrochene Bindung wieder mit Deuterium abgesättigt wird, ist von daher höher. Der bei den Bindungsbrüchen frei werdende Wasserstoff steht dagegen sehr schnell nicht mehr für eine neue Bindung zur Verfügung.

Bei Untersuchungen der verschiedenen Prozessvarianten können weitere wichtige Ergebnisse über den Einsatz von Deuterium gewonnen werden. Dafür wird zusätzlich zur Variation des Prozessgases auch der Zeitpunkt der Deuterium-Ausheilschritte innerhalb des Herstellungsprozesses variiert. Folgen dem Deuterium-Ausheilschritt zu viele weitere thermische Prozessschritte, wird die Wirksamkeit stark eingeschränkt. Werden die Deuterium-Behandlungen nur unmittelbar nach der Prozessierung des Gateoxids eingesetzt und nicht zusätzlich später im Metallisierungsprozess wiederholt, bleibt die lebensdauerverlängernde Wirkung auf das Degradationsverhalten bei BTS aus. Für diese Untersuchungen werden die Scheiben nach unterschiedlichen Prozessschritten aus der Fertigungslinie genommen. An Wafer-Material mit vier Metalllagen und Deuterium-Ausheilschritten nur direkt nach dem Gateoxidprozess und nicht mehr zusätzlich nach Prozessschritten für die Metallisierung ist der positive Einfluss des Deuteriums nicht mehr nachzuweisen. Der nach der Prozessierung der ersten Metallebene noch deutlich messbare Lebensdauererfolg bleibt aus.

Neben diesen Schwierigkeiten führen aber vor allem die mit dem Deuteriumeinsatz verbundenen hohen Kosten dazu, dass der Einsatz des Wasserstoffisotops sich bislang nicht bei der produktiven Herstellung von Halbleitertechnologien durchsetzen konnte. Solange die geforderten Lebensdauern für Halbleiterprodukte auf anderen Wegen sichergestellt werden können, wird auch weiterhin auf Deuterium verzichtet werden.

Bei der fortschreitenden Verschärfung der BTS-Problematik bei zukünftigen Technologien (Kapitel 8 und [IRTS04]) stellt der Einsatz von Deuterium jedoch eine Alternative dar. Im Rahmen der Entwicklung modernster CMOS-Technologien wird der Einsatz von Deuterium erneut in Erwägung gezogen und es werden Untersuchungen dazu durchgeführt.

7.2 Schaltungstechnische Maßnahmen

Bedingt durch die in der Vergangenheit bei jeder neuen Technologiegeneration steigenden Oxidfeldstärken (vergleiche Kapitel 4.1) und durch die Verringerung des Verhältnisses aus Transistorschwellenspannung zu Betriebsspannung stellt BTS für moderne CMOS-Technologien ein ernstes Zuverlässigkeitsproblem dar [Krishnan05, Schlünder05_2]. Es muss prozesstechnisch ein zunehmend höherer Aufwand betrieben werden, um bei gleichzeitiger Performancesteigerung die Parameterdegradation der Transistoren in Grenzen zu halten.

Die Beherrschung der Zuverlässigkeit kann daher immer weniger als eine alleinige Aufgabe für die Technologieentwicklung verstanden werden. Das Risiko eines degradationsbedingten Ausfalls der Schaltung kann auch durch verschiedene Maßnahmen im Schaltungsdesign reduziert werden (,Reliability by Design'). Dafür werden einzelne, kritische Transistoren oder Schaltungsblöcke betrachtet und eventuell abgeändert, die bei belastungsbedingter Degradation die Lebensdauer der gesamten Schaltung limitieren können. Diese ,Hot-Spots' können z.B. mit einem Aging-Simulator identifiziert werden (siehe Kapitel 6),

Um diese kritischen Blöcke vor den Auswirkungen von BT-Belastung zu schützen, können schaltungstechnische Maßnahmen ergriffen werden [Schlünder05_2, Thewes00]. Im folgenden Unterkapitel soll eine derartige Lösung aufgezeigt werden.

7.2.1 Wahl des Transistorarbeitspunktes in belastungsintensiven Betriebsmoden

In Abbildung 7.2 wird schematisch ein Beispiel für schaltungstechnische Maßnahmen aufgezeigt, mit deren Hilfe höhere Zuverlässigkeit erreicht werden kann („Reliability by Design“). Die gezeigte Erweiterung der Differenzstufe des in Kapitel 2.1 besprochenen Operationsverstärkers ermöglicht die Vermeidung von BT-Belastungen der kritischen Eingangstransistoren im Energiesparmodus („Power Down Modus“) [Thewes00].

Die Gateanschlüsse der Eingangstransistoren werden während der Power-Down-Phasen über die zusätzlichen Transistoren M14 und M15 mit der Versorgungsspannung V_{DD} verbunden. Auf diese Weise werden die sonst unter Power-Down-Bedingungen möglichen hohen elektrischen Felder über den Gateoxiden der Eingangstransistoren M4 und M5 und somit eine mögliche BT-belastungsbedingte Parameterdegradation dieser Bauelemente verhindert. Im aktiven Betrieb der Schaltung werden die zusätzlichen Transistoren gesperrt und beeinflussen die Funktion der Schaltung nicht.

Der schaltungstechnische Aufwand für diese Maßnahme fällt mit zwei zusätzlichen Transistoren gering aus. Das notwendige Steuersignal für die Kontrolle der Schalttransistoren steht bereits zur Verfügung, da es für die Aktivierung des Power-Down-Modus ohnehin benötigt wird.

Der Flächenbedarf der beiden zusätzlichen Transistoren fällt ebenfalls vernachlässigbar gering aus. Die Verbindung der Gateanschlüsse der zu schützenden Eingangstransistoren der Differenzstufe muss nicht sehr niederohmig ausgeführt werden, sodass Transistoren mit minimalen Abmessungen verwendet werden können.

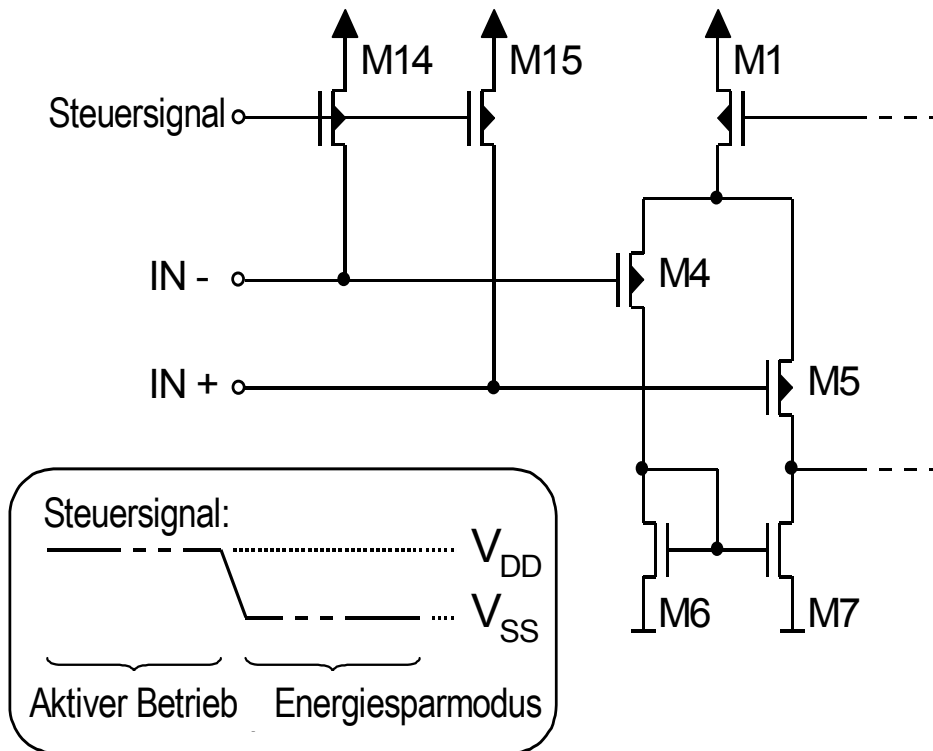


Abbildung 7.2: Schaltungsbeispiel für ‚Reliability by Design‘. Während der ‚Power-Down-Phasen‘ werden die Gateanschlüsse der p-Kanal Eingangstransistoren über zwei zusätzliche Transistoren (M14, M15) mit der Betriebsspannung V_{DD} verbunden. Auf diese Weise werden hohe elektrische Felder über den Gateoxiden der Eingangstransistoren M4 und M5 verhindert. Das Transistorpaar wird so vor BT-Belastungen geschützt.

8 Bias-Temperature-Stress-Problematik zukünftiger CMOS-Prozesse

Für die Weiterentwicklung der CMOS-Halbleitertechnologie werden bei jeder neuen Technologiegeneration kleinere Strukturgrößen für die Transistoren eingesetzt. Neben diesem reinen ‚Shrink‘-Prozess müssen jedoch insbesondere bei Transistoren mit Gatelängen weit unter 100nm weitere Maßnahmen ergriffen werden, um eine funktionale Verkleinerung zu ermöglichen und um einen Fortschritt bei der Performance der Technologie erreichen zu können.

Oft werden dafür bei der Prozessierung neue Materialien eingesetzt und grundlegende Änderungen in der Prozessierung vorgenommen, ohne die die gesetzten Technologieziele nicht erreicht werden können [ITRS04]. Neben den gezielten Verbesserungen verschiedener elektrischer Parameter können diese teilweise tief in den Herstellungsprozess eingreifenden Maßnahmen jedoch auch zu neuen oder zur Verschärfung bereits bestehender Zuverlässigkeitsproblematiken führen.

In diesem Kapitel soll das Augenmerk auf dem Einfluss verschiedener neuer Materialien und wichtiger Prozessänderungen auf die BTS-Festigkeit der Transistoren liegen. Behandelt werden drei ausgewählte aktuelle Themen. Als Erstes sollen die Auswirkungen der Einführung der ‚Dual Workfunction‘-Technik bei DRAM-Technologien betrachtet werden, die im Gegensatz zu Logik-Technologien bei vielen Halbleiterherstellern erst kürzlich vollzogen wurde. Zweitens sollen SOI-Ansätze betrachtet werden und deren BTS-Festigkeit diskutiert werden. Darüber hinaus soll als dritter Punkt ein Ausblick auf die für die 32nm-Technologiegeneration geplante Abkehr von der herkömmlichen planaren Transistorbauform gegeben werden und die damit zu erwartenden verbundenen Probleme behandelt werden.

8.1 Moderne ‚Dual Workfunction‘-Speichertechnologien

Wird bei der Ausführung der Gateelektrode von p- und n-Kanal Transistoren eine unterschiedliche Dotierung des polykristallinen Siliziummaterials verwendet, können Kurzkanaleffekte wirksamer unterdrückt werden und die Performance des p-Kanal Transistors deutlich gesteigert werden. Da bei vielen CMOS-Schaltungen die Performance der Gesamtschaltung durch die Performance der p-Kanal Transistoren limitiert wird, kann durch diese so genannte ‚Dual Workfunction‘-Technologie ein deutlicher Performancegewinn erreicht werden [GHu85].

Im Gegensatz zu Halbleitertechnologien, die für den Einsatz bei Logik- und Analogschaltungen konzipiert werden und bei denen der Einsatz dieser Technik seit vielen Technologiegenerationen üblich ist, wurde bei Speichertechnologien sehr lange auf den Einsatz unterschiedlich dotierter Gateelektroden aus Kostengründen verzichtet. Die Anforderung moderner Speicherprodukte mit immer schnelleren Zugriffszeiten und Datenraten erfordern nun jedoch, diese performancesteigernde Maßnahme auch für DRAM-Halbleitertechnologien einzuführen.

Anstatt die Gates von p- und n-MOSFETs identisch zu dotieren, werden bei Dual-Workfunction-Prozessen unterschiedliche Dotierstoffe verwendet. Für das Polysilizium des p-Kanal Transistors einer Dual-Workfunction-Technologie wird eine im Regelfall mit Bor dotierte Gateelektrode verwendet. Durch diese p-Dotierung kann der p-Kanal Transistor als elektrisch vorgespannt betrachtet werden. Dies führt zu mehreren Vorteilen:

Die bislang für die Einstellung der Einsatzspannung bei p-Kanal Transistoren einer Single-Workfunction-Technologie übliche Doppeldotierung des Kanalbereiches mit n- und p-Dotierstoffatomen kann wegfallen. Die Beweglichkeit der Ladungsträger wird durch die geringere Dotierstoffkonzentration weniger beeinträchtigt. Die Einsatzspannung des Transistors ist mithilfe der p-dotierten Gateelektrode leichter zu kontrollieren und kann niedriger eingestellt werden. Insbesondere wird hierdurch die Möglichkeit erreicht, einen zur Einsatzspannung des n-Kanal Transistors symmetrischen Wert einstellen zu können. Als weiteren positiven Effekt weist eine auf diesem Weg eingestellte Einsatzspannung eine geringere Abhängigkeit von der Temperatur auf.

Die p-Kanal Transistoren einer Dual-Workfunktion-Technologie sind üblicherweise Surface-Channel Transistoren, die p-Kanal Transistoren einer Single-Workfunktion-

Technologie üblicherweise ‚Buried Channel‘ Transistoren. Die Aufenthaltswahrscheinlichkeit der Ladungsträger im Kanal eines Surface-Channel p-Kanal Transistors weist eine engere Verteilung in unmittelbarer Nähe zum Gateoxid auf. ‚Surface Channel‘ p-Kanal Transistoren weisen eine bessere Gatekontrolle auf. So können Kurzkanaleffekte besser unterbunden werden, die Steigung im Unterschwellenstrombereich wird reduziert und das I_{ON} / I_{OFF} Verhältnis verbessert sich [GHu85].

Neben diesen positiven und für eine weitere Verkleinerung der Transistorstrukturen notwendigen Eigenschaften, können die beschriebenen Änderungen in der Prozessierung der p-Kanal Transistoren auch zu Problemen führen. Eine Folge der Prozessierungsschritte für einen Surface-Channel p-Kanal Transistor kann eine Verstärkung der Degradation bei BT-Belastungen sein. Bei der Dotierung der Gateelektrode und vor allem durch spätere Diffusion bei folgenden Prozessschritten mit erhöhten Temperaturen können Anteile der Dotierstoffatome durch das Gateoxid in den Kanalbereich des Transistors eindringen. Bedingt durch diesen in der Literatur mit ‚boron penetration‘ benannten Effekt können sehr starke Schwankungen der Einsatzspannungen der p-Kanal Transistoren entstehen. Darüber hinaus wird der Einbau von Ladungsträgern in das Gateoxid (charge trapping) begünstigt und somit die Degradation unter BT-Belastungen verstärkt [Pfiester90, Uwasawa93].

Wenn der ‚boron penetration‘-Effekt auftritt, kann statt eines reinen SiO_2 -Gatedielektrikums ein nitridiertes Siliziumoxid für die Isolierung der Gateelektrode verwendet werden, um weiterhin eine gezielte Einstellung der Einsatzspannung mit geringer Streuung gewährleisten zu können. Das Eindringen der Bor-Dotieratome des Polysiliziums in das Dielektrikum und in den Kanalbereich wird dadurch deutlich abgeschwächt, sodass eine Verschiebung der Schwellenspannung der Transistoren erfolgreich verhindert werden kann.

Ein nitridiertes Gateoxid kann sich jedoch ebenfalls negativ auf die BTS-Festigkeit der p-Kanal Transistoren auswirken. Sowohl die durch die p-Dotierung der Gateelektrode mögliche ‚boron penetration‘ als auch die als Gegenmaßnahme eingesetzten Stickstoffanteile im Gateoxid können sich negativ auf das Degradationsverhalten auswirken [Yamamoto99, Kimizuka99]. Die Lebensdauer der Transistoren nimmt ab.

8.2 Silicon on Insulator

Mit zunehmender Verkleinerung der Transistoren nehmen Kurzkanal-Effekte immer stärker zu. Einer weiteren Steigerung der Gatekontrolle insbesondere in dem Bereich unterhalb des leitenden Kanals durch eine weitere Reduktion der Gateoxiddicke sind nach unten insbesondere durch die stark ansteigenden Leckströme durch das Dielektrikum enge Grenzen gesetzt.

Eine Möglichkeit, den ‚Shrinkprozess‘ weiterhin zu ermöglichen und gleichzeitig weitere Performance steigernde Effekte zu erzielen, besteht in der ‚Silicon on Insulator‘ - Technik (SOI). Statt des üblichen Wafermaterials mit einer relativ dicken reinen Siliziumschicht (Bulk-Material) werden bei der SOI-Technik Wafer mit einer vergrabenen Oxidschicht eingesetzt. Die Siliziumschicht oberhalb des Isolators, die für die Realisierung der Transistoren genutzt wird, ist hier nur etwa 5 – 100nm dick. Die Idee, Transistoren mit dünnen Siliziumschichten auf einen Isolator aufzusetzen, wird seit sehr vielen Jahren verfolgt, scheiterte jedoch recht lange an der Schwierigkeit eine dünne Schicht einkristallinen Siliziums oberhalb eines Isolators in hoher Qualität herzustellen. Inzwischen können SOI-Wafer für die Produktion in großen Stückzahlen kommerziell erworben werden. Die SOI-Technik hat speziell bei High-Performance Anwendungen wie z.B. Mikroprozessoren einen festen Platz im Portfolio der Halbleiterhersteller eingenommen.

Abbildung 8.1 zeigt schematisch den Querschnitt eines MOSFETs in SOI-Technik. Die Diffusionsgebiete des Transistors reichen bis auf die Isolatorschicht hinab. Der Siliziumfilm weist eine Dicke im Bereich von etwa 5 bis 100nm auf, die Oxidschicht im Bereich von 200-500µm. Dieser Aufbau führt zu zahlreichen Vorteilen:

Die für die mögliche Geschwindigkeit einer Schaltung begrenzenden Kapazitäten der Diffusionsgebiete werden reduziert. Der Wegfall des Siliziumbereiches in der Tiefe unterhalb der Diffusionsgebiete führt zu einer deutlichen Reduktion der Kurzkanaleffekte. Insbesondere bei sehr dünnen Siliziumschichten wird eine vollständige Kontrolle der Siliziumschicht durch das Gate erreicht („Full Depletion“). Die Kanallänge der Transistoren kann weiter reduziert werden.

Dazu kommen weitere wichtige Vorteile der SOI-Technik für Transistoren im Schaltungsverbund.

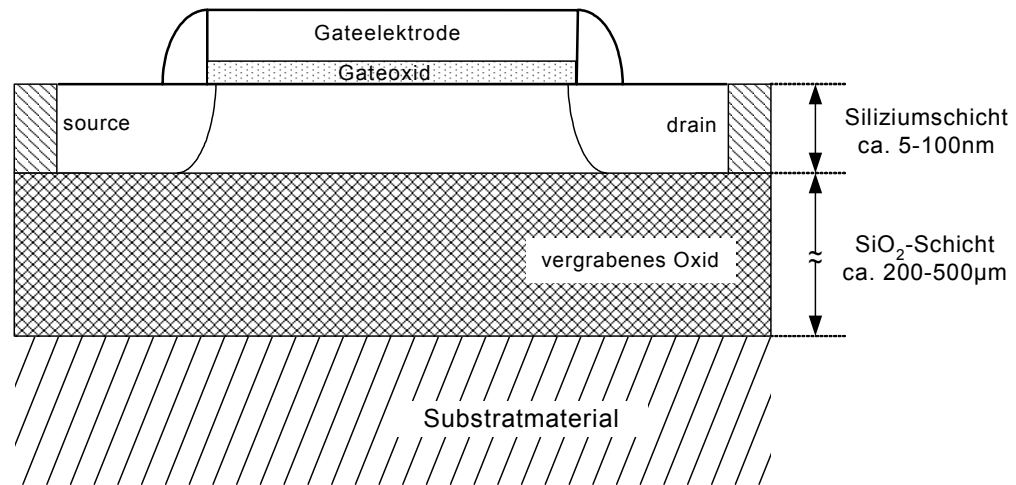


Abbildung 8.1: Schematischer Querschnitt eines Transistors in SOI-Technik. Für den aktiven Bereich des Transistors wird eine dünne Siliziumschicht benutzt, die auf einer vergrabenen Oxidschicht liegt. Die Diffusionsgebiete Drain und Source reichen bis zur vergrabenen Isolationsschicht.

Durch die vollständige Isolation der Siliziumbereiche der Transistoren untereinander werden gegenseitige Beeinflussung der Transistoren, wie z.B. Substrat-Steuereffekte oder Substrat-Rauschen, wirkungsvoll unterdrückt.

Der vollständig isolierte Siliziumbereich der Transistoren führt jedoch auch zu neuen Schwierigkeiten. Die Siliziumdioxidschicht unterhalb der Transistoren führt sowohl zu einer elektrischen als auch zu einer thermischen Isolation. Als elektrischer Sicht kann dies zu Aufladungseffekten des Transistors führen. Ladungen können nicht mehr wie bisher über das Substrat abfließen und führen zum so genannten ‚History-Effekt‘. Das Übertragungsverhalten des Bauelementes erlangt eine zusätzliche Abhängigkeit von der zeitlichen Historie. Je nach Ladungszustand des Substratbereiches in Abhängigkeit der zuvor durchlaufenen Arbeitspunkte reagiert der Transistor unterschiedlich.

Die thermische Isolation des Transistors durch das vergrabene Oxid führt zu einer deutlich veränderten Wärmekapazität des Bauelementes. Die bei Betrieb des Transistors entstehende Wärme kann nur noch langsam abfließen. Arbeitet der Transistor in leistungintensiven Arbeitspunkten, heizt sich das Bauelement in kürzester Zeit auf.

Dieser so genannte ‚Self-Heating-Effekt‘ führt zu einer Abnahme des Kanalstromes, sobald der Transistor für eine Zeitspanne im Nanosekundenbereich² durchgehend mit hohem Stromfluss betrieben wird. Simulationen ergeben, dass in analogtypischen Arbeitspunkten mit einer Self-Heating bedingten Aufwärmung von etwa zusätzlich 50°C in Vergleich zu konventionellen ‚Bulk‘-Transistoren gerechnet werden muss [Knoblinger05].

Sowohl der History-Effekt als auch das Self-Heating müssen bei der Schaltungsentwicklung berücksichtigt werden und erfordern mehr Aufwand im Vergleich zu Bulk-Technologien.

Auch Auswirkungen der SOI-Technik auf die der Zuverlässigkeit müssen in Betracht gezogen werden. Wird speziell das Degradationsverhalten von SOI p-Kanal Transistoren bei BT-Belastungen mit Transistoren mit herkömmlichem Substratmaterial verglichen, ergeben sich zunächst vergleichbare Parameterverschiebungen [Bolam00]. Das Self-Heating eines SOI-Transistors kann jedoch zu einer Erhöhung der durchschnittlichen Transistortemperatur führen. Die für die Berechnung der Lebensdauer herangezogenen Temperaturen, die aus dem Produktanforderungsprofil hervorgehen, müssen für die Berücksichtigung des Self-Heatings nach oben korrigiert werden (siehe auch Kapitel 2.6.3.2). Da eine höhere Temperatur bei BTS zu einer Verstärkung der Schädigungsmechanismen führt (vergleiche Kapitel 2.6.2.2), kann der Self-Heating-Effekt zu einer Reduktion der Lebensdauer von Transistoren in SOI-Technik führen.

8.3 Zukünftige Tansistor-Bauformen

Auch mithilfe der SOI-Technik kann der Shrink-Prozess nicht beliebig weitergeführt werden [ITRS04]. Um den für die vollständige Gatekontrolle notwendigen Full-Depletion-Effekt zu erreichen, darf die Schichtdicke des eingesetzten Siliziums nicht zu groß sein. Als Faustformel gilt, dass die erreichbare minimale Kanallänge mit Full-Depletion-Effekt etwa um den Faktor drei größer ist als die Silizium-Schichtdicke oberhalb des Oxides. Bei Transistoren in Full-Depletion-SOI Technik zeigt die Schichtdicke einen deutlichen Einfluss auf die elektrischen Parameter der Bauelemente. Eine Schwankung der Schichtdicke führt insbesondere zu

² bei dünnen Siliziumschichtdicken im Bereich unter 100nm

Schwankungen der Einsatzspannung der Transistoren. Diese Anforderungen an eine Siliziumschichtdicke von z.B. 5nm auf einer vergrabenen Oxidschicht führen zu hohen Prozesskosten.

Ein aussichtsreicher Ansatz, die Gatelänge der Transistoren weiter reduzieren und die Performance weiter steigern zu können, beruht auf der Abkehr von der rein planaren Integration von Transistorstrukturen. Um die Kontrolle des Kanalbereiches weiter zu verbessern, steuert die Gateelektrode nicht nur wie bisher von oben den Feldeffekt, sondern von mehreren Seiten des Kanalbereiches. Als dafür geeigneteste Bauform hat sich bislang der so genannte FinFET herausgestellt.

Abbildung 8.2 zeigt schematisch die Transistorbauform in einer räumlichen Darstellung. Aus dem auch hier zum Einsatz kommenden SOI-Material werden schmale Stege herausgeätzt. Diese namensgebenden Silizium-Finnen stellen das Substratmaterial des Transistors dar. Abbildung 8.3 zeigt eine Aufnahme eines Rasterelektronenmikroskops (REM) des Querschnitts einer Siliziumfinne. Das Gateoxid umschließt diese Finnen vollständig für einen Bereich, der später die Kanallänge der Bauelemente definiert. Die Gateelektrode wiederum umschließt dieses Gateoxid. So wird eine Kanalgeometrie erreicht, die in ihrer Länge der Finnenlänge und in der Kanalweite etwa zweimal der Finnenhöhe plus der Finnenbreite entspricht. Die Finnenhöhe wird üblicherweise nicht variiert, sodass größere Kanalweiten durch eine Parallelschaltung mehrerer Finnen erreicht werden müssen. Abbildung 8.4 zeigt eine REM-Aufnahme eines Transistors, bei dem eine solche Parallelschaltung mit vier Finnen realisiert wird.

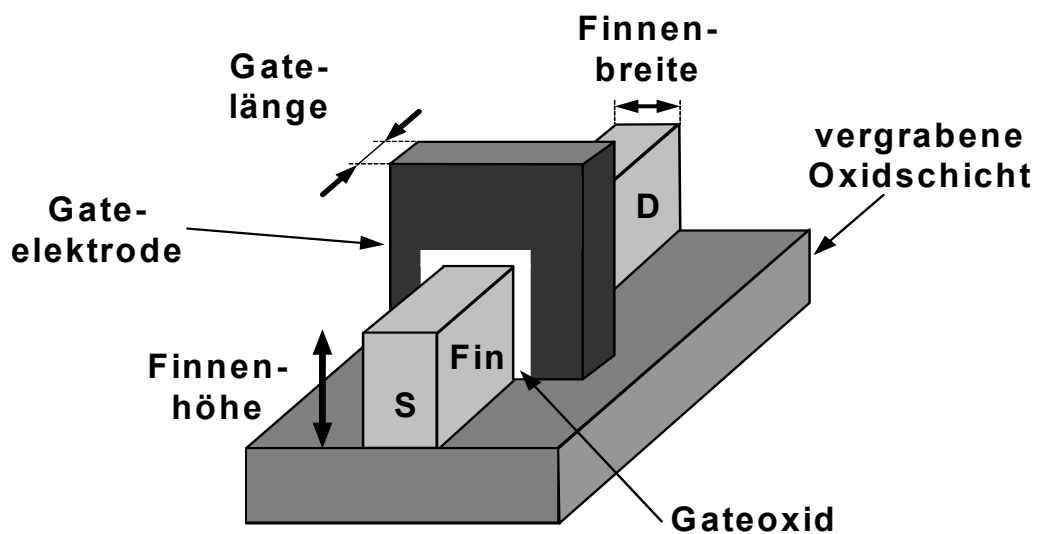


Abbildung 8.2: Schematische Darstellung eines FinFET Transistors in SOI-Technik.

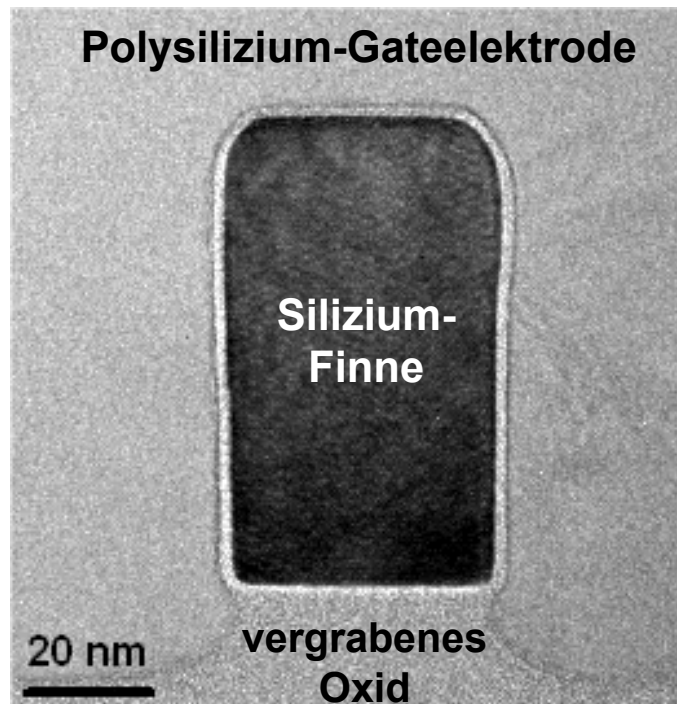


Abbildung 8.3: REM-Aufnahme eines Querschnitts durch die Silizium-Finne eines FinFET auf SOI-Material. Technologie T15.

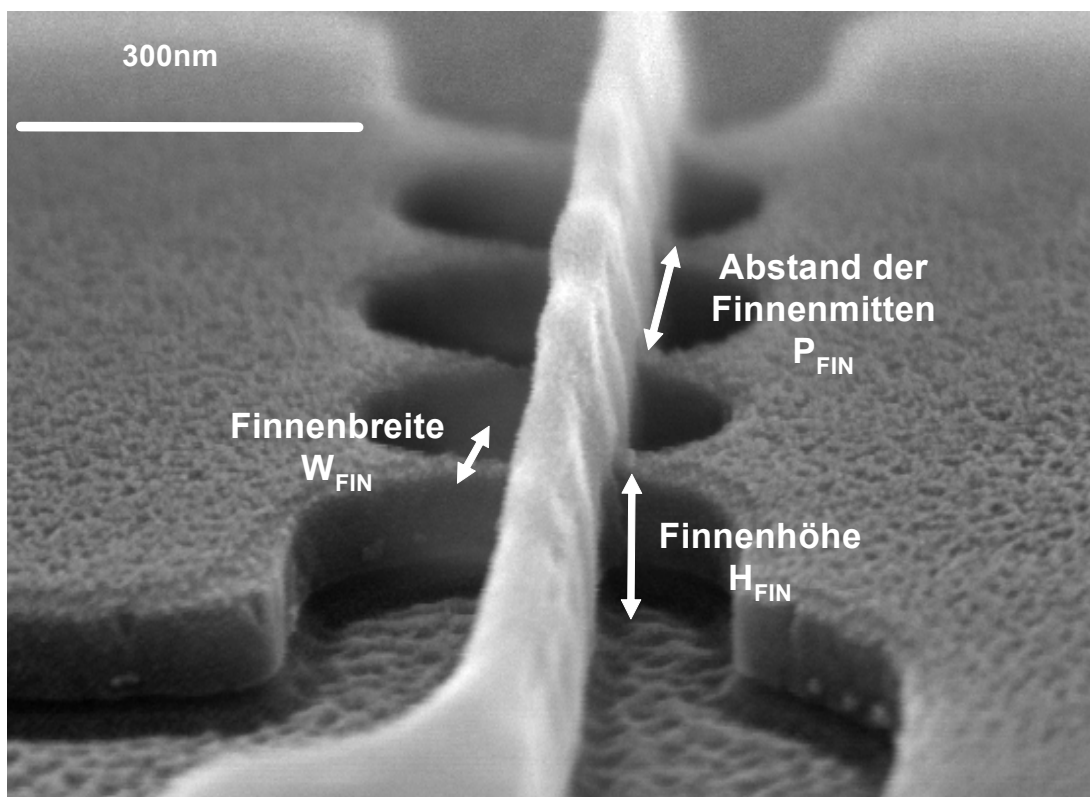


Abbildung 8.4: REM-Aufnahme eines Transistors mit vier parallel geschalteten Siliziumfinnen mit eingezeichneten Maßen für Finnenhöhe, Finnenbreite und Abstand der Finnen. Technologie T15

Die Gesamtfläche des Transistor fällt bei entsprechender Dimensionierung des Abstands der n parallel geschalteten Finnen P_{Fin} und der Finnenhöhe H_{Fin} kleiner aus als die eines planaren MOSFET mit vergleichbarer effektiver Weite W_{eff} . Bei der untersuchten Technologie T15 (Anhang A1) gilt:

$$W_{eff} / (n \cdot P_{Fin}) > 1 \quad (8.1)$$

Durch das den Kanalbereich des Transistors umschließende Gate wird eine besonders gute Kontrolle der Ladungsträger erreicht. Diese gute Kontrolle des Kanals wird im Vergleich zur planaren Bauform bereits bei größeren Gateoxiddicken erreicht.

Aufgrund des bis auf die Siliziumdicke identischen Grundmaterials zur SOI-Technik mit planaren Transistoren, genießt die FinFET-Technik alle Vorteile der zuvor beschriebenen SOI-Technik. Der Full-Depletion-Zustand wird bereits im Vergleich zur planaren SOI-Technik dickeren Siliziumschichtdicken erreicht. Der bereits beschriebene Self-Heating-Effekt tritt aufgrund der thermisch isolierenden vergrabenen Siliziumschicht auch bei FinFETs auf. Die damit verbundenen Auswirkungen auf die Zuverlässigkeit bei BTS gelten daher auch hier.

Durch die neue Bauform der Transistoren müssen aber noch weitere Aspekte bzgl. BTS beachtet werden. Werden die Transistorenfinnen in derselben Ausrichtung wie konventionelle planare Transistoren integriert, besitzen die verschiedenen Oberflächen der Finnen unterschiedliche Siliziumgitterausrichtungen [Sze81]. Wie aus Abbildung 8.5a) ersichtlich, zeigt die Fläche auf der Finnenoberseite die Ausrichtung $\{100\}$, die seitlichen Finnenaußenflächen jeweils die Ausrichtung $\{110\}$ auf.

Die Finnenoberseite besitzt dieselbe Kristallorientierung und damit dasselbe Bindungsverhalten zum Oxid wie bei üblichen planaren Transistoren. Die seitlichen Flächen weisen dagegen mit ihrer anderen Kristallorientierung auch andere Oberflächeneigenschaften auf. An der Grenzfläche tritt hier bei der Oxidation im Vergleich zu einer $\{100\}$ -Oberfläche eine größere Anzahl nicht abgesättigter Bindungen auf (Dangling Bonds), die während folgender Prozessschritte typischerweise mit Wasserstoff abgesättigt werden (vergleiche Kapitel 2.2.1).

Bei derartigen Silizium-Gateoxid-Grenzflächen muss mit einer stärkeren Degradation nach BTS gerechnet werden. Während einer BT-Belastung führt die größere Anzahl von belasteten Wasserstoffbindungen auch zu einer größeren Zahl

aufgebrochener Bindungen und somit zu mehr belastungsinduzierten Grenzflächenzuständen. Da die seitlichen Grenzflächen der Finnen typischerweise einen prozentual relativ großen Anteil der Gesamtfläche des Gateoxids des Transistors darstellen, wirkt sich der hier verstärkt auftretende Schaden signifikant auf die Parameterdegradation des Transistors aus. Ein in derselben Ausrichtung wie ein planarer Transistor integrierter FinFET kann daher ein verstärktes Degradationsverhalten bei BTS aufweisen [Maeda04].

Als Abhilfe gegen diese verstärkte Parameterdegradation von FinFETs bei BTS kann eine Rotation der Finnausrichtung auf dem Wafer vorgenommen werden. Wird der Transistor um 45° zur herkömmlichen Ausrichtung gedreht integriert, weisen alle Siliziumoberflächen der Finnen einheitlich eine $\{100\}$ -Orientierung des Kristallgitters auf (Abbildung 8.5b)). Die mit Wasserstoff passivierten Bindungen treten bei einem FinFET mit dieser Ausrichtung wieder in einer zu einem planaren Transistor vergleichbaren Anzahl auf. Die Degradation nach BTS wird gegenüber einem FinFET mit 0° Ausrichtung deutlich reduziert.

Das beschriebene Degradationsverhalten von FinFETs unterschiedlicher Ausrichtung auf dem Wafer wird experimentell durch zahlreiche Belastungsversuche an der Vorfeldtechnologie T15 bestätigt [Pompl05].

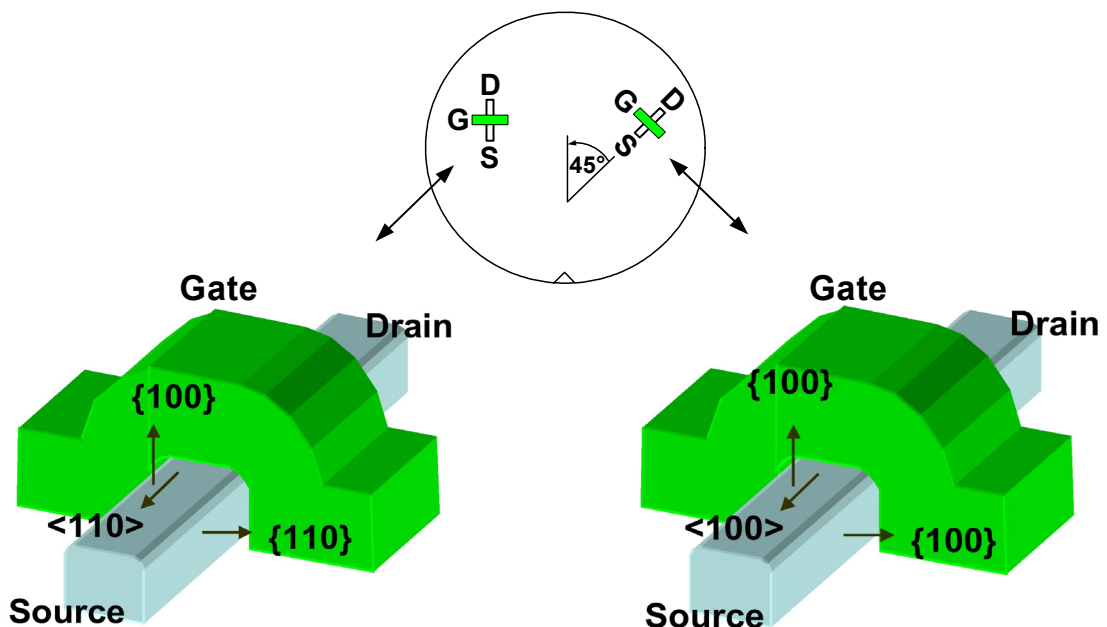


Abbildung 8.5: Schematische Darstellung von FinFETs mit unterschiedlichen Kristallorientierungen des Siliziums der Finnenoberflächen bei:
 a) herkömmlicher Integrationsausrichtung der Finne
 b) um 45° gedrehter Integrationsausrichtung der Finne

9 Zusammenfassung

Bei der Entwicklung moderner CMOS-Technologien müssen neben den Zielen der Performancesteigerung und der Reduktion der Fertigungskosten zunehmend Zuverlässigkeitsaspekte berücksichtigt werden. Die mögliche Lebensdauer von Transistoren wird heute insbesondere durch die bei der Kombination eines hohen elektrischen Gateoxidfeldes mit einer erhöhten Temperatur hervorgerufene Parameterdegradation bestimmt. Diese Belastung durch ‚Bias Temperature Stress‘ (BTS) ist insbesondere bei p-MOS Transistoren bei negativen Gatespannungen kritisch.

Die Degradation nach BTS muss bei analogen und bei digitalen Anwendungen unter jeweils speziellen Gesichtspunkten beachtet werden. Im Analogbereich ist vor allem die unsymmetrische Degradation von Transistorpaaren problematisch, an deren Matchingverhalten hohe Anforderungen gestellt werden. In Digitalanwendungen können aufgrund belastungsbedingter Parameterdegradationen der Transistoren nach BTS zeitkritische Schaltungsteile größere Verzögerungen aufweisen und somit die Funktion der Schaltung beeinträchtigen.

Die mögliche Lebensdauer einer Schaltung und die auf Basis der Produktspezifikationen geforderte Betriebszeit liegen bei modernen CMOS-Technologien immer dichter beieinander. Die Notwendigkeit, die Lebensdauer einer Schaltung möglichst genau berechnen zu können, gewinnt daher zunehmend an Bedeutung. Der bislang übliche worst-case Ansatz für die Lebensdauerextrapolation reicht in einigen Fällen nicht mehr aus. Es kann der Fall eintreten, dass die Lebensdauerberechnung auf Basis der worst-case Betrachtung nicht die notwendige Lebensdauer ergibt, unter realen Bedingungen die Transistoren jedoch durchaus eine ausreichende Lebensdauer erreichen. Die Forderung, die Lebensdauer auch auf Basis der worst-case Betrachtung zu erreichen, führt in diesen Fällen zu unnötigen Entwicklungsaufwänden („over-engineering“).

In dieser Arbeit wird auf der Basis zahlreicher experimenteller Untersuchungen an verschiedenen CMOS-Prozessen ein physikalisches Modell erarbeitet, das eine Beurteilung der Lebensdauer von p-Kanal Transistoren erlaubt. Dafür wird neben homogenen BT-Belastungen speziell das Degradationsverhalten unter inhomogenen Belastungsbedingungen untersucht.

Liegt während der BT-Belastung zusätzlich ein durch eine Drainspannung hervorgerufenen horizontales elektrisches Feld an, fällt die Degradation im Vergleich zum worst-case Fall mit alleinigem vertikalen Feld schwächer aus. Diese Belastungsbedingungen werden an Schaltungsbeispielen für Analog- und Hochfrequenzanwendungen aufgezeigt und erklärt.

Das entwickelte Modell beschreibt die Parameterdegradation nach BTS auch in Belastungsfällen mit einer von null verschiedenen Belastungs-Drainspannung. Dabei wird die Abhängigkeit der Drainstromdegradation von dem Charakterisierungsarbeitspunkt des geschädigten Transistors korrekt berücksichtigt. Die Qualität der Modellaussagen und ihre Prozessunabhängigkeit werden durch Messungen an verschiedenen Prozessen verifiziert. Die Vergleiche der modellierten Werte mit experimentell ermittelten Daten ergeben in allen Fällen eine sehr gute Übereinstimmung. Darüber hinaus wird ein vereinfachtes Modell mit nur geringen Einbußen der Genauigkeit erstellt, das auch ohne Computerunterstützung für einfache Handrechnungen benutzt werden kann. Es beschreibt die Degradation in wichtigen Analoganwendungen mit Arbeitspunkten mit niedrigen effektiven Gatespannungen mit guter Genauigkeit.

Somit kann die Lebensdauer mit dem neuen Modell produktrelevanter und exakter berechnet werden. Die dafür notwendige Vorgehensweise bei der Lebensdauerbewertung von p-Kanal Transistoren bei BTS wird diskutiert und es werden einfache Richtlinien aufgestellt.

Degradationsmodelle wie dieses, die neben dem Worst-Case auch die Belastung in verschiedenen Belastungsarbeitspunkten und die daraus resultierende Parameterdegradation in Abhängigkeit vom Charakterisierungsarbeitspunkt beschreiben können, stellen eine Voraussetzung für die ebenfalls in dieser Arbeit behandelte Simulation belasteter Schaltungen dar. Mithilfe eines Aging-Simulators wird die Funktion einer gewählten Beispielschaltung nicht nur mit konstanten Transistorparametern, sondern auch nach Belastung simuliert und mit der unbelasteten Schaltung verglichen. Die Betrachtung der Auswirkungen der Parameterdegradation basiert dabei nicht mehr auf der isolierten Betrachtung einzelner Transistoren, sondern auf der Betrachtung aller eingesetzten Transistoren im Schaltungsverbund. Die Lebensdauer der Schaltung kann so präzise ermittelt werden.

Da es immer schwieriger und kostenintensiver wird, mit prozesstechnischen Maßnahmen allein die Zuverlässigkeit speziell bei BTS sicherzustellen, muss Zuverlässigkeit von Schaltungen immer mehr als eine gemeinsame Aufgabe der

Technologieentwicklung und des Schaltungsdesigns verstanden werden. ‚Reliability by Design‘ gewinnt daher zunehmend an Bedeutung. Durch den beschriebenen Aging-Simulator können schaltungstechnische Maßnahmen zur Erhöhung der Zuverlässigkeit unterstützt werden. Mit seiner Hilfe können die notwendigen Schaltungsmodifikationen auf wenige relevante Transistoren oder Schaltungsteile beschränkt werden, so genannte Hot-Spots. Am Beispiel einer Analogschaltung wird eine derartige Vorgehensweise aufgezeigt.

Um gegenüber heutigen CMOS-Technologiegenerationen eine Weiterentwicklung zu ermöglichen, reicht für die gewünschte Performancesteigerung das alleinige ‚Shrinken‘ oft nicht mehr aus oder ist nicht mehr im erforderlichen Maße möglich. Es müssen daher neben dem reinen ‚Shrinkprozess‘ auch neue Materialien und neue Bauformen wie z.B. die in dieser Arbeit auch untersuchten FinFETs in Betracht gezogen werden. Durch neue Materialien und neue Transistor-Bauformen werden aber Zuverlässigkeitsrisiken potenziell verstärkt, wie die Beispiele in dieser Arbeit bereits zeigen. Zuverlässigkeitsaspekte müssen daher bei der Prozessentwicklung von Anfang an berücksichtigt werden.

Anhang

A1 Prozessdaten

Name	Technologie Knoten	Min. Kanal-länge	Betriebs-spannung V_{DD}	Gateoxid -dicke $d_{ox,phys}$	Einsatz-Spannung $V_{th,p}$	Entwicklungs-status
T6	0,65 μ m	0,65 μ m	5V	15nm	-800mV	produktiv
T7	0,35 μ m	0,35 μ m	3,3V	7,5nm	-600mV	produktiv
T9	0,25 μ m	0,25 μ m	2,5V	4,5nm	-450mV	produktiv
T9D2	0,25 μ m	0,25 μ m	2,5V	4,5nm	-450mV	Entwicklung
T10	0,18 μ m	0,18 μ m	1,8V	3,3nm	-400mV	produktiv
T11	0,13 μ m	0,12 μ m	1,5V	2,2nm	-330mV	produktiv
T12	90nm	80nm	1,2V	1,6nm	-310mV	produktiv
T13	65nm	60nm	1,2V	1,2nm	-300mV	Entwicklung
T15	32nm	45nm	1,8V	1,8nm	-350mV	Entwicklung

Die untersuchten Technologien T6, T7, T9, T10, T11, T12 werden bei der Infineon Technologies AG eingesetzt. Die Technologie T9D2 wurde innerhalb des DIRDA-Deuterium-Projektes entwickelt. Teilnehmer an diesem Projekt waren unter anderem ST Microelectronics, Infineon Technologies, Polimi, Gressi und Air Liquide. Die SOI-FinFET Technologie T15 entwickeln die Infineon Technologies AG und Texas Instruments gemeinsam bei International Sematech.

A2 Messaufbau und Teststrukturen

Um die Veränderung der elektrischen Eigenschaften eines Transistors durch BT-Belastungen aufnehmen zu können, bedarf es einer präzisen messtechnischen Umgebung. Der Messaufbau muss sowohl die Belastung des Transistors durch Temperatur und Spannung als auch die Messwertaufnahme für die Charakterisierung übernehmen. In diesem Anhangskapitel soll auf die Konfiguration der eingesetzten Messplätze, auf die Funktion der einzelnen benutzten Geräte und auf die vermessenen Teststrukturen eingegangen werden.

Für die Untersuchungen der Transistoren stehen zahlreiche verschiedene Messplätze zur Verfügung, die jedoch in ihrem prinzipiellen Aufbau übereinstimmen.

Der schematische Messaufbau ist in Abbildung A.1 skizziert. Die Messungen finden innerhalb eines vor Lichteinfall schützenden Metallkastens statt, der auch andere Bereiche des elektromagnetischen Spektrums abhält. Eine große Bedeutung innerhalb des Messaufbaus muss der Temperaturregelung der Proben zugewiesen werden. Alle Untersuchungen innerhalb dieser Arbeit werden auf Scheibenebene („Wafer-Level“) vorgenommen. Die Siliziumscheiben liegen während der Messungen auf einem Thermochuck, mit dem die Temperatur der Proben eingestellt und konstant gehalten werden kann. Der Durchmesser der Wafer kann bis zu 300mm (12“) betragen. Es können Temperaturen zwischen -40°C und 200°C eingestellt werden. Für Werte unterhalb der Raumtemperatur werden Chuck und Wafer fortwährend mit Stickstoff umspült, um die Bildung von Kondensationsfeuchtigkeit an Apparatur und Siliziumscheibe zu verhindern.

Der Chuck ist elektrisch vom übrigen Messaufbau isoliert. Für eine weitere Entkopplung des Substratmaterials von eventuellen Störeinflüssen kann bei Bedarf eine spezielle Folie zwischen Wafer und Chuck eingebracht werden. Die Folie ist auf der Basis von Glasfaserpolymeren aufgebaut und ermöglicht so eine gute elektrische Isolation bei einem gleichzeitig geringen Wärmewiderstand. Der spezifizierte Wärmeleitwert des Folienmaterials ist etwa um den Faktor 6 größer als der des Siliziums. Der nötige gute Wärmeübergang zwischen der Waferauflage und dem eigentlichen zu untersuchenden Transistor wird so sichergestellt.

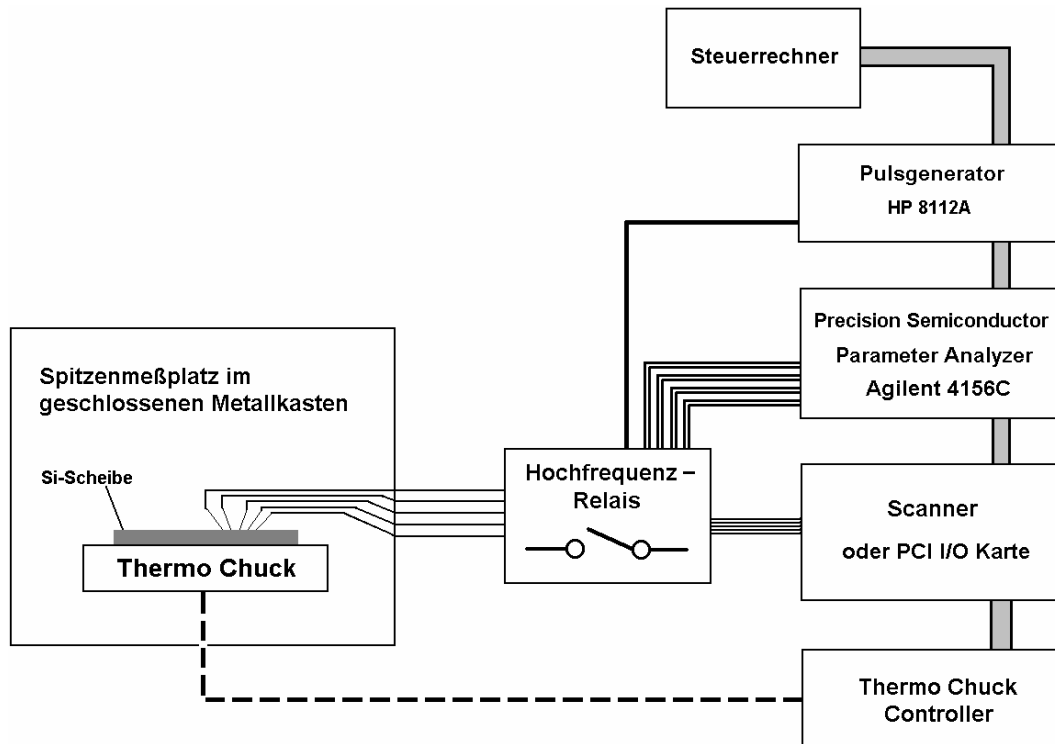


Abbildung A.1: Schematischer Aufbau der für die experimentellen Untersuchungen eingesetzten Messplätze

Das zentrale Messgerät ist ein Agilent 4156C Precision Semiconductor Parameter Analyzer. Dieses Gerät stellt die nötigen Gleichspannungen zur Verfügung und misst gleichzeitig Ströme. Innerhalb des benutzten Messbereiches liegt die Auflösungsgrenze des Gerätes bei 10fA . Um derart kleine Ströme messen zu können³, muss das gesamte Messsystem sorgfältig geplant, aufgebaut und letztendlich auch verifiziert werden. An Verkabelung, Abschirmung und Masseführung müssen besondere Ansprüche gestellt werden. Die Zuführungen zwischen dem Analyzer und dem Messobjekt erfolgen daher über so genannte Kelvin-Kabel. Um Störeinflüsse möglichst gering zu halten, besitzen diese Kabel getrennte Leitungen für Force und Sense sowie eine aktive Abschirmung. Die für die in Kapitel 3.3.3 beschriebene Charge-Pumping-Charakterisierungstechnik benötigten Signale am Gate des Transistors werden durch einen

³ Innerhalb der in Kapitel 3.3.3 beschriebenen Charge-Pumping Charakterisierungen werden Ströme im Bereich einiger hundert Femtoampere aufgezeichnet.

HP 8112A Pulse Generator erzeugt. Auch hier werden aufgrund der Frequenzen in der Größenordnung bis zu einigen MHz bestimmte Anforderungen an den Aufbau des Messplatzes gestellt. Eine HF-optimierte Kabelführung und die konsequente Einhaltung der 50 Ohm Wellenwiderstände sind dafür notwendig.

Um einen automatisierten Messablauf während der Charakterisierung der unterschiedlichen Kennparameter und der Belastung des Transistors ohne manuelles Umstecken der Messgeräte zu ermöglichen, werden Hochfrequenzrelais eingesetzt. Diese Relais besitzen ebenfalls 50 Ohm Wellenimpedanz und weisen einen Innenwiderstand kleiner als 0,3 Ohm auf. Wie aus Abbildung A ersichtlich, wird mit ihnen unter anderem die Umschaltung des Gatekontaktes auf die beiden unterschiedlichen Spannungsquellen Pulsgenerator und Analyzer durchgeführt. Ebenso werden unterschiedliche Abschlusswiderstände während der Kennlinien- und einer Charge-Pumping-Messung an das Gate geschaltet. Die Steuerung dieser Relais erfolgt ebenfalls automatisiert über eine Steuereinheit („Scanner“) oder eine PCI I/O Erweiterungskarte für den Steuerrechner.

Sämtliche eingesetzten Geräte werden über einen IEEE-Messgerätebus zentral durch einen Steuerrechner kontrolliert. Die Steuersoftware basiert auf der grafischen Programmiersprache LabVIEW. Das Programm ermöglicht dem Benutzer Zugriff auf sämtliche Parameter der Belastungs- und Charakterisierungsphasen, ohne die Messapparaturen manuell bedienen zu müssen. Die vollständige Automatisierung der Mess- und Belastungszyklen garantiert exakt eingehaltene Zeitabläufe der Belastungsphasen sowie der Charakterisierung. So wird die Reproduzierbarkeit der Messungen gewährleistet. Darüber hinaus können durch den Einsatz automatischer Systeme bei Bedarf auch mehrere Teststrukturen an verschiedenen Stellen des Wafers automatisch in Folge angefahren und vermessen werden.

Die zu messenden Transistoren befinden sich innerhalb von Teststrukturen, die speziell für Wafer-Level-Belastungsuntersuchungen entworfen wurden. Beim Layout der Strukturen wurde auf eine möglichst rückwirkungsfreie Messumgebung für die Transistoren geachtet. Dafür wurden z.B. alle Strompfade innerhalb der Teststruktur möglichst niederohmig ausgeführt, um Verfälschungen der Arbeitspunkte durch Spannungsabfälle zu vermeiden.

Bei Bedarf stehen Teststrukturen zur Verfügung, bei denen spezielle ESD-Schutzstrukturen an den Gate-Anschluss pads auf der ersten Metallebene die Gateoxide der Testtransistoren schützen. Einerseits sollen so ESD-Schäden während der Handhabung der Proben verhindert werden, andererseits schützen die

Strukturen ebenfalls vor Plasmaschäden bei allen der Prozessierung der ersten Metalllagen folgenden Schritten. Bedingt durch die Antennenverhältnisse bei Transistoren mit einem durch ein Pad nach außen geführten Gatekontakt sind die in Teststrukturen verwendeten Einzeltransistoren einem erhöhten Risiko für Schäden durch Plasmabelastungen ausgesetzt. Ohne derartige Schutzmaßnahmen leidet die Vergleichbarkeit zu den in Schaltungen eingesetzten Transistoren, die meistens durch den Schaltungsverbund besser vor Plasmaschäden geschützt sind [Simon99, Wang04].

Typische ESD-Schutzstrukturen schränken jedoch die Charakterisierungsmöglichkeiten der geschützten Transistoren ein. Über die an der Gateelektrode angeschlossene Schutzstruktur fließen geringe Sperr- und Leckströme zum Substrat. Der Gatestrom des Transistors kann daher bei derart geschützten Bauelementen nicht durch eine Messung am Gatepad des Transistors bestimmt werden, da er durch die Fehlerströme der Schutzstruktur überlagert wird.

Auch eine Charakterisierung der Transistoren mithilfe der Charge-Pumping-Methode (Kapitel 3.3.3) wird durch die Schutzmaßnahmen behindert. Typische Schutzstrukturen leiten die Gateladung bei einem Wechsel der Polarität ab. Die alternierende Umladung der Gateelektrode, wie sie durch das Trapezsignal während der Charge-Pumping-Charakterisierung erreicht werden soll, wird unterdrückt. Eine Charge-Pumping-Charakterisierung ist deshalb nicht möglich.

Für die Untersuchungen in dieser Arbeit standen sowohl ungeschützte als auch mit ESD-Schutzstrukturen versehene Transistor-Teststrukturen zur Verfügung.

A3 Modellberechnungen

Die Berechnung der Integrale der Gleichung 5.7 aus Kapitel 5.1:

$$\begin{aligned}
 (I + \Delta I) \cdot L &= \int_0^L (I + \Delta I) \, dy \\
 &= W \cdot \mu_0 \cdot C_{OX} \cdot \frac{V_G - V_{th}}{L} \int_0^L \frac{(V_G - V_{th}) \cdot \left(1 - \frac{y}{L}\right) - V_0 \cdot e^{\left[C \cdot (V_{G, stress} - V_{D, stress}) \cdot \frac{y}{L}\right]}}{1 + \Theta \cdot \left\{ (V_G - V_{th}) \cdot \left(1 - \frac{y}{L}\right) \right\}} \, dy \\
 &= \left[\int_0^L \frac{e^{\left(-C \cdot V_D \cdot \frac{y}{L}\right)}}{1 + \Theta \cdot (V_G - V_{th}) \cdot \left(1 - \frac{y}{L}\right)} \, dy + \int_0^L \frac{1}{1 + \Theta \cdot (V_G - V_{th}) \cdot \left(1 - \frac{y}{L}\right)} \, dy \right. \\
 &\quad \left. + \int_0^L \frac{\frac{y}{L}}{1 + \Theta \cdot (V_G - V_{th}) \cdot \left(1 - \frac{y}{L}\right)} \, dy \right] \cdot \frac{W \cdot \mu_0 \cdot C_{OX} \cdot (V_G - V_{th})}{L} \\
 &= \left[\left(\frac{Ei \left(C \cdot (V_{G, stress} - V_{D, stress}) \cdot \frac{L \cdot \Theta \cdot (V_G - V_{th}) - 1 - \Theta \cdot (V_G - V_{th})}{\Theta \cdot (V_G - V_{th})} \right)}{\Theta \cdot (V_G - V_{th})} \right) \right. \\
 &\quad \left. - \frac{Ei \left(-C \cdot (V_{G, stress} - V_{D, stress}) \cdot \frac{1 + \Theta \cdot (V_G - V_{th})}{\Theta \cdot (V_G - V_{th})} \right)}{\Theta \cdot (V_G - V_{th})} \right) \\
 &\quad \cdot e^{-C \cdot (V_{G, stress} - V_{D, stress}) \cdot \frac{1 + \Theta \cdot (V_G - V_{th})}{\Theta \cdot (V_G - V_{th})}}
 \end{aligned}$$

$$\begin{aligned}
& + \frac{-\ln(L \cdot \Theta \cdot (V_G - V_{th}) + 1 + \Theta \cdot (V_G - V_{th}))}{\Theta \cdot (V_G - V_{th})} + \frac{\ln(1 + \Theta \cdot (V_G - V_{th}))}{\Theta \cdot (V_G - V_{th})} \\
& + \frac{-(L \cdot \Theta \cdot (V_G - V_{th}) + \ln(L \cdot \Theta \cdot (V_G - V_{th})) - 1 - \Theta \cdot (V_G - V_{th}))}{\Theta^2 \cdot (V_G - V_{th})^2} \\
& - \frac{\ln(L \cdot \Theta \cdot (V_G - V_{th})) - 1 - \Theta \cdot (V_G - V_{th}) \cdot \Theta \cdot (V_G - V_{th})}{\Theta^2 \cdot (V_G - V_{th})^2} \\
& + \ln(-1 - \Theta \cdot (V_G - V_{th})) \cdot \frac{1 + \Theta \cdot (V_G - V_{th})}{\Theta^2 \cdot (V_G - V_{th})^2} \left. \vphantom{\frac{1 + \Theta \cdot (V_G - V_{th})}{\Theta^2 \cdot (V_G - V_{th})^2}} \right] \cdot \frac{W \cdot \mu_0 \cdot C_{OX} \cdot (V_G - V_{th})}{L}
\end{aligned}$$

Die in der Lösung vorkommende Ei-Funktion

$$Ei(x) = \gamma + \ln(x) + \sum_{n=1}^{\infty} \frac{x^n}{n \cdot n!}$$

lässt sich tabellarisch aus mathematischen Handbüchern entnehmen.

Verwendete Symbole und Abkürzungen

ΔQ_{it}	[C]	belastungsinduzierter Zuwachs der Ladung in Grenzflächenzuständen
ΔQ_{ox}	[C]	belastungsinduzierte Oxidladung
ΔV_{th}	[V]	belastungsinduzierte Einsatzspannungsverschiebung
$\Delta V_{th}(y)$	[V]	belastungsinduzierte Einsatzspannungsverschiebung als Funktion des Ortes entlang des Kanals
f	[Hz]	Frequenz des Charge-Pumping-Zyklus
ϕ_{MS}	[eV]	Differenz der Austrittsarbeiten
ΔN_{GFZ}		Anzahl der belastungsinduzierten Grenzflächenzustände
μ	[cm ² /V]	Ladungsträgerbeweglichkeit
ψ_B	[V]	Potenzialdifferenz zwischen Fermienergie der Kanalladungsträger und dem intrinsischen Fermienergie
C_{ox}	[fF/μm ²]	flächenbezogene Gatekapazität
$d_{add,verarm}$	[nm]	Beitrag zur elektrischen Gateoxiddicke oberhalb des Gateoxides aufgrund der Polydepletion
$d_{add,welle}$	[nm]	Beitrag zur elektrischen Gateoxiddicke unterhalb des Gateoxides aufgrund des Kanalabstandes
$d_{ox,phys}$	[nm]	physikalische Gateoxiddicke
E_a	[eV]	Aktivierungsenergie
g_{DS}	[μA/V]	differenzieller Drain-Source-Leitwert

g_m	[mA/V]	Steilheit (Transkonduktanz)
H_{Fin}	[μm]	Höhe der Siliziumfinne eines FinFET
I_{CP}	[A]	Charge-Pumping Strom
I_D	[A]	Drainstrom
k	[eV/K]	Boltzmannkonstante ($8,62 \cdot 10^{-5}$ eV/K)
L	[μm]	Länge des Transistorkanals
N_{GFZ}		Anzahl der gepumpten Grenzflächenzustände
N_{it}	$\left[\frac{1}{\text{cm}^2} \right]$	Grenzflächenzustandsdichte
P_{Fin}	[μm]	Abstand von Finnenmitte zu Finnenmitte
q	[C]	Elementarladung ($1,60218 \cdot 10^{-19}$ C)
Q_{it}	[C]	Ladung in Grenzflächenzuständen
Q_{OX}	[C]	Oxidladung
Q_{SC}	[C]	Ladung der freien Ladungsträger im Kanal
τ	[s]	Lebensdauer
t_{stress}	[s]	Belastungsdauer
T_{char}	$^{\circ}\text{C}$	Charakterisierungstemperatur
T_{stress}	$^{\circ}\text{C}$	Belastungstemperatur
T_{zyklus}	[s]	Periodendauer des Charge-Pumping Zyklus
V_{CP}	[V]	Spannungswert des Signals an der Gateelektrode bei Charge-Pumping-Messungen
$V_{CP,high}$	[V]	Spannungswert des oberen Trapezsignalpegels bei Charge-Pumping-Messungen

$V_{CP,low}$	[V]	Spannungswert des unteren Trapezsignalpegels bei Charge-Pumping-Messungen
$V_{D,char}$	[V]	Charakterisierungs-Drainspannung
$V_{D,stress}$	[V]	Belastungs-Drainspannung
V_{FB}	[V]	Flachbandspannung, allgemein
$V_{FB,n}$	[V]	Flachbandspannung, n-Kanal MOSFET
$V_{FB,p}$	[V]	Flachbandspannung, p-Kanal MOSFET
V_G	[V]	Gatespannung, allgemein
$V_{G,char}$	[V]	Charakterisierungs-Gatespannung
$V_{G,stress}$	[V]	Belastungs-Gatespannung
$V_{Gate,eff}$	[V]	effektive Gatespannung
V_{th}	[V]	Einsatzspannung, allgemein
$V_{th,n}$	[V]	Einsatzspannung n-Kanal MOSFET
$V_{th,p}$	[V]	Einsatzspannung p-Kanal MOSFET
$V_{th,lin}$	[V]	Schwellenspannung im Linearen Bereich
$V_{th,sat}$	[V]	Schwellenspannung im Sättigungsbereich
W	[μm]	Weite des Transistorkanals
W_{eff}	[μm]	Effektive Weite des Transistorkanals bei FinFETs
W_{Fin}	[μm]	Breite der Siliziumfinne eines FinFET

Literaturverzeichnis

- [Armin04] K. v. Armin, E. Borinski, P. Seegebrecht, H. Fiedler, R. Brederlow, R. Thewes, J. Berthold, and C. Pacha, „Efficiency of Body Biasing in 90nm CMOS for Low Power Digital Circuits, Proc. 34th ESSCIRC 2004, pp. 175-178
- [Aur98] S. Aur, T. Grider, V. McNeil, T. Holloway and R. Eklund, “Effects of advanced processes on hot carrier reliability,” Proc. IRPS, 1998, pp. 180-183
- [Baker98] R. J. Baker, H. W. Li, and D. E. Boyce, “CMOS: Circuit Design, Layout, and Simulation”, IEEE Press, Series on Microelectronic Systems, New York, USA, 1998
- [Bolam00] R. Bolam, G. Shahidi, F. Assaderaghi, M. Khare, A. Mocuta, T. Hook, E. Wu, E. Leobandung, S. Voldman, D. Badami, “Reliability issues for silicon-on-insulator”, International Electron Device Meeting, Technical Digest, 2000, pp. 131-134
- [Brower88] K. L. Brower, “Kinetics of H₂ passivation of Pb centers at the (111) Si-SiO₂ interface”, Phys. Rev. B, vol. 38, 1988, p.9657
- [Brox94] M. Brox, A. v. Schwerin, Q. Wang, and W. Weber, “A model for the time- and bias-dependence of p-MOSFET degradation”, IEEE Trans. Electron Devices, 1994, p. 1184
- [Chakravarthi04] S. Chakravarthi, A.T. Krishnan, V. Reddy, C.F. Machala and S. Krishnan; “A Comprehensive Framework For Predictive Modeling of Negative Bias Temperature Instability”, Proc. IRPS 2004, pp. 273-282

- [Chaparala00]** P. Chaparala, J. Shibley, L. Paul, „Threshold voltage drift in PMOSFETS due to NBTI and HCI“, IEEE IRW 2000, pp. 95-97
- [Chung04]** S.S. Chung, C.H. Yeh, H.J. Feng, C.S. Lai, J.-J. Yang, C.C. Chen, Y. Lin, S.C. Chen, and M.S. Liang; “The Impact of STI on the Reliability of a Narrow Width p-MOSFET’s with Advanced ALD N/O Gate Stack, IEEE TDMR Online-Journal, 2004
- [GHu85]** G. Hu and R.H. Bruce, “Design tradeoffs between surface and burried-channel FETs”, IEEE Trans. Electron Devices, vol. ED-32, 1985, p. 584
- [Goser90]** K. Goser, „Großintegrationstechnik, Teil 1: Vom Transistor zur Grundschtaltung“, Hüthig Buch Verlag GmbH, Heidelberg, 1990
- [Goser91]** K. Goser, „Großintegrationstechnik, Teil 2: Von der Grundschtaltung zum VLSI System“, Hüthig Buch Verlag GmbH, Heidelberg, 1991
- [Gregor92]** R. W. Gregor, “On the relationship between topography and transistor matching in an analog CMOS technology”, IEEE Trans. Electron Devices, 1992, p. 275
- [Groeseneken84]** G. Groeseneken, H. E. Maes, N. Betrán, and R. F. de Keersmaeker, “A reliable approach to charge-pumping mearsuements in MOS transistors, “IEEE Trans. Electron Devices, 1984, p. 42
- [Heil35]** O. Heil, British Patent 439,457; 1935
- [Heremans91]** P. Heremans, R. Bellens, G. Groeseneken, A. v. Schwerin, H. E. Maes, M. Brox, W. Weber, “The mechanisemns of hot carrier degradation”, in “Hot-carrier design constrains in MOS devices and circuits”, Hrsg.: C. T. Wang, Van Nostrand Reinhold, New York, 1991

- [Hokari88]** Y. Hokari, IEEE Transaction on Electron Devices, Vol. 35, No. 8, 1988, pp. 1299-1304
- [Hu85]** C. Hu, S.C. Tam, F.-C.Hsu, P.-K.Ko, T.-Y.Chan und K.W.Terrill, "Hot-electron-induced MOSFET degradation – Model, monitor and improvement", IEEE J. Solid-State Circuits, 1985, p.295
- [Huang92]** J. H. Huang, Z. H. Liu, M. C. Jeng, P. K. Ko, and C. Hu, "A physical model for MOSFET output resistance", IEDM Tech. Digest, 1992, p. 569
- [IFX04]** Infineon Technologies AG, "Memory Spectrum Brochure 2004-09", 1Gb DDR2 DRAM Component, www.infineon.com, September 2004
- [Intel04]** Intel Corporation, „Product Brief Intel Itanium2“, Dual Core Server Processor, www.intel.com, 2004
- [ITRS04]** International Technology Roadmap for Semiconductors, Edition 2004, <http://public.itrs.net>
- [Jeppson77]** K. O. Jeppson and C. M. Svensson, "Negative bias stress of MOS devices at high electric fields and degradation of MNOS devices" J. Appl. Phys., 1977, p. 2004
- [Kahng60]** D. Kahng, and M. M. Atalla, "Silicon-silicon dioxide field induced surface devices", IRE Solid State Device Research Conf., 1960
- [Kimizuka00]** N. Kimizuka, K Yamaguchi, K. Imai, T. Iizuka, C.T: Liu, R.C. Keller, T. Horiuchi, "NBTI enhancement by nitrogen incorporation into ultrathin gate oxide for 0.10- μm gate CMOS generation", VLSI Technology Symposium, Digest of Technical Papers, 13-15th June 2000, pp.92–93

- [Kiriata98]** T. Kiriata, M. Gall, K. Hosokawa, J-M. Dortu, H. Wong, P. Pfefferl, B.L. Ji, O. Weinfurtner, J.K. DeBrosse, H. Terletzki, M. Selz, W. Ellis, M.R. Wordemann, and O. Kiehl; „A 220-mm², Four- and Eight-Bank, 256-Mb SDRAM with Single-Sided Stched WL Architecture”; IEEE Journal of Solidf-State Circuits, Vol. 33, No. 11, November 1998, pp. 1711-1719
- [Kizilyalli97]** I. C. Kizilyalli, J. W. Lyding, and K. Hess, “Deuterium Post-Metal Annealing of MOSFET’s for Improved Hot Carrier Reliability, IEEE Electron Device Letters, vol. 18, no. 3, 1997, pp. 81-83
- [Knoblinger05]** G. Knoblinger, F. Kuttner, A. Marshall, C. Russ, P. Haibach, P. Patruno, T. Schulz, W. Xiong, M. Gostkowski, K. Schrufer, C.R. Cleavelin, “Design and Evaluation of Basic Analog Circuits in an Emerging MuGFET Technology”, IEEE International SOI Conference, Honolulu, Oct 3-6th 2005, accepted for publication
- [Ko89]** P. K. Ko, “Approaches to scaling”, in “Advanced MOS device physics”, Hersg.: N. G. Einspruch, Academic Press, San Diego (CA), 1989
- [Krishnan05]** A.T. Krishnan, “NBTI: Process, Device and Circuit”, IRPS 2005, 18 Apr., San Jose, Tutorial Notes, Section 212, pp. 1-28
- [Laksh86]** K. R. Lakshmikumar, R. A. Hadaway, M. A. Copeland, “Characterization and modeling of mismatch in MOS transistor for precision analog design”, IEEE J. Solid-State Circuits, 1986, p. 1057
- [LaRosa97]** G. La Rosa, F. Guarin, S. Rauch, A. Acovic, J. Lukaitis, E. Crabbe, “NBTI-channel hot carrier effects in pMOSFETs in advanced CMOS technologies”, Proc. IRPS, 1997, pp. 282-286

- [Linnenbank98]** C. G. Linnenbank, W. Weber, U. Kollmer, B. Holzapfel, S. Sauter, U. Schaper, R. Brederlow, S. Cyrusian, S. Kessel, R. Heinrich, E. Hoefig, G. Knoblinger, A. Hesener, and R. Thewes, "What do matching results of medium area MOSFETs reveal for large area devices in typical analog applications?", Proc. ESSDERC '98, 1998, p. 104
- [Liu93]** Z. H. Liu, C. Hu, J. H. Huang, T. Y. Chan, M. C. Jeng, P. Ko, and Y. Cheng, "Threshold Voltage Model for Deep-Submicrometer MOSFETs", IEEE Trans. Electron Devices, 86, 1993
- [Lu89]** C. Y. Lu and J. M. Sung, "Reverse short-channel effects on threshold voltage in submicrometer salicide devices," IEEE Electron Device Lett., vol. 10, Oct. 1989, pp. 446-448
- [Maeda04]** S. Maeda; J.A. Choi; J.H. Yang; Y.S. Jin; S.K. Bae; Y.W. Kim; K.P. Suh, "Negative bias temperature instability in triple gate transistors", IEEE International Reliability Physics Symposium Proceedings, 42nd Annual., 25-29 April 2004, pp.8-12
- [Makabe00]** M. Makabe, T. Kubota, T. Kitano, "Bias temperature degradation of pMOSFETs: mechanism and suppression", Proceedings IRPS 2000, pp. 205-209
- [Martin98]** A. Martin, P. O'Sullivan and A. Methewson, "Dielectrics Reliability Measurement Methods: A Review", J. Microelectronics Reliability, Vol. 38, No. 1, 1998, pp. 37-72
- [Mazuré89]** C. Mazuré, M. Orłowski, "Guidelines for Reverse Short-Channel Behavior," IEEE Electron Device Letters, Vol. 10, No. 12, Dec. 1989; pp. 556-558

- [McCrackin63]** F. L. McCrackin E. Passaglia, R.R. Stromberg, H.L. Steinberg, "Measurement of the thickness and refractive index of very thin films and the optical properties of surfaces by ellipsometry", Journal of Research of the National Bureau of Standards, 67A, 1963, pp. 363-367
- [Oehm93]** J. Oehm and K. Schumacher, "Quality Assurance and Upgrade of Analog Characteristics by Fast Mismatch Analysis Option in Network Analysis Environment", IEEE J. Solid-State Circuits, 1993, p. 865
- [Ogawa95]** S. Ogawa and N. Shiono, "Generalized diffusion-reaction model for the low field charge-buildup instability at the Si-SiO₂ interface", Phys. Rev. B, 1995, p. 4218
- [Ogawa96]** S. Ogawa, M. Shimaya, and N. Shiono, "Impact of Negative-Bias Temperature Instability on the Lifetime on Single-Gate CMOS Structures with Ultrathin (4-6 nm) Gate Oxides", Jpn. J. Appl. Phys., 1996, p. 1484
- [OrtizConde02]** A. Ortiz-Conde, F.J. García Sánchez, J.J. Liou, A. Cerdeira, M. Estrada, Y. Yue, "A review of recent MOSFET threshold voltage extraction methods", J. Microelectron. Reliability 42, 2002, p. 583
- [Pacha04_1]** C. Pacha, persönliche Mitteilung, Sept. 2004
- [Pacha04_2]** C. Pacha, M. Bach, K. v. Armin, R. Brederlow, D. Schmitt-Landsiegel, P. Seegebrecht, J. Berthold, and R. Thewes, „Impact of STI-Induced Stress, Inverse Narrow Width Effect, and Statistical VTH Variations on Leakage Currents in 120nm CMOS”, Proc. 34th ESSDERC 2004, pp. 397-400
- [Pelgrom89]** M. J. M. Pelgrom, A. C. J. Duinmaijer, A. P. G. Welbers, "Matching properties of MOS transistors", IEEE J. Solid-State Circuits, 1989, p. 1433

- [Pfiester90]** J.R. Pfiester, F.K. Baker, T.C. Mele, H. Tseng, P.J. Tobin, J.D. Hayden, J.W. Miller, C.D. Gunderson, L.C. Parrillo, "The effects of boron penetration on p+ polysilicon gated PMOS devices", IEEE Trans. Electron Devices, vol. 37, Issue: 8, Aug. 1990, pp.1842 – 1851
- [Pompl05]** T. Pompl, H. C. Mogul, C. Schlünder, M. Kerber, G. Haase, E. Ogawa, J. W. McPherson, W. Xiong, T. Schulz, K. Schrüfer, R. Cleavelin, "TDDDB and NBTI Evaluation of Gate Dielectric Integrity for CMOS Multi-Gate FET Devices with Poly and TiN Gates", submitted to IEEE IEDM, 2005
- [Rashkeev01]** S. N. Rashkeev, D.M. Fleetwood, R. D. Schrimpf, and S.T. Pantelides, "Defect Generation by Hydrogen at the Si-SiO₂ Interface", Phys. Rev. Lett., vol. 87, 2001, p. 165
- [Schlünder99]** C. Schlünder, R. Brederlow, P. Wiczorek, C. Dahl, J. Holz, M. Roehner, S. Kessel, V. Herold, K. Goser, W. Weber, and R. Thewes, "Trapping Mechanisms in Negative Bias Temperature Stressed p-MOSFETs", J. Microelectronics Reliability 1999, 39 (Proc. ESREF), pp. 821-826
- [Schlünder03]** C. Schlünder, R. Brederlow, B. Ankele, A. Lill, K. Goser, and R. Thewes, "On the Degradation of P-MOSFETs in Analog and RF Circuits under Inhomogenous Negative Bias Temperature Stress", Proc. IEEE IRPS, 2003, pp. 5-10
- [Schlünder05_1]** C. Schlünder, R. Brederlow, B. Ankele, W. Gustin, K. Goser, and R. Thewes, "Effects of Inhomogenous Negative Bias Temperature Stress on p-Channel MOSFETs", J. Microelectronics Reliability, vol. 45, Special Section NBTI, 2005, pp. 39-45
- [Schlünder05_2]** C. Schlünder, "Mixed Signal Circuit Reliability – From Device Stress Conditions to IC-Failure", Invited Tutorial, IEEE IRPS, San Jose, Tutorial Notes, Section 121, Apr. 17th 2005, pp. 1-18

- [Schuegraf93]** K. F. Schuegraf, C. C. King and C. Hu, in Proceedings of Technical Papers of the International Symposium on VLSI Technology, Systems, and Applications, 1993, pp. 86-90
- [Schumacher87]** K. Schumacher, „Integrationsgerechter Entwurf analoger MOS-Schaltungen“, R. Oldenburg Verlag, München, Wien, 1987
- [Sell02]** B. Sell, “Interface Characterization of Metal-gate MOS-Structures”, Dissertation, TU Hamburg-Harburg, Germany, 2002
- [Shockley48]** W. Shockley and G. L. Pearson, “Modulation of conductance of thin films of semiconductors by surface charges”, Phys. Rev. 74, 1948, p. 232
- [Simon99]** P. Simon, J.-M. Luchies, W. Maly, “Antenna ratio definition for VLSI circuits [plasma etch damage]”; 4th International Symposium on Plasma Process-Induced Damage, 9-11 May, 1999, pp. 16 - 20
- [Stesmans00]** A. Stesmans, “Integration of Pb defects at the (111) Si/SiO₂ interface with molecular hydrogen: Simultaneous action of passivation and dissociation”, J. Appl. Phys., vol 38, 2000, p. 489
- [Stolk98]** P. A. Stolk, F. P. Widdershoven, and D. B. M. Klaassen, “Modeling Statistical Dopant Fluctuations in MOS Transistors”, IEEE Trans. Electron Devices, 1998, p. 1960
- [Sze81]** S. M. Sze, “Physics of Semiconductor Devices”, Second Edition, NY, John Wiley & Sons, 1981
- [Thewes95]** R. Thewes, “Degradation analoger CMOS-Schaltungen durch heiße Ladungsträger“, Dissertation, Universität Dortmund, Juli 1999, Kapitel 3 und 4

- [Thewes97_1] R. Thewes, M. Brox, K. F. Goser, and W. Weber, "Hot-Carrier Degradation of p-MOSFETs Under Analog Operation", IEEE Trans. Electron Devices, 1997, p. 607
- [Thewes97_2] R. Thewes and W. Weber, "Effects of Hot-carrier Degradation in Analog CMOS Circuits", Proc. Conf. "Insulating Films on Semiconductors" (INFOS), and Microelectronic Engineering 36, 1997, p. 285
- [Thewes99_1] R. Thewes, R. Brederlow, C. Schlünder, P. Wiczorek, A. Hesener, B. Ankele, P. Klein, S. Kessel and W. Weber, "Device Reliability in Analog CMOS Applications", IEEE IEDM, Tech. Digest, 1999, pp. 81-84
- [Thewes99_2] R. Thewes, G. H. Walter, R. Brederlow, C. Schlünder, A. v. Schwerin, R. Jurk, C. G. Linnenbank, G. Lengauer, D. Schmitt-Landsiedel, and W. Weber, "Channel Length Dependence of Hot-Carrier Degradation of LATID-n-MOSFETs under Analog Operation", Proc. IEEE IRPS, 1999, p. 233
- [Thewes00] R. Thewes, R. Brederlow, C. Schlünder, P. Wiczorek, B. Ankele, A. Hesener, J. Holz, S. Kessel, and W. Weber, "MOS Transistor Reliability under Analog Operation", J. Microelectronics Reliability 40, 2000, pp. 1545-1554
- [Tiebout01] M. Tiebout, "Low power, low phase noise, differently tuned quadratur VCO-Design", IEEE Journal of Solid-State Circuits, vol. 36, July 2001, pp. 1018-1024
- [TieSch93] U. Tietze, Ch. Schenk, "Halbleiter-Schaltungstechnik", Springer Verlag, Berlin, Heidelberg u.w., 10. Auflage, 1993

- [Uwasawa93]** K. Uwasawa, T. Mogami, T. Kunio, M. Fukuma, "Scaling limitations of gate oxide in p+ polysilicon gate MOS structures for sub-quarter micron CMOS devices", International Electron Device Meeting, Technical Digest., 5-8 Dec. 1993, pp.895-898
- [Walter98]** G. H. Walter, W. Weber, R. Brederlow, R. Jurk, C.G. Linnenbank, C. Schlünder, D. Schmitt-Landsiegel, and R. Thewes, "Precise quantitative evaluation of the hot-carrier induced drain series resistance degradation in LATID-n-MOSFETs", Proc. ESREF '98 and J. Microelectron. Reliability, vol 38, 1998, p. 1063
- [Wang04]** Z. Wang, "Detection of and Protection against Plasma Charging Damage in Modern IC Technology", Dissertation, University of Twente, Niederlande, 2004
- [Wuensche02]** S. Wuensche, M. Jacunski, H. Streif, A. Sturm, J. Morrish, M. Roberge, M. Clark, T. Nostrand, E. Stahl, S. Lewis, J. Heath, M. Wood, T. Vogelsang, E. Thoma, J. Gabric, M. Kleiner, M. Killian, P. Poechmueller, W. Mueller, G. Bronner; "A 110nm 512Mb DDR DRAM with Vertical Transistor Trench Cell"; VLSI Symposium on Circuits, Digest of Technical Paper.; 13-15th June 2002, pp. 114-118
- [Yamamoto99]** T. Yamamoto, K. Uwasawa, T. Mogami, "Bias temperature instability in scaled p+ polysilicon gate p-MOSFET's", Trans. on Electron Devices, vol 46 ,Issue: 5, May 1999, pp. 921-926

Lebenslauf

Persönliche Daten

Name: Christian Schlünder
geboren am: 12.11.1972 in Menden (Sauerland)
Familienstand: ledig
Anschrift: Laimer Platz 1a, 80689 München
E-mail: christian@schluender.de

Schulischer Werdegang

1979 – 1983 St. Josef-Grundschule Lendringsen
1983 – 1992 Städt. Walram-Gymnasium Menden
Juni 1992 Allgemeine Hochschulreife

Technischer Zivildienst

Sep. 1992 – Okt. 1993 Technische Versorgung u. Instandhaltung
Paracelsus Klinik Hemer

Studium

Okt. 1993 Beginn des Studiums Elektrotechnik,
Universität Dortmund, Schwerpunkt Elektronik

Jan. 1997 – Dez. 1997 Studienbegleitende Tätigkeit: Studentische Hilfskraft
Lehrstuhl für Nachrichtentechnik

März 1998 – Juli 1998 Werkstudententätigkeit: Siemens AG, München
Zentralabteilung Technik (ZT ME 2):

Okt. 1998 – März 1999 Diplomarbeit: „Degradations- und Relaxations-
mechanismen in MOS-Transistoren nach Bias
Temperature Stress“. Durchführung der Arbeit bei
Siemens AG München (Zentralabteilung Technik) /
Infineon Technologies AG, (Corporate Research Dept.)

Berufsstart

Apr. 1999 – Juli 1999 Selbstständigkeit: Beratungstätigkeit für Infineon
Technologies AG, München, Corporate Research Dept.

Juli 1999 – Jan. 2000 Angestellter: Infineon Technologies AG, München
Corporate Research Department (CPR3)

Feb. 2000 – dato Infineon Technologies AG, München
Central Reliability Methodology (RM DEV)

Danksagung

An dieser Stelle möchte ich meinen Dank an alle aussprechen, die zum Gelingen dieser Arbeit beigetragen haben. Im Einzelnen möchte ich danken:

Herrn Prof. Dr.-Ing. Karl Goser für die Ermöglichung und Betreuung dieser Arbeit;
Herrn Prof. Dr.-Ing. Jürgen Oehm für die Arbeit als Koreferent.

Herrn Dr.-Ing. Roland Thewes für die Betreuung seitens der Infineon Technologies AG;

den Herren Dr. rer. nat. Jörg Berthold, Dr.-Ing. Dipl.-Phys. Ralf Brederlow, Dr.-Ing. Dipl.-Phys. Christian Pacha, Dr. rer. nat. Hans Reisinger, Dr.-Ing. Guido Stromberg, Dr. rer. nat. Heiko Nielen, Dipl.-Ing. Andreas Martin, Dipl.-Ing. Werner Muth, Dipl.-Ing. Manfred Budde, Dipl.-Ing. Stefan Holzhauser, Dipl.-Phys. Andrea Mitchell und Dipl.-Ing. Franz Ungar für die zahlreichen technischen und nichttechnischen Diskussionen und Anregungen;

meinem Vorgesetzten der zurückliegenden Jahre Dr. rer. nat. Arnulf Lill und meinem jetzigen Vorgesetzten Dr. rer. nat. Wolfgang Gustin für ihre jederzeitige Unterstützung der Arbeit;

der Infineon Technologies AG München insbesondere der Abteilung RM für die Bereitstellung des technisch gut ausgestatteten Umfeldes.

Nicht unerwähnt lassen möchte ich alle bislang noch nicht genannten Mitarbeiter der Abteilungen RM DEV, RM MON, CPR FEC und CPR ET, die mir immer hilfsbereit gegenüberstanden.

Ein besonderer Dank richtet sich an meine Partnerin Dipl.-Soz.päd./-Soz.arb. Sonja Jabs, die mich vor allem in der Endphase dieser Arbeit unterstützte.