

Charakterisierung von MOS-Transistoren vor und nach Gateoxiddurchbruch

Vom Promotionsausschuß der
Technischen Universität Hamburg-Harburg
zur Erlangung des akademischen Grades
Doktor-Ingenieur
genehmigte Dissertation

von
Alejandro AVELLÁN HAMPE
aus Mudanya

2004

Erster Gutachter:

Prof. Wolfgang Krautschneider

Zweiter Gutachter:

Prof. Wolfgang Bauhofer

Vorsitzender des Promotionsausschusses:

Prof. Jörg Müller

Tag der Einreichung:

24. Juni 2003

Tag der mündlichen Prüfung:

16. Januar 2004

Vorwort zur Online-Ausgabe

Die vorliegende Dissertation ist im Jahr 2004 im VDI-Verlag als Fortschritt-Bericht Nr. 371 in der Reihe 9 Elektronik, ISBN 3-18-337109-X, ISSN 0178-9422 erschienen. Die Auflage ist nun ausverkauft und nicht mehr lieferbar. Daher habe ich mich entschlossen, das Originaldokument jetzt unter **urn:nbn:de:gbv:830-tubdok-8105** als Online-Ausgabe zur Verfügung zu stellen.

München, im November 2009

Vorwort

Der Präsident der Technischen Universität Hamburg-Harburg, Professor Nedeß erklärte in seiner Einführung anlässlich einer Preisverleihung im Jahre 2003, daß hierzu auch die Eltern des Preisträgers eingeladen wurden, da diese „dafür Sorge getragen haben, daß derjenige überhaupt existiert und durch ihre stete Unterstützung seinen weiteren Weg bis zu diesem Tage ermöglicht haben“. Diesen Worten möchte ich mich anschließen und in diesem Sinne an erster Stelle meinen Eltern für die Entstehung des vorliegenden Werks danken.

Mein Betreuer Professor Krautschneider hat mir die Sicherheit seiner steten Solidarität gegeben und durch seine uneingeschränkte Unterstützung und Weitsicht alle meine Projekte vorangetrieben. Dies ist mehr, als der durchschnittliche Doktorand erwarten kann.

Mit fachlichem Rat und genialen Eingebungen stand mir in vielen, auch schwierigen Diskussionen Dietmar Schröder zur Seite. Ben Sell hat durch die Herstellung der von mir verwendeten Proben den experimentellen Teil dieser Arbeit überhaupt erst ermöglicht und durch gemeinsame Tage im Labor und den kontinuierlichen Gedankenaustausch die Lösung vieler Probleme beflügelt, dies gilt ebenfalls für seinen „Nachfolger“ Stefan Jakschik.

Professor Bauhofer und den Mitarbeitern seines Arbeitsbereichs möchte ich für die Quasi-Adoption in ihren Reihen danken, ebenso gilt mein Dank den Kollegen und Mitarbeitern meines eigenen Arbeitsbereichs Mikroelektronik an der Technischen Universität Hamburg-Harburg.

Prólogo

Aunque la mayor parte de mi formación, mi personalidad y mis conocimientos se los debo al gran país de Alemania y su pueblo admirable, en puntos decisivos de mi vida han intervenido personajes de mi estirpe iberoamericana, Raquel Mancera y Enrique Miranda notablemente, para mencionar a los más importantes en el plan personal y profesional. A ellos dos en especial, y a todos los otros que quedan sin ser nombrados les quiero agradecer de todo corazón por enriquecer mi vida, por enseñarme, y hacer de mi lo que soy y lo que seré.

Frankfurt am Main, Enero de 2004

Pero tú eres la reina

A Laura

Inhaltsverzeichnis

Verzeichnis verwendeter Formelzeichen	VII
Abstract	XI
1 Einleitung	1
2 Meßverfahren	5
2.1 Ladungspumpenverfahren	6
2.1.1 Zeitaufgelöste Analysen	8
2.1.2 Untersuchungen im Frequenzbereich	10
2.1.3 Ladungspumpstrommessungen bei Strukturen mit Tunneloxiden . .	13
2.1.4 Fazit	15
2.2 Ladungsbasierte Kapazitäts-Spannungs-Messung	15
2.3 Stromverstärker	17
2.3.1 Prinzip und Aufbau	17
2.3.2 Realisierung	22
2.3.3 Anwendung	25
3 Kennlinien nach Oxiddurchbruch	27
3.1 Durchbruchsuntersuchungen an MOS-Kapazitäten	29
3.1.1 Harter Durchbruch	30
3.1.2 Weicher Durchbruch	39
3.1.3 Fazit aus den Messungen an MOS-Kapazitäten	48
3.2 Durchbrüche an MOS-Transistoren	50
3.2.1 Harter Durchbruch	50
3.2.2 Weicher Durchbruch	56
3.3 Fazit	59

4	Random Telegraph Signals und Rauschen	60
4.1	Einführung in die Random Telegraph Signals	62
4.1.1	Grundlagen von RTS im Drainstrom von MOSFETs	65
4.1.2	Überblick zu RTS in MOS-Tunneldioden	70
4.2	Beobachtung von RTS in MOSFETs mit durchgebrochenen Gateoxiden . .	71
4.2.1	Lokalisierung des Durchbruchs bei MOS-Transistoren	72
4.2.2	Messung der Amplituden und Zeitkonstanten	73
4.3	Modellentwicklung	75
4.3.1	Herleitung aus der Multiphonontheorie	76
4.3.2	Modellimplementierung und Ergebnisse	80
4.3.3	Amplituden	82
4.4	Zusammenfassung	83
5	Oxiddurchbrüche in Schaltungen	84
5.1	Analoge Verstärker	86
5.1.1	Differenzverstärker	86
5.1.2	Stromverstärker	89
5.2	Digitale Gatter	90
5.3	Fazit	93
6	Zusammenfassung und Ausblick	94
Anhang		
A	Näherung für die LambertW(x) Funktion	96
B	MOS-Transistormodell in Spectre HDL	97
Literaturverzeichnis		100

Verzeichnis verwendeter Formelzeichen

Symbol	Definition
A	Querschnittsfläche
A_G	Gatefläche
A_S	Durchbruchsquerschnitt
c_n	Einfangfaktor
C	Kapazität
D	Diffusionskonstante
$D(E)$	Transmissionswahrscheinlichkeit
D_{it}	Grenzflächenzustandsdichte
e	Eulersche Zahl
e_n	Emissionsfaktor
E	Energie
E_0	Maximale Barrierenenergie
E_B	Einfangbarriere
E_C	Energieniveau der Leitungsbandkante
E_F	Fermi-Niveau
E_{FC}	Quasi-Fermi-Energieniveau auf Kathodenseite
E_g	Bandlücke zwischen Leitungs- und Valenzband
E_R	Gitterrelaxationsenergie
E_t	Energie in transversaler Richtung
E_T	Energieniveau der Störstelle
f	Frequenz
f_T	Störstellenbesetzungsfaktor
$F(E)$	Fermi-Dirac-Verteilung
g_d	Drainleitwert

Symbol	Definition
g_m	Steilheit
h	Plancksches Wirkungsquantum
\hbar	$= h/2\pi$
I	Strom
I_B	Substratstrom
I_{CP}	Ladungspumpstrom
I_D	Drainstrom
I_F	Strom in Durchlaßrichtung des pn-Übergangs
I_G	Gatestrom
I_S	Sourcestrom
I_{SD}	Drain+Source-Strom
J	Stromdichte
J_{GR}	Generations-Rekombinations-Stromdichte
J_S	Diffusions-Stromdichte
k	Boltzmannkonstante
K	Transistorkennlinienkonstante
L	Kanallänge
m	Ruhemasse des Elektrons
n	Ladungsträgerdichte
n_D	Diodenidealitätsfaktor
n_i	intrinsische Ladungsträgerdichte
n_T	Störstellendichte
n_T^-	Dichte besetzter Störstellen
n_T^0	Dichte unbesetzter Störstellen
N	Anzahl freier Ladungsträger
N_C	Effektive Zustandsdichte im Leitungsband
N_{Poly}	Gatedotierung
$N_{Substrat}$	Substratdotierung
N_T	Anzahl der Störstellen
P_{cap}	Einfangwahrscheinlichkeit
q	Elementarladung
Q_{pc}	gepumpte Ladung pro Zyklus
R_C	Einfangrate

Symbol	Definition
R_{DB}	Widerstand des Durchbruchspfad
R_E	Emissionsrate
S	Spektrale Rauschleistungsdichte
S_{HR}	Huang-Rhys-Koppelkonstante
t	Zeit
t_B	Barrierenbreite
t_{ox}	Gateoxiddicke
T	Temperatur
v_{th}	thermische Geschwindigkeit der Ladungsträger
V_{AC}	Amplitude der sinusförmigen Ladungspumpspannung
V_B	Substrat-Potential
V_{Basis}	DC-Offsetspannung beim Ladungspumpen
V_D	Drain-Potential
V_{DD}	Positive Versorgungsspannung
V_{DS}	Drain-Source Spannung
V_G	Gate-Potential
V_{GD}	Gate-Drain Spannung
V_{GS}	Gate-Source Spannung
V_{ox}	Oxidspannung
V_S	Source-Potential
V_{SS}	Negative Versorgungsspannung
V_t	Temperaturspannung
V_{TH}	Schwellspannung
V_{Trapez}	Amplitude der trapezförmigen Ladungspumpspannung
VBP	Verstärkungsbandbreiteprodukt
W	Kanalweite
W_{RLZ}	Weite der Raumladungszone
z	Entfernung der Störstelle zur Grenzfläche
z_B	Entfernung der Barriere zur Grenzfläche
α	Krümmungsparameter der Barriere
α_s	Streufaktor
β	$= 1 - z_B/t_{ox}$

Symbol	Definition
γ	Temperaturkoeffizient der Barrierenhöhe
λ	Kanallängenmodulation
μ	Beweglichkeit
σ	Einfangwirkungsquerschnitt
τ_c	Einfangzeitkonstante
τ_e	Emissionszeitkonstante
ϕ_B	Volumenpotential des Substrats
ϕ_{MS}	Austrittsarbeitsdifferenz zwischen Gate und Substrat
ϕ_S	Oberflächenpotential des Substrats
Φ	Barrierenhöhe
ψ_S	Bandverbiegung an der Substratoberfläche
ω	Phononenfrequenz

Abstract

Charakterisierung von MOS-Transistoren vor und nach Gateoxiddurchbruch

Untersuchungen zum Verhalten von MOS-Transistoren vor und nach Gateoxiddurchbruch werden präsentiert. Dabei werden unterschiedliche Durchbrucharten klassifiziert und Modelle für die DC-Kennlinien sowie für das Rauschen (Random Telegraph Signals) nach Durchbruch entwickelt. Mit Hilfe der Modelle wird abschließend die Funktion ausgewählter analoger und digitaler Schaltungen analysiert. Es zeigt sich, daß je nach Anwendung und Durchbruchstyp in Schaltungen auch nach Durchbruch eines Transistors noch Zuverlässigkeitsreserven vorhanden sein können.

Characterization of MOS transistors before and after gate oxide breakdown

The characteristics of MOS transistors before and after gate oxide breakdown are studied. Different categories of breakdown are identified and models for the dc characteristics as well as the noise (random telegraph signals) after breakdown are developed. The models are implemented to analyze the effect of breakdown on different analog and digital circuits. It is shown that, according to the application and breakdown type, circuits may be functional after breakdown of a transistor.

1

Einleitung

Viele Gebrauchsgegenstände der industrialisierten Welt haben sich im letzten Jahrhundert rasant verändert. Eine wachsende Nachfrage nach mehr Komfort bei niedrigeren Kosten führen zum Beispiel dazu, daß vieles, was früher Sonderausstattung von Fahrzeugen der Oberklasse war, wie elektrische Fensterheber, Navigationssystem oder Airbag, heute schon in der Mittelklasse zur Standardaustattung gehören. Mobiltelefone, mit denen ursprünglich der Luxus der mobilen Kommunikation zu hohen Preisen ermöglicht wurde, sind heute extrem leistungsfähige Minicomputer geworden, die viele zusätzliche Funktionen wie Versand von Bildern, Agenda, Timer und Spiele enthalten. Ganz zu schweigen von der imposant angewachsenen Rechnerleistung, die einen heute handelsüblichen Personalcomputer in Dimensionen aufsteigen läßt, die noch vor kurzer Zeit ausschließlich wissenschaftlichen oder militärischen Hochleistungsrechnern vorbehalten waren. Aus unserer heutigen Welt ist die Mikroelektronik nicht mehr wegzudenken. Der stetig ansteigende Durst nach Information, ihrer Verarbeitung und Speicherung treibt in einer Spirale gleichzeitig die Möglichkeiten immer weiter voran und schafft sich somit wieder neue Nachfrage, weil schon morgen möglich ist, was gestern noch undenkbar schien.

Die fortschreitende Verbreitung der Digitaltechnik in allen Anwendungsbereichen, sei es Telefon, Compact Disc oder Photographie, hat eindeutig zu dieser rasanten Entwicklung beigetragen, die ganz stark die Handschrift der CMOS (Complementary Metal Oxide Semiconductor) Technologie trägt. In dieser Technologie hergestellte Bauelemente werden in ihrer Funktion durch Miniaturisierung nach bestimmten Skalierregeln nicht beeinträchtigt, im Gegenteil, eine solche Verkleinerung der Strukturen führt zu einer Verbesserung der meisten ihrer elektrischen Eigenschaften. Aus diesem Grund kann die Integration der Schaltkreise fortwährend erhöht und somit auf der gleichen Fläche immer mehr Rechenleistung

untergebracht werden. Auch für analoge Anwendungen ist die CMOS-Technologie inzwischen immer besser geeignet, so daß sich ihr marktbeherrschender Status weiter verfestigt. Durch unermüdliche Forschungsanstrengungen von Wissenschaftlern und Entwicklern auf der ganzen Welt erfolgt ungefähr alle 18 Monate eine Verdopplung der Packungsdichte von dynamischen Speichern als dem Zugpferd der Integration, so daß nun Strukturgrößen unter 100nm für die minimale Kanallänge und unter 3nm für die Oxiddicke erreicht sind. Die gleichbleibende Qualität und Zuverlässigkeit dieser winzigen Bauelemente zu gewährleisten ist daher eine stete Herausforderung. Der Übergang zur Nanotechnologie geht dabei mit dem Auftreten neuer physikalischer Effekte einher. Im Rahmen dieser Arbeit werden dabei besonders die in Bild 1.1 gezeigten Aspekte behandelt. Dies sind vor allem quantenmechanische Effekte und insbesondere der Tunneleffekt, der zum Stromfluß

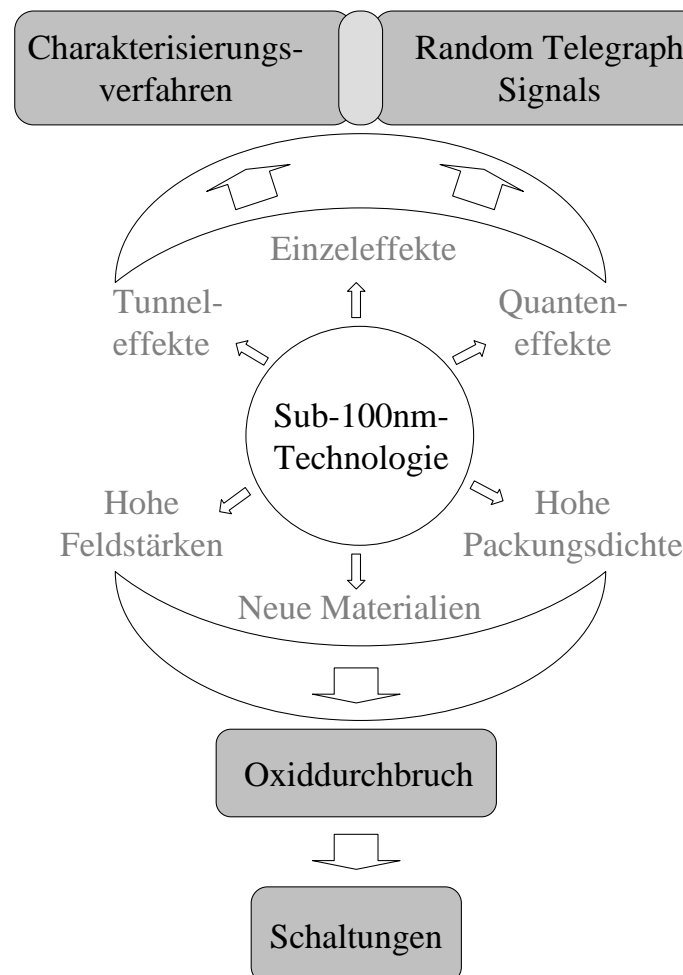


Bild 1.1: Auswirkungen der fortschreitenden Skalierung der Bauelemente, grau unterlegt sind die in dieser Arbeit behandelten Aspekte.

durch die nur noch wenige Atomlagen dicken Isolatorschichten führt. Die Entwicklung wird über kurz oder lang zu einem Umdenken in der Betrachtung der MOS-Technologie führen, denn die Transistoren werden nicht mehr im klassischen Sinne als stromlos steuerbare Elemente betrachtet werden können. Ein weiterer Aspekt der Miniaturisierung ist, daß die Eigenschaften des Bauelements von nur noch wenigen Atomen und mobilen Ladungsträgern bestimmt werden. Die bisherige Betrachtungsweise der physikalischen Vorgänge mit gleichmäßigen Dichte- und Verteilungsfunktionen muß also zunehmend durch diskrete Einzeleffekte zumindest ergänzt werden. Es gilt daher, sowohl die Auswirkung der neuen physikalischen Effekte auf klassische Meßverfahren zu untersuchen, als auch die neuen Effekte gezielt zur Etablierung neuartiger Verfahren zu verwenden.

Der zweite hier behandelte Themenbereich ist die zunehmende Belastung der MOS-Bauelemente mit fortschreitender Integration. Mit jeder neuen Technologiegeneration werden die dünnen Isolatorschichten einem höheren elektrischen Feld ausgesetzt. Andererseits nimmt auch die Packungsdichte stetig zu, so daß eine enorme Anzahl von Transistoren, bald 1 Milliarde, auf einem Chip Platz finden werden. Gleichzeitig kommen vermehrt alternative Materialien mit meist geringerer Zuverlässigkeit zum Einsatz. Unter diesen Umständen erscheint ein Gateoxiddurchbruch während des Betriebs immer wahrscheinlicher, so daß es erforderlich wird, die Auswirkungen eines solchen Durchbruchs auf ganze Schaltungen zu kennen. In diesem Sinne ist es notwendig, konkrete Modelle über die Eigenschaften eines Bauelements mit durchgebrochenem Isolator zur Verfügung zu haben, die die Entwicklung von fehlertoleranten Schaltungen und deren Überprüfung durch den Designer ermöglichen.

Bild 1.2 zeigt die einzelnen Aspekte der hierzu im Rahmen dieser Arbeit durchgeführten Untersuchungen und deren Verzahnung. Als Charakterisierungsverfahren kommen unterschiedliche Verfahren in Betracht, Strom-Spannungs-Kennlinien, das Ladungspumpenverfahren, Messung der Kapazitäts-Spannungs-Kurven und das Rauschen, dessen Sonderform, sogenannte Random Telegraph Signals, hier behandelt wird. Kapitel 2 widmet sich zunächst klassischen Meßverfahren und der Auswirkung der Skalierung auf diese Verfahren. Es wird ein Stromverstärkerkonzept vorgestellt, mit dem die zu messenden Defektströme bereits auf dem Chip verstärkt werden können. In den darauf folgenden Kapiteln steht die Entwicklung physikalischer Modelle für die Eigenschaften des MOS-Transistors nach Gateoxiddurchbruch im Vordergrund. Kapitel 3 beschäftigt sich eingehend mit der Beschreibung der Transportvorgänge nach einem Gateoxiddurchbruch. Es werden unterschiedliche Arten des Durchbruchs aufgezeigt und eine Klassifizierung anhand der elektrischen Kennlinien durchgeführt. Kapitel 4 beschäftigt sich mit sogenannten Random Telegraph Signals.

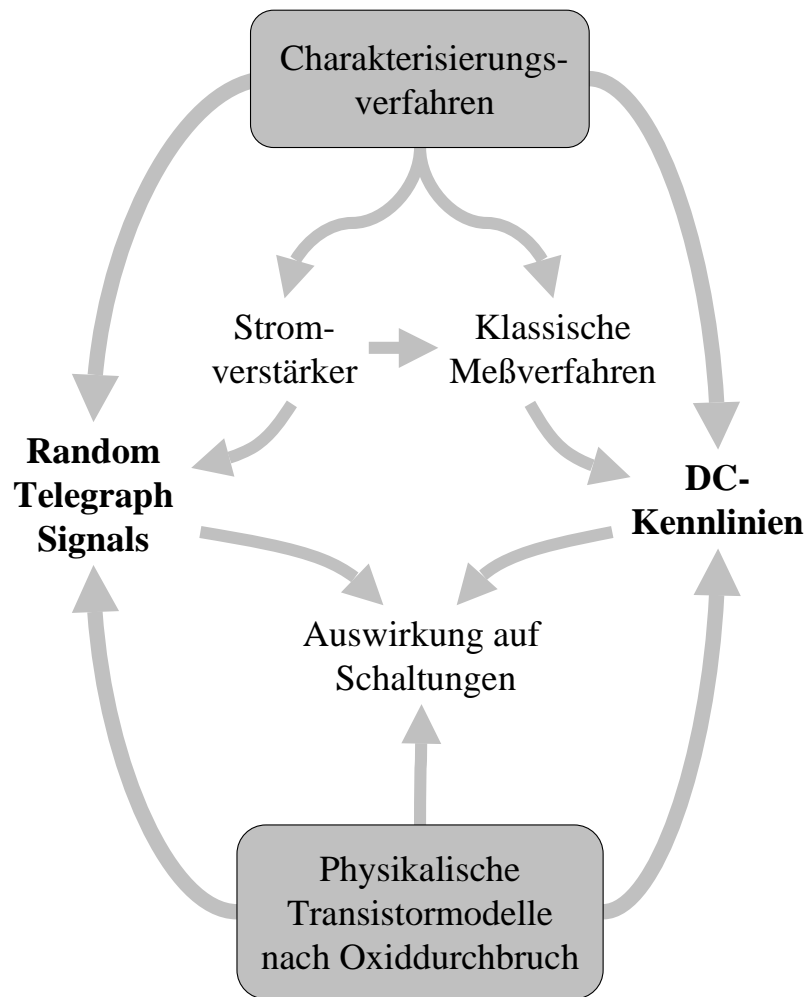


Bild 1.2: Hauptaspekte der vorliegenden Arbeit

Dies ist eine nach Durchbruch und bei Transistoren mit geringen Abmessungen häufig auftretende Rauschform, bei der Einzelladungseffekte besonders deutlich sichtbar werden. In Kapitel 5 werden die so entwickelten Modelle in ausgewählte digitale und analoge Schaltungen eingesetzt. Zusammen ergibt sich ein Bild von den neuen Herausforderungen, die sich mit der fortschreitenden Skalierung in den Nanometerbereich ergeben.

2

Meßverfahren

Qualitätskontrolle und -sicherung sind elementare Bestandteile der Fertigung jeglicher Art von Produkten. In der Halbleiterindustrie wird eine Lebensdauer elektronischer Schaltkreise von mindestens zehn Jahren angestrebt. Hierfür muß das Gesamtsystem, also Halbleiterbauelemente, Verdrahtungsebenen und Verpackung betrachtet werden. Wird der Fokus auf die Ebene des MOS-Transistors gerichtet, kann die Frage nach seiner Qualität, abgesehen von den eigentlichen Strom-Spannungskennlinien des Bauelements, auf die Reinheit der verwendeten Materialien, also Silizium und Siliziumdioxid, und den Übergängen zwischen diesen Materialien bezogen werden. Hier sollen einerseits Verunreinigungen jenseits der gewollten Einbringung von Dotieratomen vermieden werden und andererseits die Regelmäßigkeit der Kristallstrukturen und deren Übergänge möglichst perfekt sein. Kristallversetzungen und Verunreinigungen führen zu einer Störung der Bandstruktur und zu besetzbaren Energiezuständen, Störstellen, im verbotenen Band des Halbleiters und Isolators. Die Ermittlung der Qualität läuft daher auf eine Bestimmung der Störstellendichte hinaus. Hierfür sind geeignete Meßverfahren notwendig, die mehreren, sich teilweise widerlaufenden Interessen genügen müssen. So sollte eine Charakterisierungsmethode einfach und robust zu handhaben sein, um auch im industriellen Umfeld reproduzierbare Ergebnisse zu liefern. Andererseits sollte sie möglichst genaue Aussagen über die Qualität der untersuchten Probe erlauben. Weiterhin muß sie schnell und nicht zerstörend anwendbar sein. Die genannten Kriterien werden am Besten von elektrischen Methoden erfüllt, von denen einige sehr ausgereift sind und seit langer Zeit verwendet werden. Genannt seien hier Ladungspumpenverfahren, Kapazitäts-Spannungs-Messung, Leitwertmethode und Deep Level Transient Spectroscopy. Mit fortschreitender Skalierung der Bauelemente ergeben sich jedoch in zunehmendem Maße Schwierigkeiten bei der Anwendung dieser Testver-

fahren. Da die Meßmethoden auf der Auswertung der durch die Störstellen hervorgerufenen Defektströme beruhen, sind für die sichere Analyse Mindestabmessungen der Teststrukturen notwendig, die nicht skalierbar sind. Somit ergeben sich gemittelte Aussagen, die bei Sub-100nm-Bauelementen nur noch bedingt zutreffen. Die im Zuge der Qualitätsverbesserung der Prozeßtechnologie abnehmenden Defektdichten in modernen Bauelementen rücken zudem immer näher an die Auflösungsgrenze der Verfahren heran. Auch bei etablierten Meßmethoden ist es also sinnvoll, an ihrer Verbesserung zu forschen. Eine wichtige weitere Veränderung, die durch die Skalierung der Gateoxidstärke eintritt, ist die Zunahme des direkten Tunnelstroms durch das Oxid, der sich nicht nur den Kennlinien des Bauelements, sondern auch den gemessenen Defektströmen überlagert. Dies ist eine ernst zu nehmende Gefahr für die weitere Anwendung der klassischen Meßverfahren.

In diesem Kapitel werden die hierzu erhaltenen Forschungsergebnisse am Ladungspumpenverfahren und der Kapazitäts-Spannungsmessung, sowie ein Stromverstärkerkonzept zur direkten Verstärkung der Defektströme vorgestellt.

2.1 Ladungspumpenverfahren

Das Ladungspumpenverfahren ist eines der etabliertesten Verfahren, die zur Bestimmung der Grenzflächenzustandsdichte von MOS-Transistoren verwendet werden. Es beruht darauf, daß über grenzflächennahe Haftstellen ein Nettostrom zwischen den Source-/Drain-Gebieten und dem Substrat eines MOS-Transistors fließt, wenn an das Gate eine trapezförmige Spannung angelegt wird, durch die der Transistor zwischen Akkumulation und Inversion hin- und hergeschaltet wird [1, 2]. Bild 2.1 zeigt schematisch den Meßaufbau. Der Ladungspumpstrom ist abhängig von der Gatefläche A_G , der mittleren Grenzflächenzustandsdichte $\overline{D_{it}}$, der Frequenz f und der Flankensteilheit des angelegten Signals. Werden Anstiegs- und Fallzeit der Trapezspannung t_r und t_f mit einem festen Faktor α_F stets proportional zur Periodendauer gehalten, d. h. $t_r = t_f = \alpha_F/f$, dann berechnet sich die gepumpte Ladung pro Zyklus Q_{pc} nach der Beziehung [3]

$$\frac{dQ_{pc}}{d \log f} = A_G \cdot \frac{2qkT \cdot \overline{D_{it}}}{\log e}, \quad (2.1)$$

wobei q die Elementarladung, k die Boltzmannkonstante, T die Temperatur und e die Eulersche Zahl sind. In Bild 2.2 sind gemessene Ladungspumpströme an einer ungestreßten Struktur dargestellt, die mit einer von B. Sell entwickelten Matlab-Routine ausgewertet

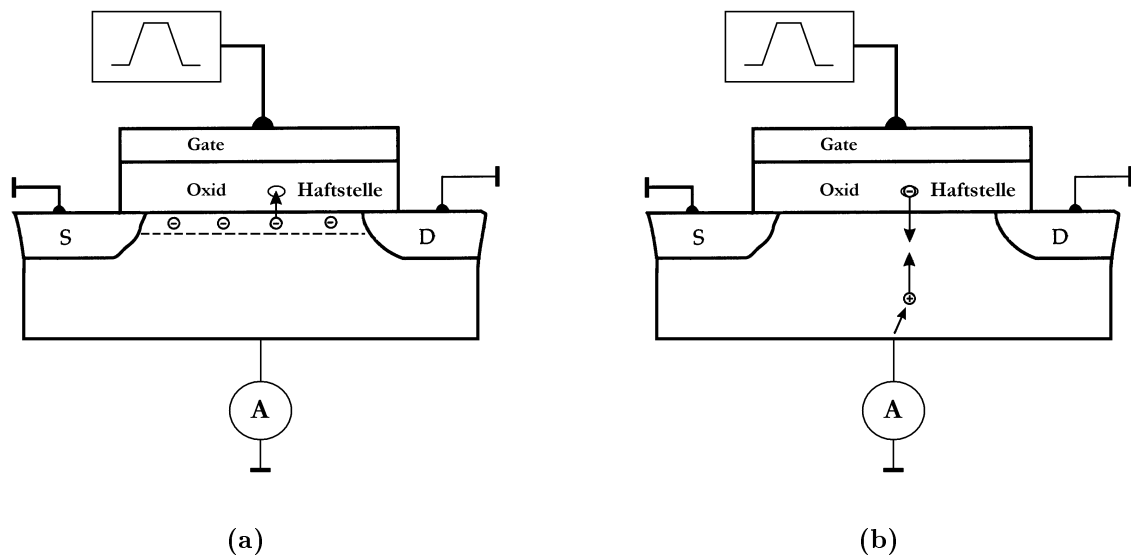


Bild 2.1: Prinzip des Ladungspumpenverfahrens. Der nMOS-Transistor befindet sich (a) in Inversion und (b) in Akkumulation.

wurden. Die Offsetspannung V_{Basis} liegt dabei in der Mitte des Trapezsignals mit der Amplitude V_{Trapez} . Aus der Steigung der Pumpladung über dem Logarithmus der Frequenz ergibt sich die mittlere Grenzflächenzustandsdichte $\overline{D_{it}}$, in diesem Fall der bei heutigen Prozessen übliche Wert von $\sim 10^{10} \text{ eV}^{-1} \text{ cm}^{-2}$.

Wegen der Flächenabhängigkeit des Ladungspumpstroms ergeben sich für den zu charakterisierenden Transistor bestimmte Mindestabmessungen, um die Grenzflächenzustands-

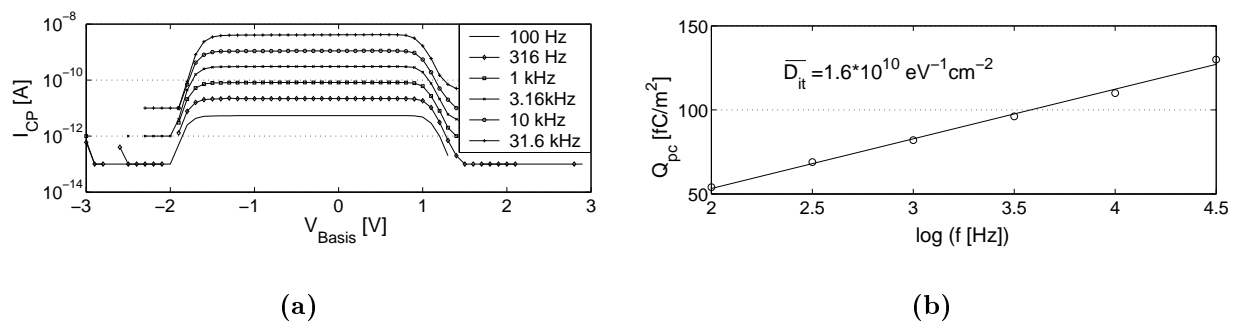


Bild 2.2: (a) Gemessene Substratströme I_{CP} eines p-Kanal-Transistors (6Capp1) in $0.35 \mu\text{m}$ AMS-Technologie mit Gatelänge $L=10 \mu\text{m}$, Gateweite $W=1000 \mu\text{m}$, $V_S = V_D = V_B = 0 \text{ V}$ bei unterschiedlichen Frequenzen ($t_r = t_f = 0.2/f$, $V_{Trapez} = 3 \text{ V}$), (b) Berechnete Pumpladung pro Zyklus Q_{pc} .

dichte verlässlich extrahieren zu können. Dies ist zwar relativ unkritisch, da $\overline{D_{it}}$ bei einer Fläche von einigen μm^2 noch ermittelt werden kann, trotzdem ist es lohnenswert, über Verbesserungen des Verfahrens nachzudenken. Von Haddara wurde vorgeschlagen, die Grenzflächenzustandsdichte durch Anwendung des Ladungspumpenverfahrens bei Anlegen einer Drain-Source-Spannung V_{DS} und Messung des Sourcestroms zu bestimmen [4, 5]. Dieser Ansatz wurde im Rahmen dieser Arbeit weiter verfolgt, die Ergebnisse von Haddara konnten aber nicht bestätigt werden.

Um über mögliche Konzepte zur Eliminierung des parasitären Tunnelstroms durch das Gateoxid Aufschluß zu erhalten, wurden zeitaufgelöste Untersuchungen durchgeführt. Diese werden zunächst präsentiert, bevor die in Anlehnung an Haddara und die an Strukturen mit Tunneloxiden durchgeführten Experimente dargestellt werden.

2.1.1 Zeitaufgelöste Analysen

Zum Verständnis des Ladungspumpenverfahrens trägt eine Betrachtung im Zeitbereich bei [6, 7]. Die Bilder 2.3(a) und (b) zeigen die erhaltenen Verläufe durch Simulation mit MINIMOS 6.1 an einem nMOS-Transistor und Messung an einem pMOS-Transistor. Simulation und Messung stimmen gut überein. Die Ströme sind Verschiebestrome der MOS-

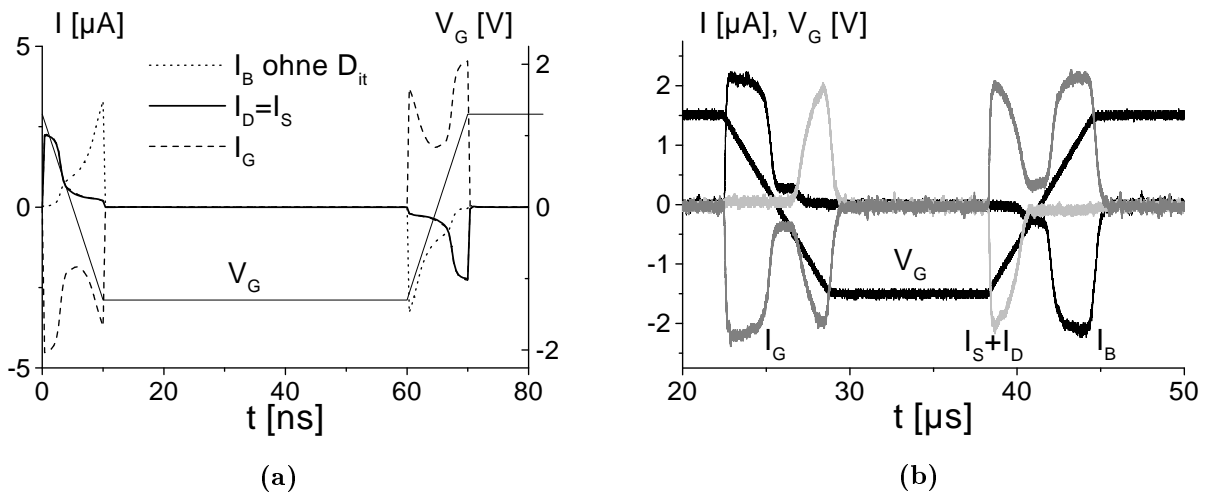


Bild 2.3: (a) Mit MINIMOS 6.1 simulierte Ströme eines nMOS-Transistors ($L=0.35\mu\text{m}$, $W=10\mu\text{m}$, $t_{ox} = 7.5\text{nm}$) beim Ladungspumpen ohne Grenzflächenzustände, (b) Gemessene Source/Drain-, Gate- und Substratströme am ungestreßten pMOS-Transistor 6Capp1 mit $L=10\mu\text{m}$, $W=1000\mu\text{m}$. Messung mit einer I-U-Wandlerschaltung der Verstärkung 10^6 und einem LeCroy LC584AM Oszilloskop, s. Abschnitt 4.2. I_G wurde in einem zweiten Schritt nach Anlegen der Pumpspannung an Source/Drain und Substrat aufgenommen.

Kapazität, an der Gateelektrode wird stets die Spiegelladung von der Ladung im Kanal bereitgestellt. Aus diesem Grund ist der Verlauf von Bulk- und Source-/Drainströmen beim p-Kanal-Transistor gegenüber dem n-Kanal-Transistor vertauscht. Am gemessenen Transistor konnte nach Erhöhung der Grenzflächenzustandsdichte durch homogene Injektion auf $5 \cdot 10^{10} \text{eV}^{-1} \text{cm}^{-2}$ kein Einfluß auf die Form der Meßkurven festgestellt werden.

Durch Vergleich der Simulation mit und ohne Grenzflächenzustände ist deren Auswirkung einfacher quantifizierbar. Aus Bild 2.4 kann der Ablauf beim Ladungspumpen abgeleitet werden. Zu Beginn des Zyklus bei $t = 0 \text{s}$ befindet sich der nMOS-Transistor

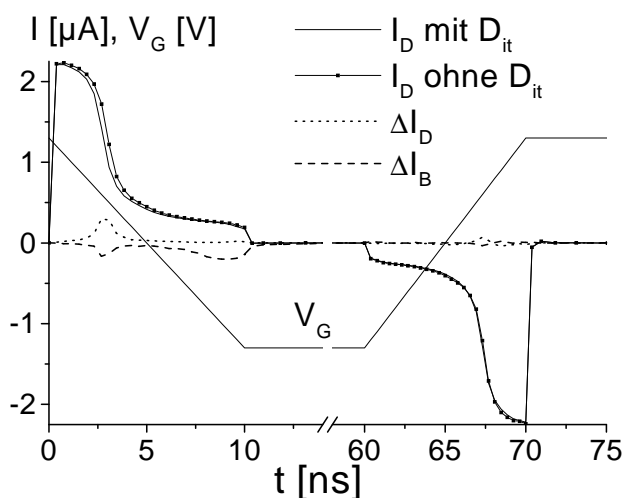


Bild 2.4: Vergleich der Ströme bei Simulation mit und ohne Akzeptorstörstellen (energetisch gaußverteilt, $\text{AUGD}=1 \cdot 10^{12} \text{eV}^{-1} \text{cm}^{-2}$, $\text{AUGS}=0.1$, $\text{AUGX}=0.5$ [8]).

in Inversion. Es kann davon ausgegangen werden, daß sämtliche Störstellen besetzt sind. Beim Übergang zur Akkumulation fließen die Minoritätsladungsträger aus dem Kanal wieder zurück in die Source-/Drain-Gebiete. Es erfolgt bei $t \approx 2.5 \text{ns}$ eine Rekombination mit den Majoritätsladungsträgern des Substrats, wahrscheinlich zum Teil auch aus den flachen Störstellen, weswegen sich ein gegenüber der Simulation ohne Störstellen niedrigerer Drain- und höherer Bulkstrom ergibt. Die tieferen Störstellen mit einer höheren Zeitkonstante rekombinieren erst später mit den Majoritätsladungsträgern des Substrats. Dies macht sich im Bulkstrom bis zur Entladung der langsamsten Störstelle bei $t \approx 15 \text{ns}$ bemerkbar. Beim erneuten Übergang in die Inversion wird das Substrat im Kanal wieder von Majoritätsladungsträgern freigeräumt, geringe Rekombinationsströme machen sich auch hier bemerkbar. Von der Inversionsschicht aus werden nun die Störstellen wieder aufgefüllt. Es wird deutlich, daß der Ladungspumpstrom während eines Zyklus selbst bei der sehr hohen simulierten Grenzflächenzustandsdichte von maximal $1 \cdot 10^{12} \text{eV}^{-1} \text{cm}^{-2}$ um Größenordnun-

gen unter dem durch die MOS-Kapazität hervorgerufenen Verschiebestrom bleibt. Aus den Verläufen selber ist zudem eine Extraktion der Grenzflächenzustandsdichte nur schwer möglich. Daß das Ladungspumpenverfahren tatsächlich so hervorragend und empfindlich funktioniert, ist vor diesem Hintergrund bemerkenswert.

In Anlehnung an Haddara [4,5] wurde nun der Source-Strom bei $V_{DS} > 0$ aufgenommen. Die Verläufe für unterschiedliche Eingangssignale und Frequenzen sind in Bild 2.5 dargestellt. Zusätzlich zu dem beim Ladungspumpen beobachteten Verschiebe- und Pumpstrom

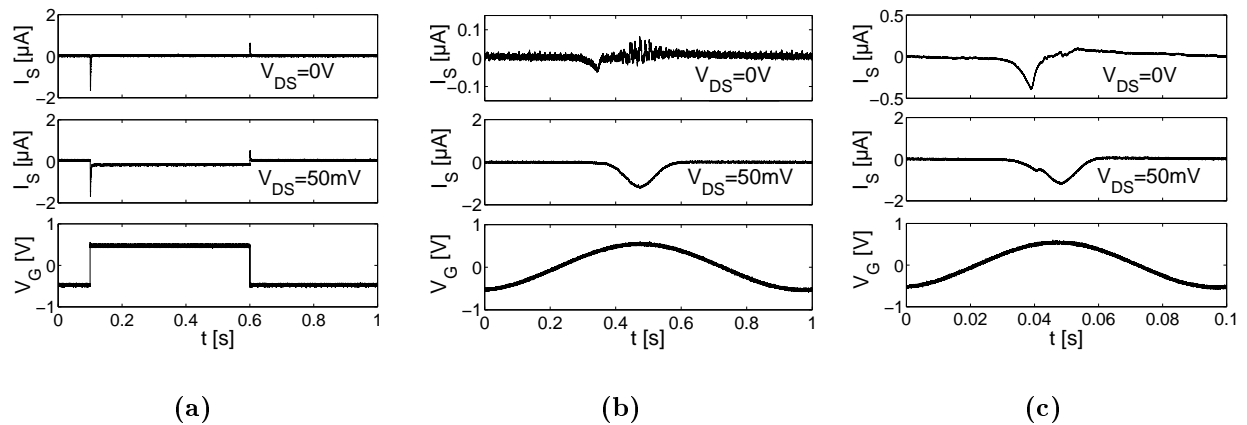


Bild 2.5: Gemessener Source-Strom eines n-Kanal-Transistors (8Capn1) mit $L=10\mu\text{m}$, $W=1000\mu\text{m}$ und $\overline{D_{it}} = 4 \cdot 10^{11} \text{eV}^{-1} \text{cm}^{-2}$ bei (a) rechteckförmiger Gatespannung mit $t_r = t_f = 40\text{ps}$ und bei sinusförmiger Gatespannung der Frequenz (b) 1kHz bzw. (c) 1 MHz.

kommt nun der Drain-Source-Stromfluß bei eingeschaltetem Transistor hinzu, der besonders bei der sinusförmigen Gatespannung schwer vom kapazitiven Strom zu trennen ist. Trotz der sehr hohen Grenzflächenzustandsdichte decken sich die Verläufe mit SPICE-Simulationsergebnissen aus Cadence Spectre von ungestreßten Strukturen. Auch hier ist also der durch die Grenzflächenzustände hervorgerufene Strom meßtechnisch schwer direkt zugänglich.

2.1.2 Untersuchungen im Frequenzbereich

Der durchschnittliche Drain-Source-Strom bei Anlegen einer Rechteckspannung an das Gate und $V_{DS} > 0$ ist theoretisch frequenzunabhängig, da hierfür das Tastverhältnis ausschlaggebend ist. Nur wenn die Frequenz in den Bereich der Anstiegszeit des Rechteckimpulses kommt, wird der durchschnittliche Strom abnehmen. Dies bestätigen SPICE-Simulationen mit Cadence Spectre. Das Anlegen eines Rechteckimpulses mit schneller Anstiegszeit ist

allerdings kritisch, da sich schnell Überschwinger ergeben. Ebenso zeigte sich, daß die verwendeten Meßgeräte, die praktisch als Integratoren fungieren und den Mittelwert bilden, unterschiedlich auf die sich ergebenden Stromformen und -spitzen reagieren. Dies liegt wahrscheinlich daran, daß durch das Hinzukommen des Drain-Source Kanalstroms im eingeschalteten Zustand sehr viel mehr Größenordnungen des Stroms durchfahren werden als beim reinen Ladungspumpen. Aus diesem Grund wurde für die weiteren Untersuchungen eine sinusförmige Gatespannung V_{AC} der Amplitude 1.6V des Frequenzgenerators HP3314A verwendet, wobei die im Folgenden als V_{Basis} bezeichnete Offsetspannung in der Mitte des Sinussignals liegt. In Bild 2.6 sind die frequenzabhängigen Transferkennlinien dargestellt. Die Mittelwertbildung bei den AC-Signalen erfolgte durch das Picoamperemeter HP4140B, mit dem auch die DC-Ströme gemessen wurden.

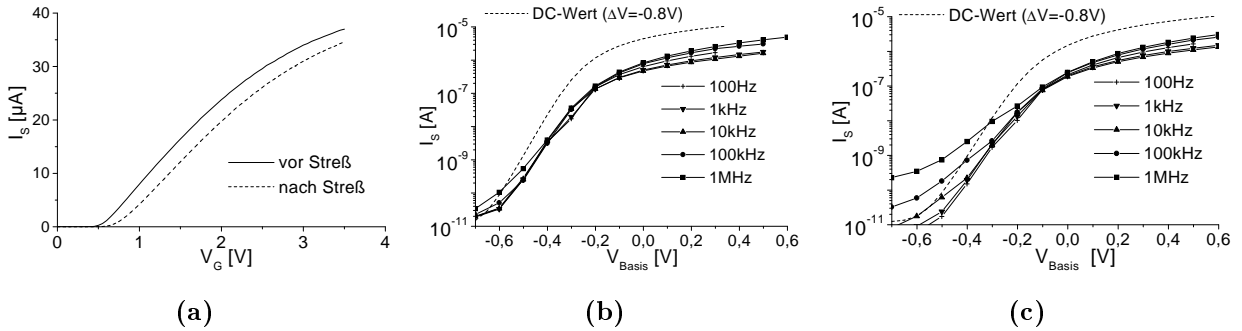


Bild 2.6: (a) DC-Transferkennlinie des n-Kanal-Transistors 8FLn4 mit $L=5\mu\text{m}$ und $W=10\mu\text{m}$ vor und nach 3minütigem Streß bei 8V ($\overline{D_{it}}$ nach Streß $4 \cdot 10^{11} \text{eV}^{-1} \text{cm}^{-2}$) und frequenzabhängige Transferkennlinien (b) vor und (c) nach Streß. Die DC-Werte sind um -0.8V verschoben.

Deutlich zu sehen ist die Schwellspannungsverschiebung durch feste Oxidladungen sowohl in der DC-Kennlinie als auch in den frequenzabhängigen Messungen. Die Differenz der AC-Kurven zur DC-Transferkennlinie in den Bildern 2.6(b) und (c) kommt hierbei durch das „Tastverhältnis“ der Wechselfspannung zustande. Eine Veränderung der AC-Kennlinien nach Streß bei negativen Basisspannungen ist ebenfalls erkennbar. Die Bilder 2.7(a) und (b) zeigen die selben Meßdaten, diesmal aufgetragen über der Frequenz, vor und nach Streß. Bei Anlegen einer rechteckförmigen Spannung ergeben sich ähnliche Verläufe. Die Veränderung der Kennlinien nach Streß läßt zunächst vermuten, daß der Drain-Source-Kanalstrom tatsächlich von der Frequenz des angelegten Wechselfsignals beeinflusst wird und somit ein Rückschluß auf umladbare Grenzflächenzustände möglich ist. Die Bilder 2.7(b) und (c) zeigen aber eindrucksvoll, daß die beobachtete Frequenzabhängigkeit nach Streß vollständig auf den „klassischen“ Ladungspumpstrom zurückgeführt werden kann, der in einer zweiten

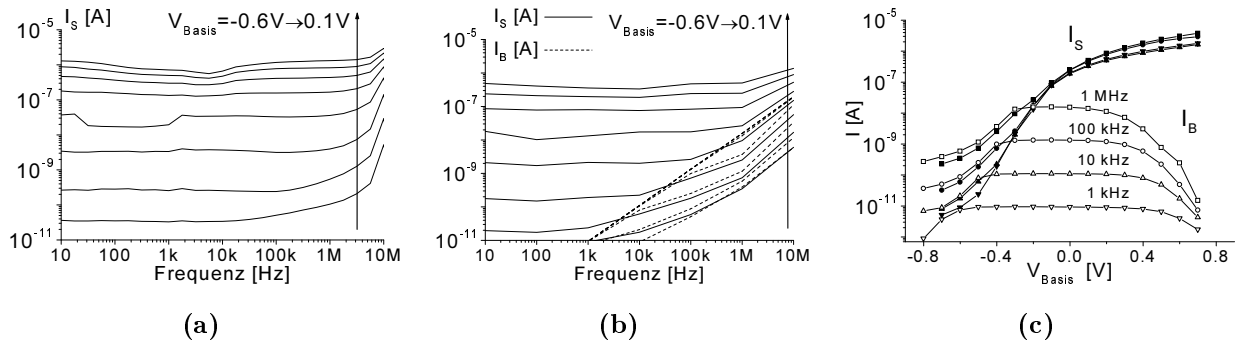


Bild 2.7: Darstellung des Source- und Bulkstroms des Transistors von Bild 2.6 über der Frequenz: (a) vor und (b) nach Streß bei unterschiedlichen Offsetspannungen V_{Basis} von -0.6 bis 0.5 V in Schritten von 0.1 V, (c) Vergleich des gemessenen Source- und Bulkstroms als Funktion der Basisspannung bei unterschiedlichen Frequenzen.

Messung unter gleichen Bedingungen am Substrat erfaßt wurde.

Es liegt nahe, die Amplitude der Wechsellspannung zu verringern, um so die Abhängigkeit der Steilheit des Transistors von der Frequenz bei umladbaren Grenzflächenzuständen zu untersuchen. Die Steilheit wird durch die Grenzflächenzustände reduziert. Allerdings haben die Ladungsträger bei hohen Frequenzen keine Zeit mehr, die Grenzflächenzustände zu füllen und wirken nicht mehr auf die Kennlinie des Transistors. Diese wird dann nur noch durch die festen Oxidladungen bestimmt. Dadurch nimmt die Steilheit bei hohen Frequenzen zu. Bild 2.8 zeigt die Differenz zwischen DC-Kennlinie vor Streß und dem AC-Mittelwert des Stroms bei Wechsellspannung nach Streß. Es zeigt sich, daß die Differenz zu hohen Frequenzen abnimmt, der Transistor also im Mittel einen höheren Strom treibt, weil die umladbaren Grenzflächenzustände keine Rolle spielen. Laut Simonne et al. macht sich bei dieser Art von Messung allerdings erst eine Grenzflächenzustandsdichte

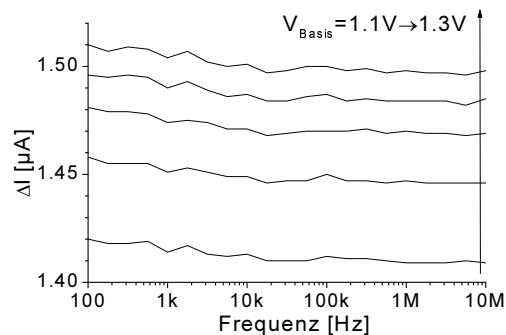


Bild 2.8: Differenz $\Delta I = I_{DC}$ (vor Streß) - I_{AC} (nach Streß) des n-Kanal-Transistors 5FLn4 mit $L=5\mu\text{m}$ und $W=10\mu\text{m}$, wobei $\overline{D_{it}}$ nach Streß $4 \cdot 10^{11} \text{eV}^{-1} \text{cm}^{-2}$, $V_{DS} = 10 \text{mV}$ und $V_{AC} = 50 \text{mV}$.

über $1 \cdot 10^{11} \text{eV}^{-1} \text{cm}^{-2}$ bemerkbar [9, S. 596-603]. Diese Aussage beruht indes auf älteren Daten, so daß möglicherweise bei heutigen Prozessen mit höheren Steilheiten ein stärkerer Einfluß zu beobachten wäre. Weitergehende Untersuchungen in diese Richtung wurden jedoch nicht angestellt.

2.1.3 Ladungspumpstrommessungen bei Strukturen mit Tunneloxiden

Der durch dünne Gateoxide fließende direkte Tunnelstrom verteilt sich, wie Bild 2.9 zeigt, aufgrund der unterschiedlichen Dotierung und Fläche nicht gleichmäßig auf Substrat und Source-/Drain-Gebiete. Die analytische Beschreibung der einzelnen Stromanteile ist noch

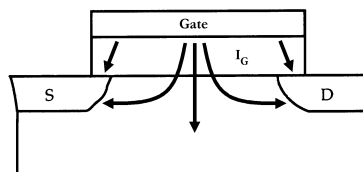


Bild 2.9: Aufteilung des Gatestroms auf Drain, Source und Substrat.

Gegenstand der Forschung [10–14]. Bei sehr dünnen Oxiden kommen Oxiddickenschwankungen und möglicherweise störstellengestütztes Tunneln hinzu, die eine weitere schwer zu kontrollierende örtliche Verteilung des Stroms nach sich ziehen könnten. Bild 2.10 zeigt beispielhaft die Aufteilung des statischen Stroms auf Substrat sowie Source-/Drain-Gebiete in einem n- und pMOS-Transistor. Es zeigt sich, daß beim nMOS-Transistor der größte Anteil des Gatestroms zu den Source-/Drain-Gebieten fließt, während er beim pMOS eher zum Substrat fließt. Bei beiden Transistortypen ist das Verhalten um $V_G = 0$ nicht symmetrisch. Dies hat natürlich Folgen für den Ladungspumpstrom, dem sich der Tunnelstrom überlagert und so die Messung verfälscht. Wird der Ladungspumpstrom sowohl an Substrat als auch an den Source-/Drain-Gebieten gemessen, so ergibt sich Bild 2.11. Es wird deutlich, daß der Ladungspumpstrom jeweils an den Elektroden mit dem auch im statischen Fall niedrigeren Tunnelstrom unverfälschtere Ergebnisse liefert, für den nMOS also im Substrat und für den pMOS im Source-/Drain-Strom. Ebenso ist die Unsymmetrie um 0V erkennbar.

Bei noch dünneren Oxiden nimmt der Tunnelstrom rapide zu und macht somit eine Messung auch an der Elektrode mit geringerem Tunnelstrom kompliziert. Bei sehr hohen Gateströmen dürfte es deshalb zunehmend schwieriger werden, das kleine Ladungspumpsignal herauszufiltern.

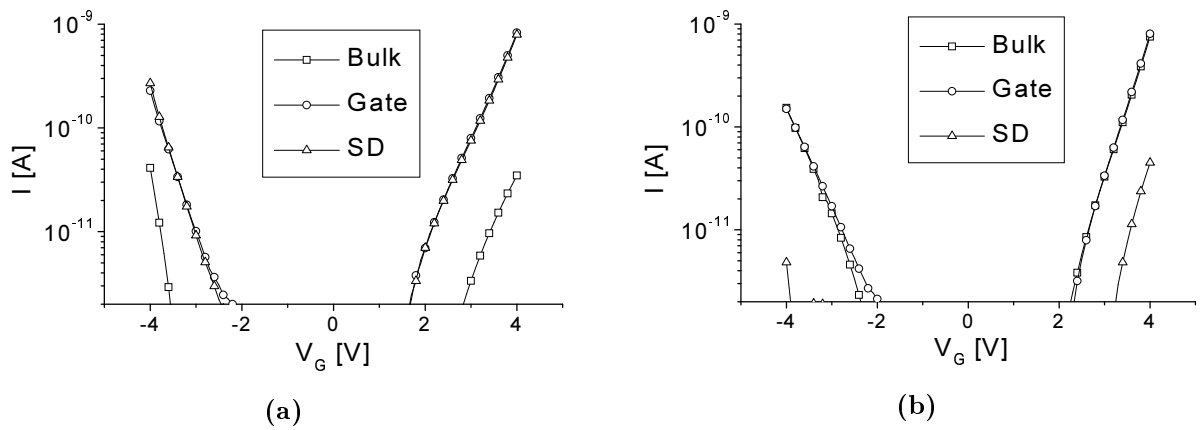


Bild 2.10: Tunnelströme in einem (a) nMOS, (b) pMOS-Transistor mit 2.7nm Oxiddicke ($L=0.5\mu\text{m}$, $W=10\mu\text{m}$, Infineon 0.25 μm dual-gate Prozeß); Source und Drain sind zusammengeschaltet (SD).

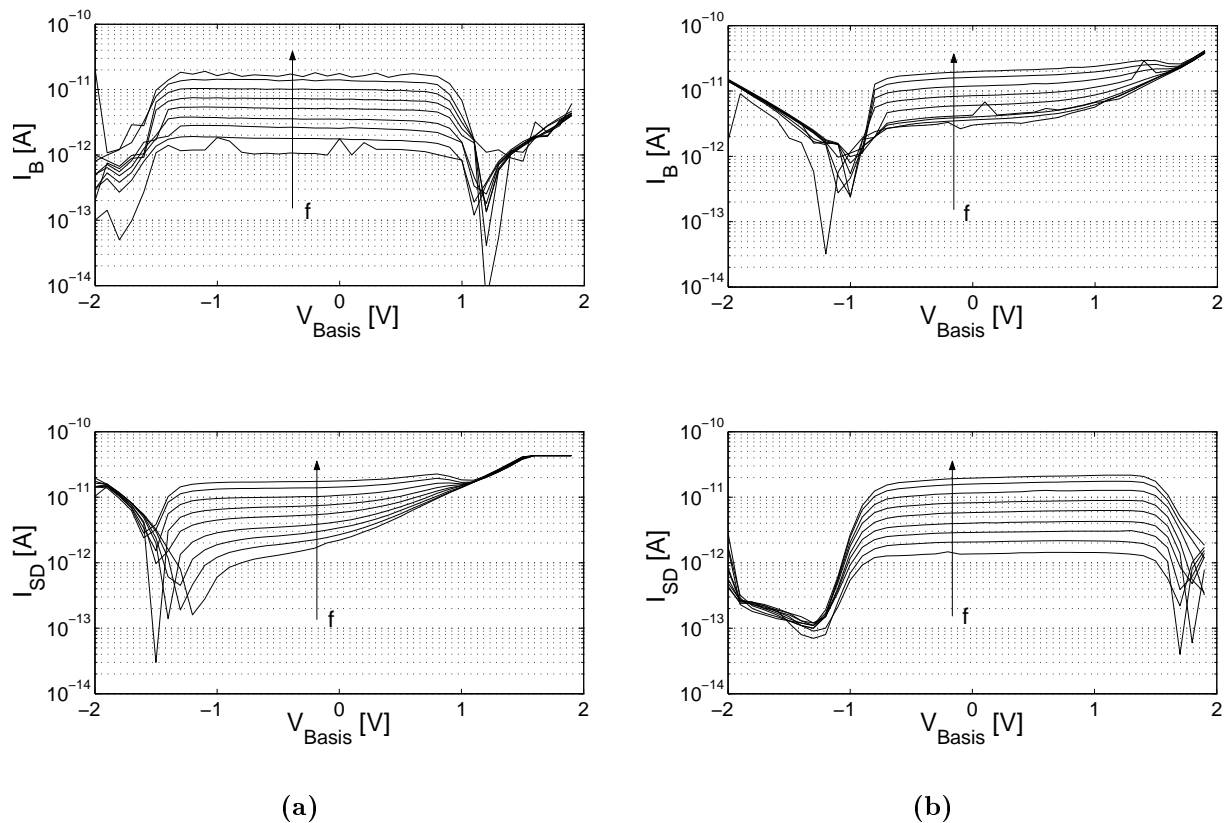


Bild 2.11: Gemessene Ladungspumpströme bei Frequenzen zwischen 10 und 100 kHz des (a) nMOS, (b) pMOS-Transistors von Bild 2.10.

2.1.4 Fazit

Es wurden Messungen zum Ladungspumpenverfahren im Zeit- und Frequenzbereich bei Strukturen mit und ohne Tunneloxiden vorgestellt. Es zeigte sich, daß der kapazitive Strom deutlich über dem durch Grenzflächenzustände hervorgerufenen Pumpstrom liegt und ihn überdeckt. Im Zeitbereich ist eine Erforschung des Ladungspumpstroms dadurch sehr schwierig. Die von Haddara vorgeschlagene Verbesserung des Verfahrens [4, 5] konnte nicht bestätigt werden. Der klassische Ladungspumpstrom kann im Substrat oder den Source-/Drain-Gebieten gemessen werden und wird durch das Anlegen einer kleinen Spannung V_{DS} praktisch nicht beeinflusst. Als alternatives Verfahren kann die Steilheit des Transistors durch Anlegen einer kleinen Wechselspannung frequenzabhängig gemessen werden. Obwohl hier möglicherweise Potential für vergleichende Analysen zwischen frischem und gestreßtem Zustand eines Bauelements steckt, ist fraglich, ob so der Absolutwert der Grenzflächenzustandsdichte ermittelt werden kann. Bei Strukturen mit Tunneloxiden wurde gezeigt, daß sich der Tunnelstrom unterschiedlich auf Source/Drain und Substrat aufteilt. Durch geeignete Wahl der Geometrie oder durch Verschaltung mehrerer Transistoren könnte dies für die Anwendung des Ladungspumpenverfahrens auch noch für sehr dünne Oxide ausgenutzt werden.

2.2 Ladungsbasierte Kapazitäts-Spannungs-Messung

Bei der Kapazitäts-Spannungs-(CV)-Messung wird die differentielle Kapazität eines Zweipols über der Spannung bestimmt. Man unterscheidet zwischen quasistatischer (QSCV), Niederfrequenz- (LFCV) und Hochfrequenz-CV-Messung (HFCV). Hiermit können die wichtigsten Parameter des MOS-Systems wie Oxiddicke, Substratdotierung, Flachband- und Schwellspannung und, durch Vergleich von LFCV- und HFCV-Kurven, auch die Grenzflächenzustandsdichte bestimmt werden [15, 16]. Aus diesem Grund ist die Messung der CV-Kurve ein sehr mächtiges Standardverfahren. Bei der QSCV-Methode wird die MOS-Kapazität durch eine langsame Spannungsrampe von Akkumulation in Inversion gefahren und dabei die Ladungsänderung gemessen, aus der die Kapazität berechnet wird. Beim LF- und HFCV-Verfahren wird der Gatespannung, die durch den Spannungsbereich gefahren wird, eine kleine Wechselspannung von einigen 10mV überlagert und die Impedanz der MOS-Struktur für jeden Spannungspunkt bestimmt. LF- und HFCV-Methode unterscheiden sich durch die Frequenz der überlagerten Wechselspannung. Typischerweise wird die HFCV-Kurve bei Frequenzen größer als 100kHz und die LFCV-Kurve unter 100Hz

gemessen. Dies führt zu einem unterschiedlichen Verlauf in Inversion. Bei reinen MOS-Kapazitäten kann die Inversionsschicht nur mit der Zeitkonstanten der Generationsrate aufgebaut werden. Bei MOS-Transistoren hingegen können durch die angrenzenden Diffusionsgebiete sehr schnell Minoritätsladungsträger für die Inversion bereitgestellt werden, so daß man bei allen Frequenzen eine „LFCV“-Kurve erhält.

Um CV-Kurven verlässlich messen zu können, sind große Flächen zur Bereitstellung eines ausreichend großen Verschiebestroms notwendig. Auch wird meist ein Lock-in-Verstärker verwendet, der den gemessenen Wechselstrom nach Betrag und Phase aufteilen kann.

Ein alternatives Verfahren zur Messung von HFCV-Kurven wurde zusammen mit B. Sell entwickelt, von dem auch die Messungen durchgeführt wurden [17, 18]. Die in einem 0.6µm-AMS-Prozess realisierte Teststruktur ist vereinfacht in Bild 2.12(a) gezeigt. Die vorgeschla-

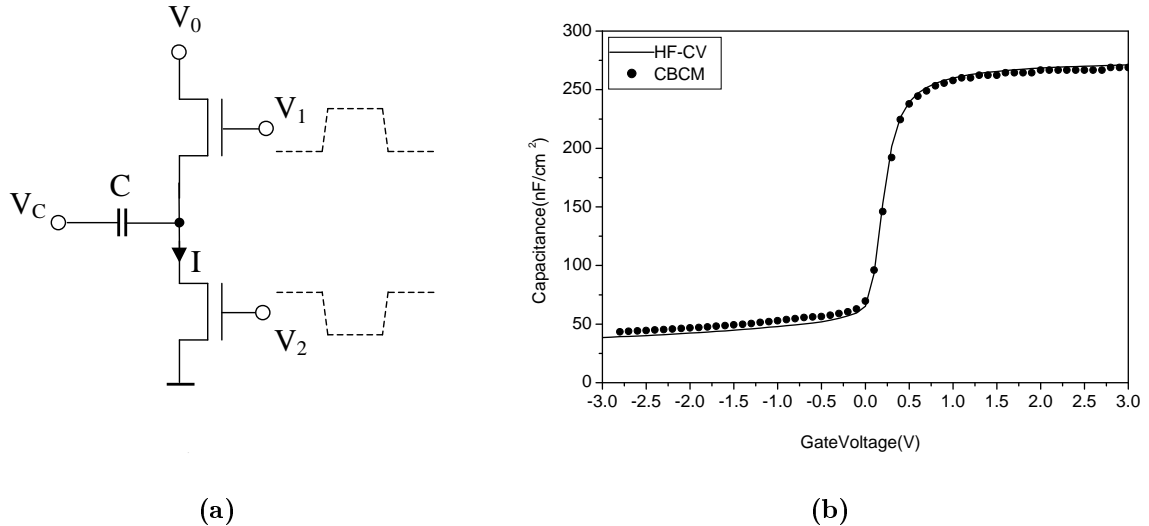


Bild 2.12: (a) Prinzip der ladungsbasierten Kapazitäts-Spannungs-Meßmethode, (b) Vergleich der Meßergebnisse mit dem HFCV-Standardmeßverfahren jeweils bei 1MHz an einer $(150\mu\text{m})^2$ pMOS-Kapazität mit einer Oxiddicke von 11.8nm [17].

gene ladungsbasierte Kapazitäts-Spannungs-Meßmethode beruht darauf, daß zwei Transistoren abwechselnd über die Spannungen V_1 und V_2 an- und ausgeschaltet werden und dabei die zu untersuchende MOS-Kapazität auf- und entladen. Die über der Kapazität abfallende effektive Gatespannung V_G ergibt sich aus den anliegenden Spannungen als

$$V_G = V_C + V_0/2 \quad (2.2)$$

und die gesuchte Kapazität C berechnet sich aus dem Verhältnis des gemessenen Stroms I , der durch die Kapazität gepumpt wird, und der Frequenz f , mit der die Transistoren

geschaltet werden:

$$C = \frac{I}{f \cdot V_0} \quad (2.3)$$

Trotz der Einfachheit der Methode zeigen die Meßergebnisse keinen Unterschied im Vergleich zum klassischen HFCV-Verfahren, wie in Bild 2.12(b) zu sehen ist. Durch Vergleich mit einer Referenzstruktur, an der nur die parasitären Effekte gemessen werden, kann die Auflösung für kleine Strukturen noch verbessert werden.

2.3 Stromverstärker

Eine naheliegende Möglichkeit, um auch schwache Meßsignale noch auswerten zu können, ist deren Verstärkung. Durch räumliche Nähe des Verstärkers zur Signalquelle können störende Einflüsse soweit wie möglich vermieden werden. Der Extremfall der Nähe wird durch sofortige Verstärkung der zu untersuchenden Größen schon auf dem Siliziumchip und der damit verbundenen minimalen Zuleitungskapazitäten erreicht, wodurch außerdem die Entkopplung der äußeren Meßtechnik vom zu untersuchenden Objekt erreicht wird. Durch eine Strom-Spannungswandlung wiederum kann das Messen der interessierenden Ströme über Shuntwiderstände oder externe I-U-Wandler vermieden und direkt mit spannungsbasierten Meßgeräten gearbeitet werden. Aus diesen Gründen wurde ein einfacher, genuiner Stromverstärker in CMOS-Technologie zur direkten On-Chip-Stromverstärkung und Strom-Spannungswandlung implementiert [19]. Ziel war es, einen robusten Verstärker mit geringen Abmessungen zu entwerfen, der durch seinen einfachen Aufbau möglichst leicht skalierbar sein sollte. Vorgabe war eine I-U-Verstärkung von ca. $10^6 = 120\text{dB}$ und eine Grenzfrequenz größer als 10kHz. Hierfür wurde das Prinzip des Stromverstärkers ausgewählt.

2.3.1 Prinzip und Aufbau

Für einen Verstärker können in der Elektrotechnik vier grundlegende Verstärkerklassen definiert werden [20]:

- Spannungsgesteuerte Spannungsquelle (Spannungsverstärker)
- Stromgesteuerte Stromquelle (Stromverstärker)
- Stromgesteuerte Spannungsquelle („transresistance amplifier, current feedback amplifier“, Strom-Spannungs-Verstärker, Transimpedanzverstärker)

- Spannungsgesteuerte Stromquelle („transconductance amplifier, OTA“, Spannungs-Strom-Verstärker)

Diese können wiederum durch äußere Beschaltung jede mögliche Verstärkerfunktion ausführen, mit einem Spannungsverstärker kann also zum Beispiel auch Strom verstärkt werden. Durch das endliche Verstärkungsbandbreiteprodukt VBP realer Verstärker ergeben sich bei äußerer Beschaltung unterschiedliche Übertragungsfunktionen (geschlossene Schleifenverstärkung), die je nach Verwendungszweck eine bestimmte Verstärkerklasse vorteilhaft gegenüber einer anderen machen. Insbesondere gilt das für die Einstellung der Grenzfrequenz, also des dominierenden Pols in Abhängigkeit von der Rückkopplungsschleife. Dies kann vergegenwärtigt werden, wenn man die Übertragungsfunktion der geschlossenen Schleife mit einem dominierenden Pol als

$$\ddot{U}(s) = k_1 \cdot \frac{1}{1 + \frac{k_2 s}{\text{VBP}}} \quad (2.4)$$

schreibt, wobei sich k_1 und k_2 durch das Rückkopplungsnetzwerk ergeben und je nach Verstärkerklasse und -funktion voneinander unabhängig sein können [20]. Innerhalb des Verstärkungsbandbreiteprodukts des realen Verstärkers kann also die geschlossene Schleifenverstärkung über k_1 unabhängig von der Grenzfrequenz VBP/k_2 eingestellt werden.

In der Praxis ist der Spannungsverstärker, der als Differenzverstärker ausgeführt und synonym als Operationsverstärker bezeichnet wird, der gebräuchlichste Verstärker für jede Art von Anwendung. In der Beschaltung als Spannungsverstärker gilt hier $k_1 = k_2$, woraus die bekannte Einschränkung eines konstanten Verstärkungsbandbreiteprodukts der geschlossenen Schleife folgt. Durch intelligente Realisierungen in Bipolar- oder CMOS-Technologien sind inzwischen jedoch weitere Bauformen aus obiger Liste zum Leben erweckt worden. Bedingt durch technologische Verbesserungen und der Zuwendung zum Strom als informationstragendem Signal setzen sich vermehrt diese alternativen Konzepte durch, wobei teilweise unterschiedliche Definitionen und Begriffe die Orientierung erschweren. Zusätzlich muß unterschieden werden zwischen Verstärkern anderer Bauform, die mit gleicher äußerer Beschaltung anstatt des Spannungs-(Operations)-Verstärkers eingesetzt werden sollen und aufgrund bestimmter Merkmale einen Vorteil versprechen. Auf diese Weise können die bekannten Schaltungskonzepte weiter verwendet und nur die Vorteile des neuen Verstärkertypus voll ausgenutzt werden. Bei bestimmten Typen müssen jedoch auch deren Beschränkungen berücksichtigt werden, die dann ebenfalls in die äußere Beschaltung mit einfließen und somit eine intensivere Beschäftigung mit der neuartigen Verstärkerklasse

notwendig machen. Dies ist natürlich ein Hemmnis für deren Anwendung.

In kommerziell vertriebenen Verstärkern spielt neben dem Differenzverstärker der Transimpedanzverstärker die wichtigste Rolle, der in gleicher äußerer Beschaltung äquivalent verwendet werden kann. Die prinzipiellen Vorteile des Transimpedanzverstärkers sind eine nur von der Rückkopplungsimpedanz abhängige Bandbreite sowie die erheblich höhere Anstiegsgeschwindigkeit. Der Unterschied in der slew rate beträgt mehr als eine Größenordnung, da beim Transimpedanzverstärker theoretisch Groß- und Kleinsignalverhalten keinen Unterschied aufweisen. Dies hängt jedoch auch davon ab, ob der Verstärker nicht-invertierend oder invertierend betrieben wird, da die beiden Eingänge stark unsymmetrisch sind. Die Schleifenverstärkung kann unabhängig von der Bandbreite eingestellt werden und das äquivalente Eingangsruschen ist niedriger als bei einem Verstärker mit Differenzeingangsstufe. Nachteile sind schlechtere DC-Eigenschaften und hoher Ruheeingangsstrom [21].

Beim Operationsverstärker mit Differenzeingangsstufe wird die angelegte Differenzspannung in einen Strom gewandelt, verstärkt und dann wiederum in eine Spannung zurückgewandelt und ausgegeben. Die hier verwendete Architektur stützt sich im Gegensatz dazu auf das Prinzip des Stromverstärkers, bei dem zunächst der Strom in einem Stromspiegel in eine Spannung transformiert und verstärkt und daraufhin wieder in einen Strom zurückgewandelt wird [22–24]. Bild 2.13 zeigt die verwendete Schaltung, die eine Abwandlung der Konzepte des Transimpedanzverstärkers bzw. Stromverstärkers darstellt. Die vier Transistoren I0, I1, I9 und I10 stellen einen Spannungsfolger dar, der das an in_p angelegte Potential auf inn überträgt. Die Stromspiegel I11-I13 bzw. I8-I12 führen zu einer Strom-Spannungsverstärkung des Eingangsstroms an inn , die verstärkte Spannung kann an sp abgegriffen werden. Diese Spannung wird an I15 in einen Strom gewandelt und verstärkt, dieser wird an out gegen Ground gemessen. Über die Offsetspannung an I14 kann der Ausgangsstrom für den statischen Fall zu Null ausgeglichen werden.

Die Kleinsignalverstärkung läßt sich aus dem in Bild 2.14 gezeigten vereinfachten Ersatzschaltbild ableiten. Hierbei liegt in_p an einer konstanten Spannung, die Stromeinspeisung des AC-Stroms erfolgt an inn . Wird perfektes Matching angenommen, so kann für die Transistoren I8 bis I13 gleiches g_m und g_d angesetzt werden. Der Strom i_{inn} teilt sich dann jeweils zur Hälfte auf den oberen und den unteren Zweig I8, I9 und I10, I11 auf. Über Transistor I8 bzw. I11 fällt also die Spannung

$$v_8 = \frac{i_{inn}}{2} \cdot \frac{1}{g_{m,8}} = v_{11} = \frac{i_{inn}}{2} \cdot \frac{1}{g_{m,11}} = \frac{i_{inn}}{2g_m} \quad (2.5)$$

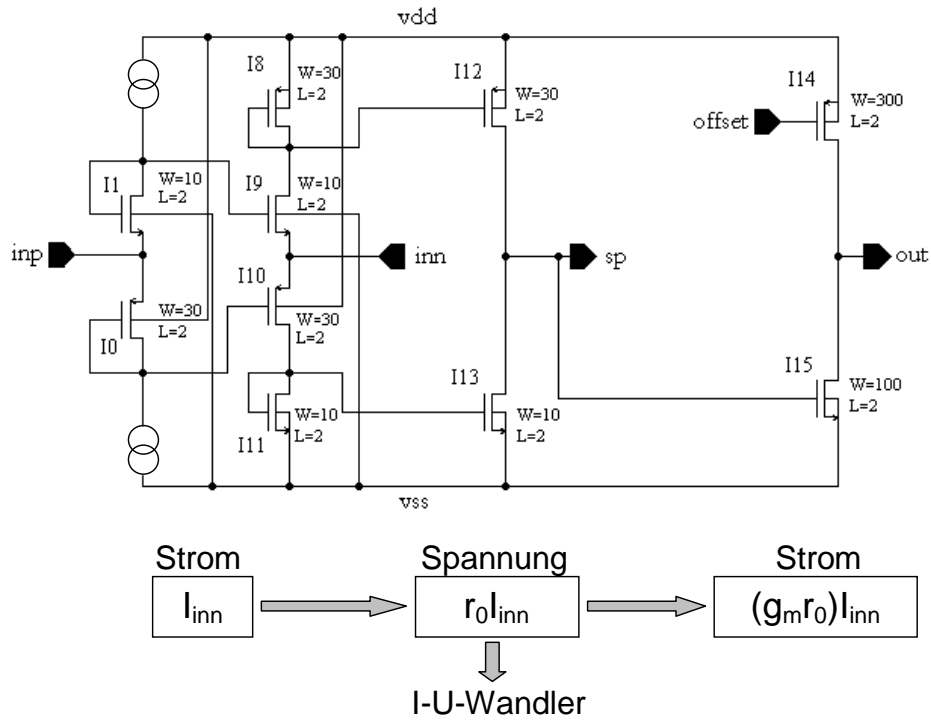


Bild 2.13: Aufbau und Prinzip des entwickelten Stromverstärkers, L und W in μm (s. auch Bild 5.5 auf S. 89).

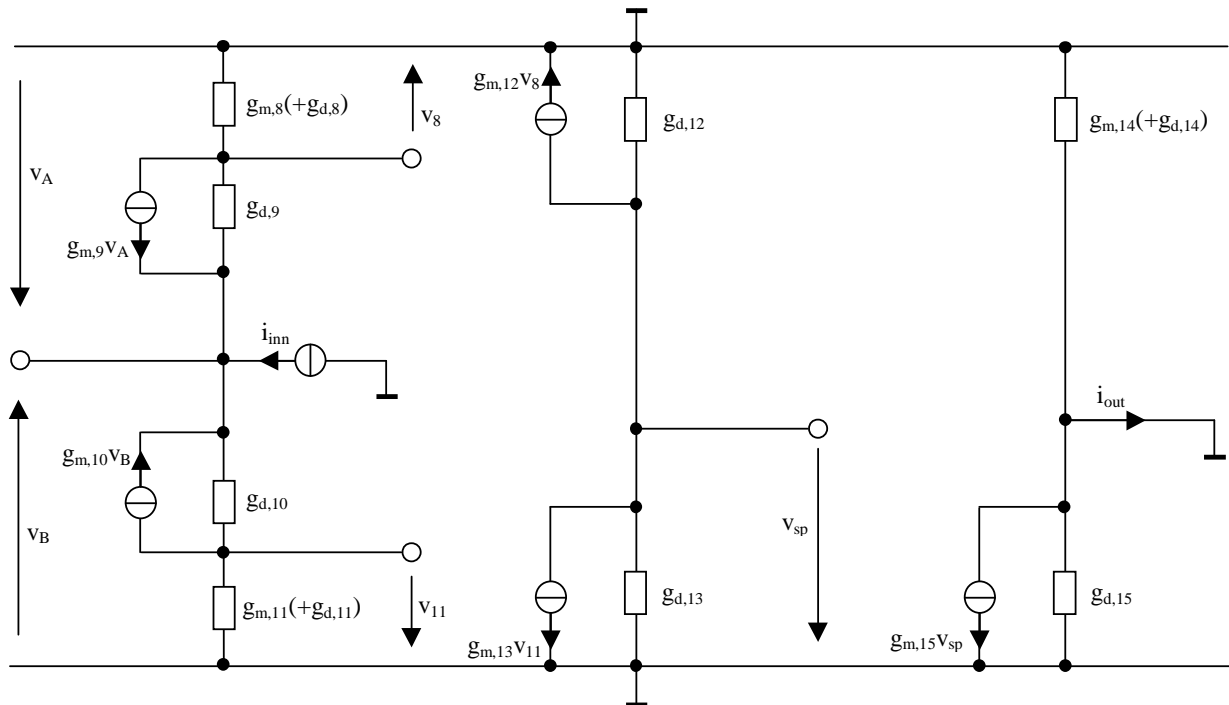


Bild 2.14: Einfaches Kleinsignalersatzschaltbild des Stromverstärkers.

ab, die die Steuerspannung für die Transistoren I12 und I13 darstellt. Die beiden Transistoren führen den gleichen Strom, das Potential an sp ergibt sich zu

$$v_{sp} = -\frac{g_{m,12} \cdot v_8 + g_{m,13} \cdot v_{11}}{g_{d,12} + g_{d,13}} = -\frac{2g_m}{2g_d} \frac{i_{inn}}{2g_m} = -\frac{i_{inn}}{2g_d}. \quad (2.6)$$

Der durch die Stromquellen in Bild 2.13 eingeprägte DC-Strom beträgt $20\mu\text{A}$ in $0.6\mu\text{m}$ Technologie. Die Kanallängenmodulation kann mit $\lambda = 2 \cdot 10^{-2}\text{V}^{-1}$ angesetzt werden. Mit $g_d \approx \lambda I_D = 0.4\mu\text{AV}^{-1}$ folgt also für die I-U-Verstärkung

$$-\frac{v_{sp}}{i_{inn}} = \frac{1}{2g_d} \approx 1.25 \cdot 10^6 \frac{\text{V}}{\text{A}} = 122\text{dB}(\Omega). \quad (2.7)$$

Die Stromverstärkung erfolgt durch I15. Der Strom fließt vollständig zum Kleinsignal-Kurzschluß an out ab. Es folgt

$$i_{out} = -g_{m,15} v_{sp} = \frac{g_{m,15}}{2g_d} i_{inn}.$$

Der durch den Transistor I14 eingeprägte Strom beträgt 8.6mA . Daraus ergibt sich $g_{m,15} \approx \sqrt{2KI_D W/L} = 10.2\text{mAV}^{-1}$ mit $K = 120\mu\text{AV}^{-2}$ beim nMOS-Transistor in $0.6\mu\text{m}$ Technologie. Die I-I-Verstärkung berechnet sich somit zu

$$\frac{i_{out}}{i_{inn}} = \frac{g_{m,15}}{2g_d} \approx 1.28 \cdot 10^4 = 82\text{dB}. \quad (2.8)$$

Mit dem gezeigten Konzept ist auch ohne äußere Beschaltung eine Strom-Spannungswandlung mit hoher Verstärkung möglich, so daß die Problematik der Herstellung integrierter Widerstände entfällt. Aufgrund des Vorteils der höheren slew rate wurde es für die Aufnahme von Signalen mit steilen Flanken als geeignet empfunden und anhand der im Rahmen dieser Arbeit untersuchten Random Telegraph Signals erprobt (s. Kapitel 4). Um auch in einfachen CMOS-Technologien hergestellt werden zu können, wurde auf eine Frequenzkompensation des Verstärkers verzichtet. Dadurch entfallen Kapazitäten als Schaltelemente, die in der Regel mit Hilfe einer zusätzlichen Polysiliziumebene hergestellt werden und eine weitere Maskenebene erforderlich machen. Eine Optimierung in Bezug auf Rauschen oder Leistungsaufnahme wurde nicht durchgeführt.

2.3.2 Realisierung

Drei leicht unterschiedliche Varianten der Stromverstärkerarchitektur, eine davon bestehend aus zwei hintereinandergeschalteten kaskodierten Verstärkern, wurden in einem $0.6\mu\text{m}$ Standard-CMOS-Prozess von AMS erfolgreich verwirklicht. Die einfachste, in Bild 2.13 gezeigte Variante wurde außerdem in einem $0.35\mu\text{m}$ Prozess mit identischem Layout realisiert, um das Skalierverhalten zu untersuchen. Das Layout des Verstärkers in beiden Technologien ist in Bild 2.15 zu sehen. Es ist nicht durch Guard-Ringe oder bezüglich Matching

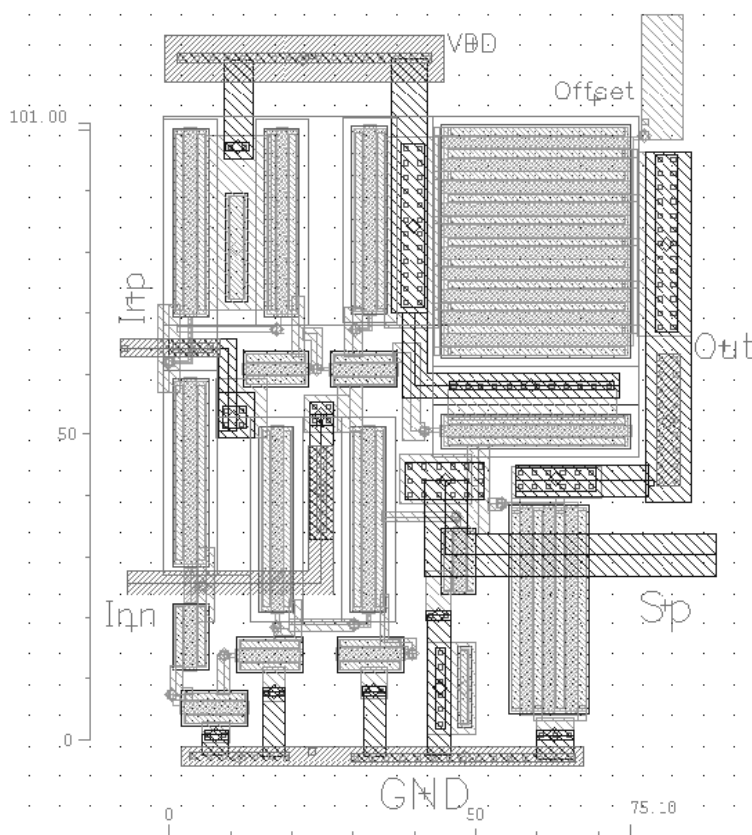


Bild 2.15: Layout des in 0.6 und $0.35\mu\text{m}$ Technologie realisierten Stromverstärkers, Abmessungen in μm .

optimiert, trotzdem war der Betrieb einwandfrei und zeigt somit die Qualität heutiger Prozesse. Die Abmessungen dieser Variante sind ca. $100\mu\text{m} \cdot 75\mu\text{m}$, wobei das Layout auch die zur Verwirklichung der Konstantstromquellen verwendeten sechs Transistoren enthält.

DC-Messungen wurden mit einem Keithley SCS-4200 Semiconductor Characterization System realisiert. AC-Analysen wurden mit einem ZVRE Vektoranalysator von Rohde&Schwarz durchgeführt, die Einspeisung erfolgte über einen $1\text{M}\Omega$ Widerstand und der

Ausgangsstrom wurde über der 50Ω Eingangsimpedanz des Analysators gemessen. Die Messungen wurden mit der SPICE-Simulation mit SpectreS verglichen. Diese beinhaltet die aus dem Layout extrahierten parasitären Kapazitäten. Die DC-Verstärkung für den Verstärker in $0.6\mu\text{m}$ Technologie ist in Bild 2.16 gezeigt. Gute Übereinstimmung der Meß-

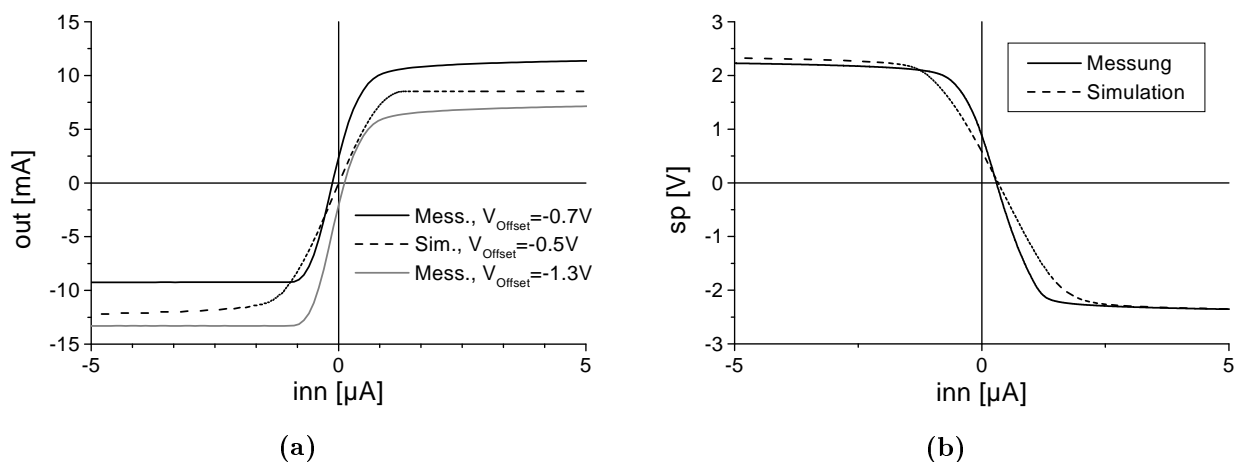


Bild 2.16: (a) I-I-DC-Verstärkung und (b) I-U-DC-Verstärkung des Stromverstärkers in $0.6\mu\text{m}$ Technologie.

und Simulationskurven war für beide Technologien festzustellen. Die Frequenzgänge sind in Bild 2.17 gezeigt. Der Vergleich mit den Meßwerten ist in der $0.6\mu\text{m}$ Technologie recht gut, in der $0.35\mu\text{m}$ Technologie allerdings weniger befriedigend. Bei dem $0.35\mu\text{m}$ Prozeß ergaben sich außerdem verstärkt Latch-up-Probleme. Hier macht sich sicherlich das ungünstige Layout ohne Guard-Ringe oder Optimierung bezüglich Matching bemerkbar, was bei einer moderneren Technologie mit höheren Steilheiten und geringerer Versorgungsspannung wahrscheinlich einen höheren Einfluß hat. In Tabelle 2.1 sind die wichtigsten Eckdaten des Stromverstärkers in beiden Technologien gezeigt. Die gemessene I-U- und I-I-Verstärkung stimmen gut mit den in Abschnitt 2.3.1 berechneten Werten überein. Eine deutliche Reduzierung der Fläche ist unter Beibehaltung der Spiegelverhältnisse ohne Weiteres möglich.

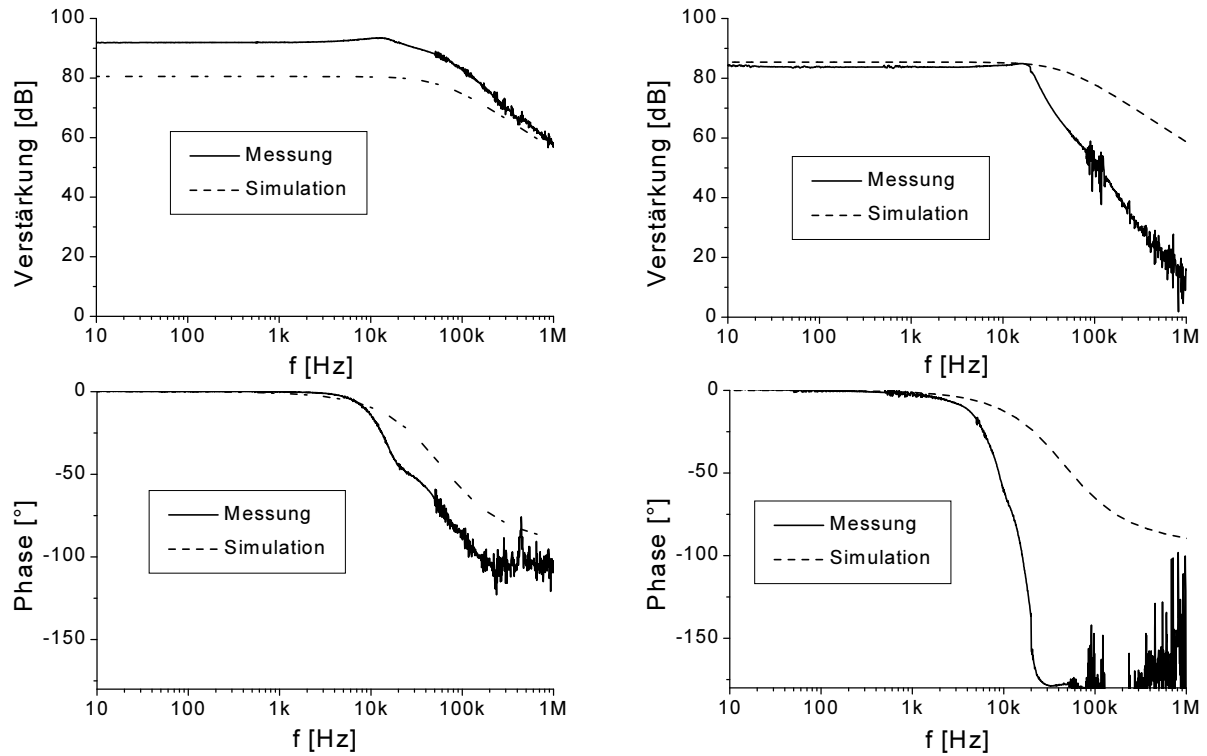


Bild 2.17: Frequenz- und Phasengänge für 0.6µm (links) und 0.35µm (rechts) Technologie.

Tabelle 2.1: Kenndaten des Stromverstärkers

Technologie	0.6µm		0.35µm	
	Messung	Simulation	Messung	Simulation
V_{DD}/V_{SS} [V]	±2.5	±2.5	±1.65	±1.65
V_{Offset} [V]	-1	-0.5	-0.35	-0.2
Offene I-I-Verstärkung [dB]	92	81	84	85
Grenzfrequenz (Phase=-45°) [kHz]	21	56	8	47
Grenzfrequenz (3dB) [kHz]	43	56	21	47
Eingangsimpedanz von inn [kΩ]	4	4	5	4
Offene I-U-Verstärkung [dB(Ω)]	120	125	-	130

Das Rauschen wurde mit einem SR756 Stanford Research Dynamic Signal Analyzer gemessen. Bild 2.18 zeigt das am Spannungsausgang sp gemessene Rauschen. Der $1/f$ -förmige Verlauf (s. Kapitel 4) bei niedrigen Frequenzen ist deutlich erkennbar. Hier sind Verbesserungen notwendig, um das relativ hohe Rauschen zu reduzieren.

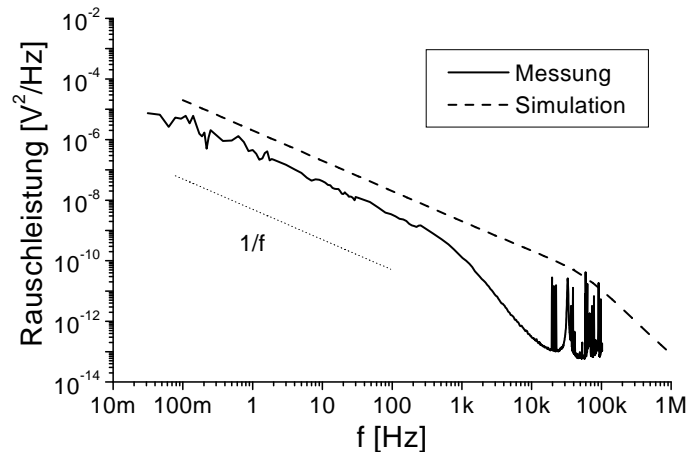


Bild 2.18: Gemessene und simulierte Rauschleistungsdichte an sp der $0.6\mu\text{m}$ Technologie bei 0A Eingangsstrom (inp auf 0V).

2.3.3 Anwendung

Die Funktionalität des Stromverstärkers zur Untersuchung von Fluktuationen konnte erfolgreich demonstriert werden [19]. Hierfür wurde in der $0.6\mu\text{m}$ Technologie auf dem Chip ein pMOS-Transistor wie in Bild 2.19 gezeigt an den Stromverstärker angeschlossen. Die DC-Spannungen wurden mit dem Keithley SCS-4200 bereitgestellt und am Ausgang sp des Stromverstärkers der I-U-gewandelte und mit Faktor 10^6 verstärkte Drainstrom des

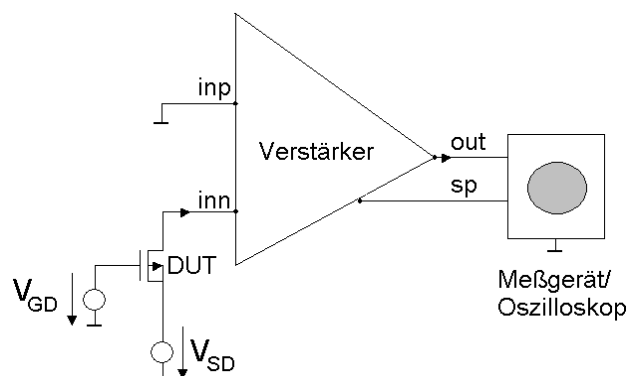


Bild 2.19: Schaltung zur Vermessung eines pMOS-Transistors.

pMOS-Transistors aufgenommen. Die Spannung an sp konnte direkt mit einem LeCroy Oszilloskop LC584AM gemessen werden. Der pMOS-Transistor wurde einem homogenen 10-minütigem Streß von $V_{GS} = 10V$ ausgesetzt. Durch die Schädigung des Oxids ergeben sich Grenzflächenzustände, die zu Fluktuationen im Drainstrom führen können (s. Kapitel 4). Bild 2.20 zeigt einen so gemessenen Stromsprung. Der Vorteil der höheren zeitlichen

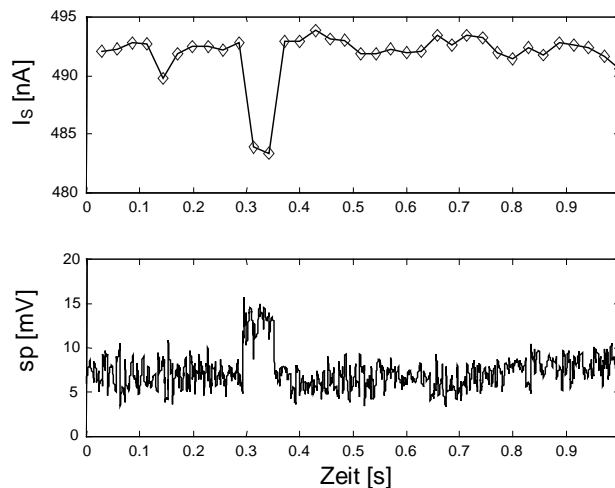


Bild 2.20: Oben: Direkt mit dem SCS-4200 gemessener Sourcestrom des gestreßten Transistors, Samplezeit 30ms; Unten: Verstärkt-gewandeltes und mit dem LC584AM gemessenes Signal, Samplezeit 4 μ s, Nachbearbeitung unter Matlab mit 50Hz Notch-Filter und Wavelet-Denoising-Toolbox (s. Abschnitt 4.2); $V_{GD} = -1.092V$.

Auflösung des Oszilloskops ist deutlich zu sehen. Der Stromsprung wird bei der direkten Strommessung mit dem Keithley SCS-4200 nur mit zwei Samplepunkten erfaßt, während beim I-U-gewandelten Signal die Flanken scharf nachgebildet werden. Strommeßgeräte haben im Vergleich zu Spannungsmeißgeräten nur eine sehr geringe Bandbreite, für höherfrequente Fluktuationen liegt der Vorteil der Meßmethode mit Verstärker und Oszilloskop dadurch auf der Hand.

Da die Abhängigkeit von Prozeßschwankungen und vom Arbeitspunkt in der offenen Konfiguration von Bild 2.19 erheblich ist, muß allerdings über Maßnahmen zur Rückkopplung und Stabilisierung der Verstärkung nachgedacht werden. Hierbei müßten dann auch Rauschaspekte mit berücksichtigt werden. Abschließend läßt sich sagen, daß die Verstärkung von Defektströmen allgemein, und das gezeigte Stromverstärkerkonzept den Anwendungsbereich von Standardmeßverfahren erweitern und somit den höheren schaltungstechnischen Aufwand rechtfertigen könnten.

3

Kennlinien nach Oxiddurchbruch

Mit zunehmender Integration und Komplexität der Schaltungen wird auch die Fehlerwahrscheinlichkeit und -anfälligkeit immer größer. Bereits jetzt können auf einem Chip mehrere Millionen Transistoren untergebracht werden, die alle innerhalb eines engen Toleranzbereiches funktionieren müssen. Bei immer kleiner werdenden Strukturen mit Sub-100nm-Kanallänge wirken sich statistische Schwankungen von Dotieratomen und Ladungsträgern immer stärker aus und führen zu Fluktuationen von Kenngrößen der Bauelemente wie Schwellspannung und Steilheit, die durch Verbesserungen am Prozeß nicht eliminiert werden können [25]. Dies führt zu dem bekannten Problem, daß analoge Schaltungen nicht im selben Maße wie die digitalen Gatter skaliert werden können. Auch digitale Schaltungen könnten aber in zunehmendem Maße betroffen sein. Möglicherweise verstärkt sich hierdurch der Trend, fehlertolerante Schaltungen zu entwickeln, die in einem noch weiteren Bereich als den jetzt im Rahmen der „corner Analyse“ berücksichtigten Prozeßschwankungen funktionsfähig sind. Die Steigerung wäre eine Schaltung, die selbst bei durchgebrochenem Gateoxid eines oder mehrerer Transistoren noch in einem gewissen Toleranzbereich arbeitet. Natürlich wird die Stromaufnahme einer solchen Schaltung ansteigen und über die weitere Ausbreitung des Fehlers auf andere Teile der Schaltung bzw. die weitere Zerstörung des Bauelements bei anhaltender Verwendung ist ebenfalls nachzudenken. Deswegen werden weiterhin alle Anstrengungen unternommen werden, um die Wahrscheinlichkeit eines Durchbruchs im realen Betrieb zu minimieren. Bei dünneren Oxiden nehmen die Zuverlässigkeitsreserven jedoch stetig ab. Dünnere Oxide sind intrinsisch unzuverlässiger, da eine geringere kritische Defektdichte zur Erzeugung eines Durchbruchspfades im Oxid notwendig ist. Bild 3.1 zeigt schematisch die Entstehung eines Durchbruchspfades nach dem Perkulationsmodell [26, 27]. Es wird deutlich, daß bei dünneren Oxiden ein Durch-

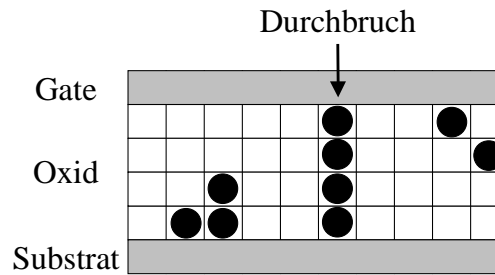


Bild 3.1: Schematische Darstellung der Defektgeneration (schwarze Punkte) im Oxid und Vollen- dung des Durchbruchpfads [26,27].

bruchspfad schneller entstehen kann. Andererseits wird die Voraussage der Lebensdauer unter Betriebsbedingungen aus Streßexperimenten immer unsicherer [28–30]. Dies führt dazu, daß ein Gateoxiddurchbruch bei dünnsten Oxiden möglicherweise nicht mehr unter vertretbarem Aufwand vermieden werden kann.

Unabhängig von der Auswirkung auf Schaltungen ist mit der Analyse des Verhaltens von MOS-Transistoren nach Oxiddurchbruch unterschiedlicher Art jedoch auch die Hoff- nung verbunden, Aufschluß über physikalische Grundmechanismen des MOS-Systems und unterschiedliche Transportarten zu erhalten. Da nach dem Durchbruch auch eine starke Zunahme des Rauschens zu verzeichnen ist, können auch Grundmechanismen von Rausch- quellen an der Silizium-Siliziumdioxid-Grenzfläche abgeleitet werden (s. Kapitel 4).

Der Gateoxiddurchbruch bedeutet, wie die Bilder 3.1 und 3.2 zeigen, den Verlust der Isolierfähigkeit in einem Teilstück des Oxids und somit einen niederohmigen leitenden Pfad vom Gate zum Kanal oder Source-/Drain-Überlappbereich. Bei Oxiden unter 5nm Dicke wird eine Sonderform des Durchbruchs beobachtet, der sogenannte weiche Durchbruch, bei dem die Isolierfähigkeit des Oxids nicht vollständig verschwindet und die Stromhöhe im Allgemeinen unter der des harten Durchbruchs liegt. Bild 3.3(a) zeigt Beispiele für das

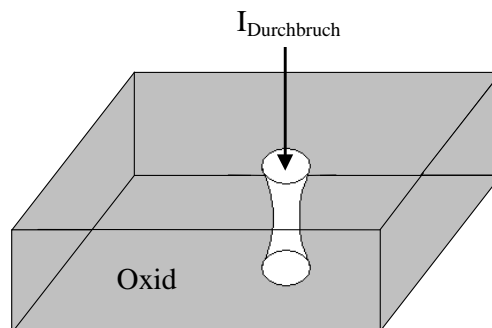


Bild 3.2: Schematische Darstellung des Durchbruchkanals durch das Oxid.

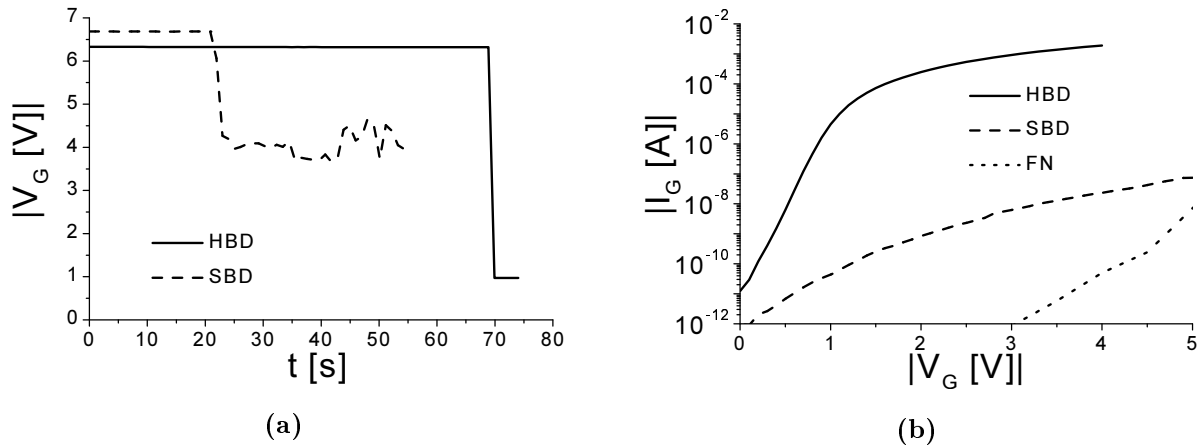


Bild 3.3: (a) Harter (HBD) und weicher (SBD) Durchbruch von MOS-Strukturen mit 4.4nm dicken Oxiden in Akkumulation bei konstanter Strombelastung, (b) typische Größenordnungen der Gatestromverläufe von MOS-Kapazitäten in Akkumulation nach Durchbruch im Vergleich zum Fowler-Nordheim-Tunnelstrom (FN) vor Durchbruch.

Verhalten bei Durchbruch durch konstante Strombelastung des Oxids. Das Niveau der Spannung nach dem ersten Sprung dient dabei als erstes Kriterium für die Restisolierfähigkeit und der Einstufung als hartem oder weichem Durchbruch. In Bild 3.3(b) werden die unterschiedlichen Strom-Spannungskennlinien der beiden Durchbrucharten deutlich. Für das Bauelement bedeutet der Durchbruch außer dem einsetzenden zusätzlichen Gatestrom, der sich je nach Durchbruchsort und -art auf die Elektroden aufteilt, eine Veränderung der DC-Stromkennlinien und auch, wie in der gestrichelten Kennlinie von Bild 3.3(a) erkennbar, erhöhtes Rauschen. Bei sehr dünnen Oxiden unter 2.5nm Dicke erfolgt auch der harte Durchbruch nicht mehr schlagartig wie in Bild 3.3(a), sondern progressiv mit einer meßbaren Geschwindigkeit [31, 32].

In diesem Kapitel werden zuerst weiche und harte Durchbrüche an MOS-Kapazitäten und Modelle für die auftretenden Gatestromkennlinien vorgestellt und sodann MOS-Transistoren untersucht.

3.1 Durchbruchsuntersuchungen an MOS-Kapazitäten

Um die grundlegenden Transportmechanismen nach Durchbruch zu untersuchen, bietet sich die im Vergleich zum Transistor deutlich einfachere MOS-Kapazität an. Hieran wird zunächst ausschließlich der einsetzende Gateleckstrom analysiert.

3.1.1 Harter Durchbruch

Es wurden Durchbrüche an Kapazitäten mit unterschiedlichen Dotierungen erzeugt. Standardmäßig wurden Kapazitäten aus einem dual-gate Prozeß mit n+Poly-Gate auf p-Substrat (nMOS) bzw. p+Poly-Gate auf n-Substrat (pMOS) verwendet. Zum Vergleich wurden Kapazitäten mit n+Poly-Gate auf n-Substrat herangezogen. Bild 3.4 zeigt die sich nach Durchbruch ergebenden Kennlinien. Die Strukturen mit jeweils entgegengesetzter Dotie-

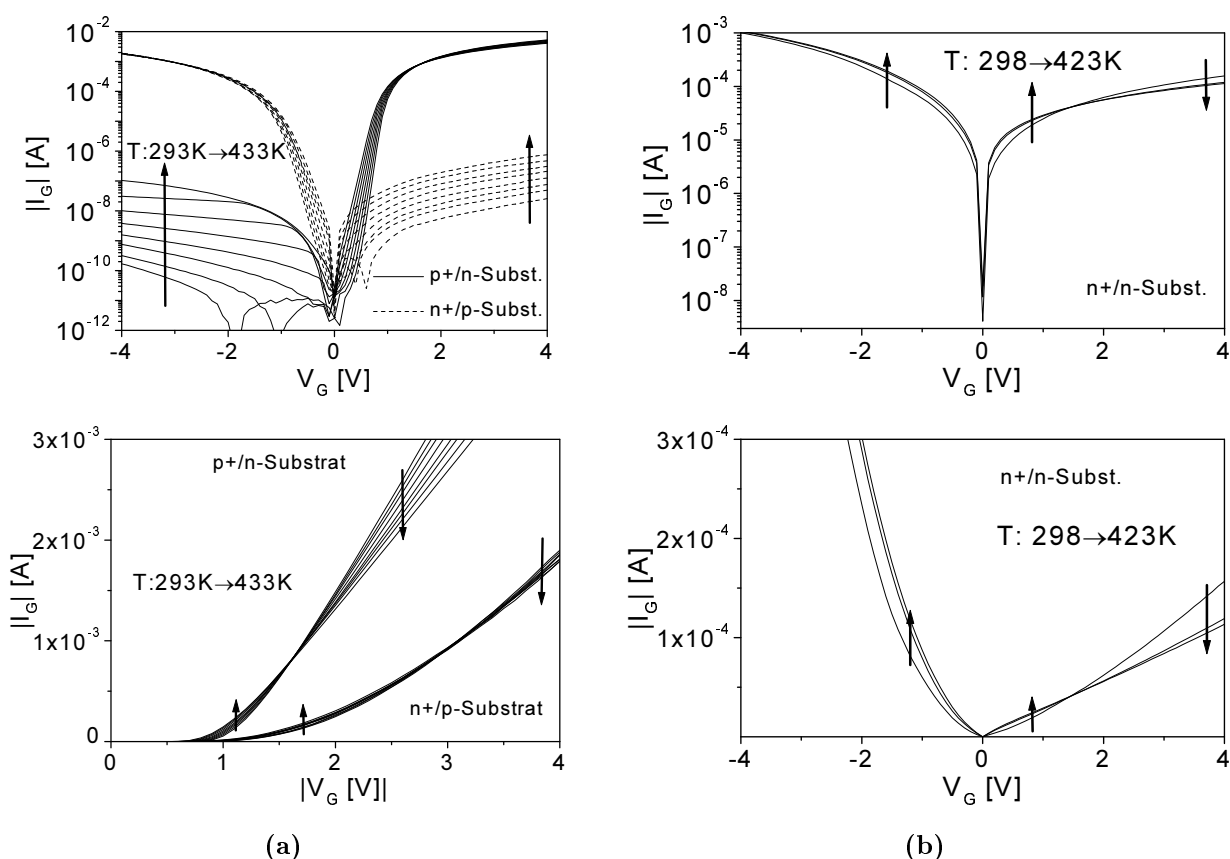


Bild 3.4: Gemessene I-U-Kennlinien in logarithmischer (oben) und linearer Darstellung (unten) von MOS-Kapazitäten aus einem Infineon 0.25 μm Prozeß nach hartem Durchbruch bei unterschiedlichen Temperaturen: (a) n+Poly-Gate auf p-Substrat bzw. p+Poly-Gate auf n-Substrat [33], (b) n+Poly-Gate auf n-Substrat. Messung (b) durchgeführt von B. Sell. Daten der Kapazitäten: $t_{ox} = 4.4\text{nm}$, $N_{Substrat} = 3 \cdot 10^{17}\text{cm}^{-3}$, $N_{Poly} = 7 \cdot 10^{19}\text{cm}^{-3}$, $A_G = 0.1\text{mm}^2$.

rung der Elektroden (Bild 3.4(a)) verhalten sich symmetrisch zueinander und weisen niedrigeren Strom in Inversion und hohen Strom in Akkumulation auf. In Akkumulation zeigt sich bei der Temperaturabhängigkeit ein Kreuzungspunkt, bei niedrigen Spannungen ist die Temperaturabhängigkeit positiv, bei hohen Spannungen negativ. Dieses Temperaturverhalten sieht man ebenfalls bei den Kapazitäten mit gleichartiger Dotierung (Bild 3.4(b)).

Bei diesen Proben ist jedoch auch in Inversion ein hoher Stromfluß vorhanden. Auffallend ist, daß die Kennlinie nach Durchbruch nur bei gleichartiger Dotierung praktisch ohmsch ist, während sie bei entgegengesetzter Dotierung erst zu höheren Spannungen linear wird.

Der Durchbruch wurde mit dem Programm Atlas, wie in Bild 3.5(a) gezeigt, durch Wegnahme des Oxids in einem kurzen Abschnitt, also durch Kontaktierung der Elektroden in diesem Bereich, simuliert. Die Simulationsergebnisse zeigen die selben Charakteristika

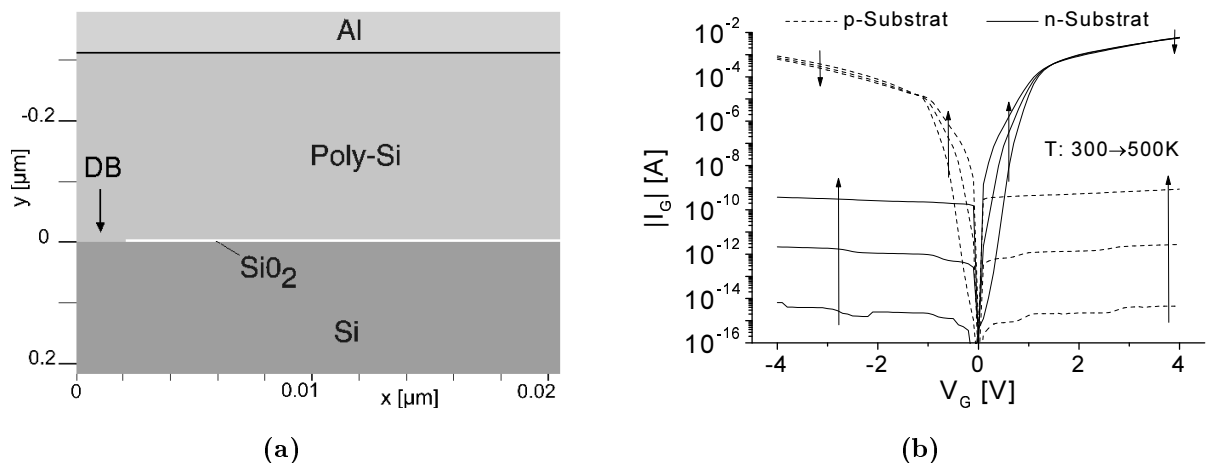


Bild 3.5: (a) Mit Atlas von Silvaco simulierter Durchbruch (DB) in zylindrischen Koordinaten, (b) Simulierte I-U-Kennlinien bei unterschiedlichen Temperaturen (Radius des Durchbruchs 2nm beim p-Substrat und 10nm beim n-Substrat; Verwendete Modelle: mos, impact selb [34]).

sowohl für entgegengesetzte Dotierung (Bild 3.5(b)) als auch für gleichartige Dotierung. Zur Modellierung des Verhaltens der für die Praxis bedeutenderen dual-gate Dotierung wurden zwei sich ergänzende Deutungsversuche entwickelt, die im Folgenden vorgestellt werden. Das Verständnis für die single-gate Dotierung ergibt sich aus den Modellen.

Durchbruch als transparente Barriere und Begrenzung des Stromflusses durch die zur Verfügung stehende Ladung an den Elektroden

Hierbei wird angenommen, daß durch den Durchbruch ein vollständig durchsichtiger leitender Pfad zwischen den beiden Elektroden der durchgebrochenen Kapazität entsteht und die I-U-Kennlinie nur durch die zur Verfügung stehende Ladung an den Elektroden zustandekommt. Der Strom ist somit nur ein Abbild der Ladungsträgerkonzentration. Demnach ist der Strom in Inversion begrenzt durch die Generationsrate, die mit der Temperatur zunimmt und in Akkumulation begrenzt durch die sich an der Halbleiter-Oxid-Grenzschicht

akkumulierende Ladung. In Inversion ist der Strom praktisch spannungsunabhängig, da die Generationsrate durch die angelegte Spannung nicht beeinflusst wird. Der in Bild 3.4(a) gemessene Anstieg des Stromes mit der Spannung ist wahrscheinlich eher auf die Art der Messung zurückzuführen (Geschwindigkeit und Spannungsrichtung). Bei anderen Messungen (s. auch Bild 3.8) war dieser Strom praktisch konstant. Die Temperaturabhängigkeit wird, wie auch der Vergleich mit den Simulationen zeigt, richtig wiedergegeben. Simulationen mit MINIMOS 6.1 ergaben, daß die Ladungsträgerkonzentration in Akkumulation als Funktion der Temperatur einen ebensolchen Kreuzungspunkt wie die gemessenen Stromkennlinien aufweist, hervorgerufen durch eine bei niedrigen Spannungen positive und bei hohen Spannungen negative Temperaturabhängigkeit [33]. Wird der Durchbruch als transparent angesehen, stellt er keine Barriere für den Stromfluß dar, und der Strom ergibt sich aus der zur Verfügung stehenden Ladung und der Durchbruchfläche. Der Strom vom Substrat durch die Barriere berechnet sich zu [35]:

$$J = \frac{4\pi q m}{h^3} \cdot \int_{E_C}^{\infty} F(E - E_{FC}) \cdot \int_0^{E - E_C} D(E - E_t) dE_t dE \quad (3.1)$$

Hierbei ist h das Plancksche Wirkungsquantum, m die Masse des Elektrons, $F(E)$ die Fermi-Dirac-Verteilung, $D(E)$ die Übergangswahrscheinlichkeit durch die Barriere und E_t die Energie der Ladungsträger quer zur Durchbruchflußrichtung. Wird die Übergangswahrscheinlichkeit zu 1 gesetzt, für die Fermi-Dirac-Funktion die Boltzmann-Näherung verwendet und die Quasi-Fermi-Energie geschrieben als

$$E_{FC} = E_C + q\phi_s - \frac{E_g}{2}, \quad (3.2)$$

ergibt sich für die Stromdichte [33]

$$J = \frac{4\pi q m}{h^3} (kT)^2 \exp\left(\frac{q\phi_s - E_g/2}{kT}\right), \quad (3.3)$$

wobei E_g die Bandlücke in Silizium und ϕ_s das Oberflächenpotential¹⁾ ist. Unter der Annahme, daß das Oberflächenpotential durch den Durchbruch nicht zu stark beeinflusst wird, kann der Wert für ein ungeschädigtes Oxid angesetzt werden. Dies wurde mit MINI-

¹⁾Es wird hier die Notation von Nicollian und Brews [36, S. 50ff] verwendet, wonach ϕ_s das Oberflächenpotential, bezogen auf das Fermi-Niveau E_F , und $\psi_s = \phi_s - \phi_B$ die Bandverbiegung an der Oberfläche ist. ϕ_B ist dabei das Volumenpotential im Substrat weit entfernt von der Grenzfläche (s. auch Bild. 4.12 auf S. 79).

MOS 6.1 simuliert. Wird die Temperaturabhängigkeit von E_g vernachlässigt, folgt der in Bild 3.6(a) gezeigte Verlauf, der alle wichtigen Merkmale der Spannungs- und Temperaturabhängigkeit in Akkumulation reproduziert. Bei niedrigen Spannungen muß der Stromfluß von der Anodenseite aus berücksichtigt werden, um für die Spannung $V_G = 0V$ den physikalisch richtigen Wert $I=0A$ zu erhalten. Für höhere Spannungen kann dieser Anteil jedoch vernachlässigt werden. Der Vergleich mit den Meßdaten in Bild 3.6(b) zeigt eine recht gute Übereinstimmung. Der Querschnitt, der zur Abstimmung mit den Meßdaten notwendig ist, ist allerdings relativ groß.

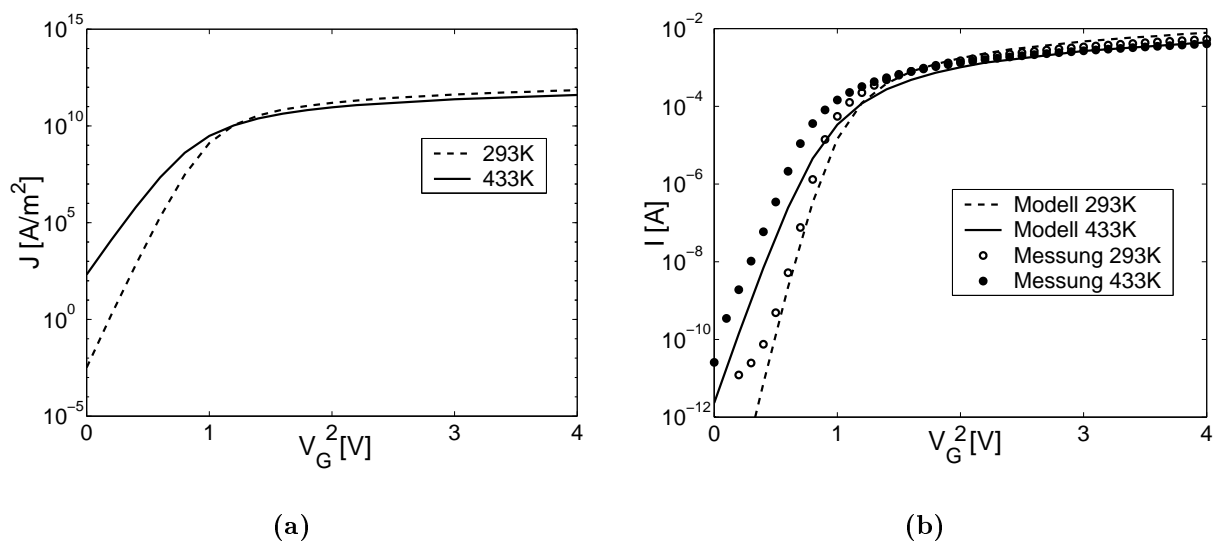


Bild 3.6: (a) Nach Gleichung 3.3 berechneter Strom durch einen harten Durchbruch in Akkumulation für die nMOS-Kapazität, (b) Vergleich mit den Meßdaten aus Bild 3.4 bei einem angenommenen Querschnitt des Durchbruchs A_S von $9 \cdot 10^{-13} m^2$.

Durchbruch als pn-Übergang

Hierbei spielt die restliche den Durchbruch umgebende MOS-Kapazität für das Strom-Spannungsverhalten keine Rolle. Die Kennlinie wird als Diodenkennlinie mit Reihen- und Parallelwiderständen aufgefaßt. Der niedrigere Strom in Inversion wird auf das Sperrverhalten der Diode zurückgeführt. Dieser Ansatz wurde von Bearda et al. ebenfalls gefunden [37]. Der Diodenstrom ergibt sich aus der Generations-Rekombinations-Komponente J_{GR} , die bei niedrigen und negativen Spannungen dominiert, und dem Diffusionsstrom J_S . Es ergeben sich folgende Gleichungen [38]:

$$J_{GR} = qn_i W_{RLZ} \frac{\sigma v_{th,x} n_T}{2} \quad (3.4)$$

$$J_S = q\sqrt{D_x\sigma v_{th,x}n_T} \cdot \frac{n_i^2}{N_{Substrat,x}} \quad (3.5)$$

Hierbei steht der Index x für den Substrattyp n oder p. $\sigma = \sigma_n = \sigma_p$ ist der Einfangquerschnitt, v_{th} die thermische Geschwindigkeit der Ladungsträger, D die Diffusionskonstante, n_i die intrinsische Ladungsträgerdichte, n_T die Störstellendichte, $N_{Substrat}$ die Substratdopierung und W_{RLZ} die Weite der Raumladungszone. Definitionen und Näherungen sind im Standardwerk von Sze [38] enthalten, für die Beweglichkeit wird ein empirisches Modell von Arora et al. [39] verwendet. Die Stromdichte J ergibt sich daraus zu [40]

$$J = J_S \left[\exp\left(\frac{V}{V_t}\right) - 1 \right] + J_{GR} \exp\left[\left(\frac{V}{2V_t}\right) - 1\right] \quad (3.6)$$

mit V_t der Temperaturspannung. Bei hohen Strömen in Vorwärtsrichtung muß zusätzlich der serielle Widerstand R_S des Strompfades berücksichtigt werden, was zu einer Abweichung vom idealen Verhalten von Gleichung 3.6 führt. Der Strom in Vorwärtsrichtung I_F ergibt sich dann zu

$$I_F = A_S \cdot \left(J_S \left[\exp\left(\frac{V - I_F R_S}{V_t}\right) - 1 \right] + J_{GR} \left[\exp\left(\frac{V - I_F R_S}{2V_t}\right) - 1 \right] \right) \quad (3.7)$$

mit A_S dem Durchbruchquerschnitt.

Die implizite Gleichung 3.7 kann vereinfacht werden, wenn angenommen wird, daß der Generations-Rekombinations-Term nur bei niedrigen Spannungen vorhanden ist, bei denen der Serienwiderstand vernachlässigt werden kann. Bei höheren Spannungen wird nur der Diffusionsterm berücksichtigt, in diesem Fall kann die -1 in der Klammer vernachlässigt und die Gleichung mittels der *LambertW* Funktion gelöst werden. Es folgt folgende Gleichung in Vorwärtsrichtung:

$$I_F = \frac{V_t}{R_S} \text{LambertW} \left[\frac{J_S A_S R_S}{V_t} \exp\left(\frac{V}{V_t}\right) \right] + J_{GR} A_S \exp\left[\left(\frac{V}{2V_t}\right) - 1\right] \quad (3.8)$$

In Anhang A ist eine einfache analytische Näherung für die *LambertW* Funktion in drei Teilstücken aufgeführt.

Der Serienwiderstand kann nach einer von Werner vorgeschlagenen Methode aus den Meßdaten durch Berechnung der Kleinsignalleitfähigkeit G bestimmt werden [41]:

$$\frac{G}{I_F} = \frac{1}{n_D V_t} (1 - G R_S) \quad (3.9)$$

Hierbei ist $1 < n_D < 2$ der Idealitätsfaktor der Diode, der aus der Steigung bei niedrigen Spannungen folgt. Es ergibt sich die in Bild 3.7 zeigte Abhängigkeit. Die Übereinstim-

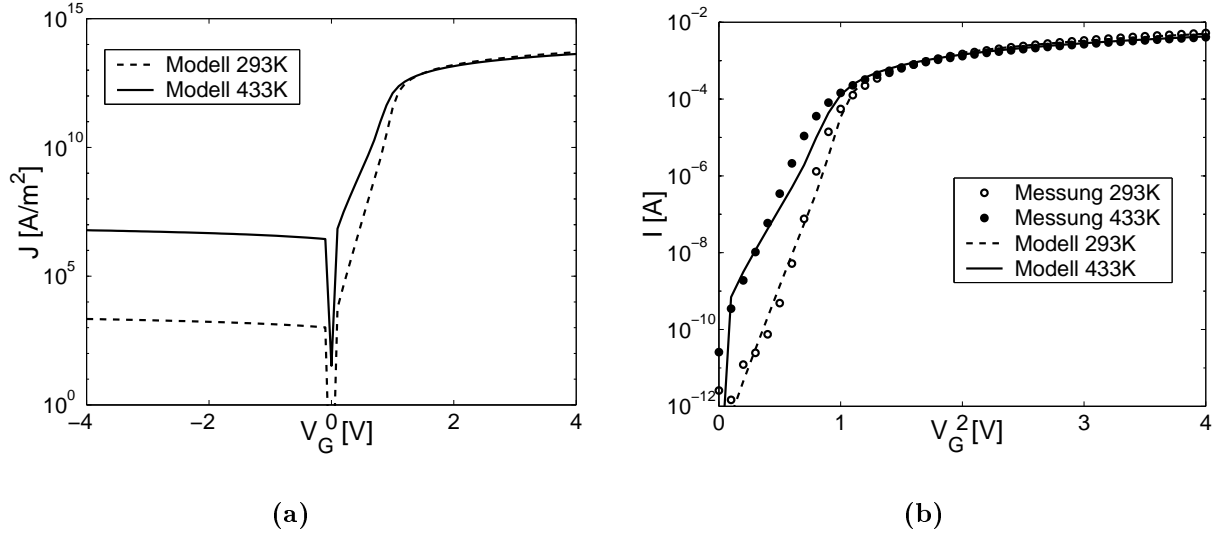


Bild 3.7: (a) Nach Gleichung 3.8 berechnete Stromdichte eines pn-Übergangs, (b) Vergleich mit den Meßdaten in Akkumulation aus Bild 3.4 bei einem angenommenen Querschnitt des Durchbruchs A_S von $1 \cdot 10^{-16} \text{m}^2$, $N_{Substrat} = 3 \cdot 10^{17} \text{cm}^{-3}$, $N_{Poly} = 7 \cdot 10^{19} \text{cm}^{-3}$, $n_T = 5 \cdot 10^{18} \text{cm}^{-3}$, $\sigma = 5 \cdot 10^{-12} \text{cm}^{-2}$, $R_S(293\text{K}) = 580\Omega$, $R_S(433\text{K}) = 694\Omega$.

mung ist sehr gut, wobei hier der Fitparameter Serienwiderstand vorhanden ist, dessen Temperaturabhängigkeit und absolute Größe noch theoretisch erfaßt werden müßte, ebenso wie die Parameter Einfangquerschnitt und Störstellendichte. Bei den untersuchten Kurven ergab sich ein Idealitätsfaktor $n_D \approx 1.5$, was die für die Berechnung angenommene hohe Defektdichte rechtfertigt. Auch der verwendete Wert für den Einfangquerschnitt liegt innerhalb der Spanne zwischen 10^{-9} und 10^{-16}cm^{-2} , die sich bei der Untersuchung von Random Telegraph Signals in Abschnitt 4.3.2 ergeben. Für die Berechnung des Serienwiderstands müßte der Widerstand des Substrats und die Flußform des Stroms modelliert werden. Die Durchbruchsfläche von 10^{-16}m^2 erscheint realistisch.

Nach Werner kann zur Erklärung eines höheren Meßstroms in Sperrichtung ein Parallelwiderstand herangezogen werden [41]. Wie bereits erwähnt zeigten die Meßdaten aber in Sperrichtung inkonsistentes Verhalten, weswegen auf eine weitergehende Modellierung verzichtet wurde.

Messungen an Proben mit Aluminiumoxid als Dielektrikum

Zum Vergleich wurden Messungen an Proben mit Al_2O_3 als Dielektrikum und n+Poly-Gate auf p-Substrat durchgeführt. Bild 3.8(a) zeigt erwartungsgemäß ein sehr ähnliches Verhalten der Durchbruchskennlinien wie bei den bisher betrachteten Siliziumoxidproben

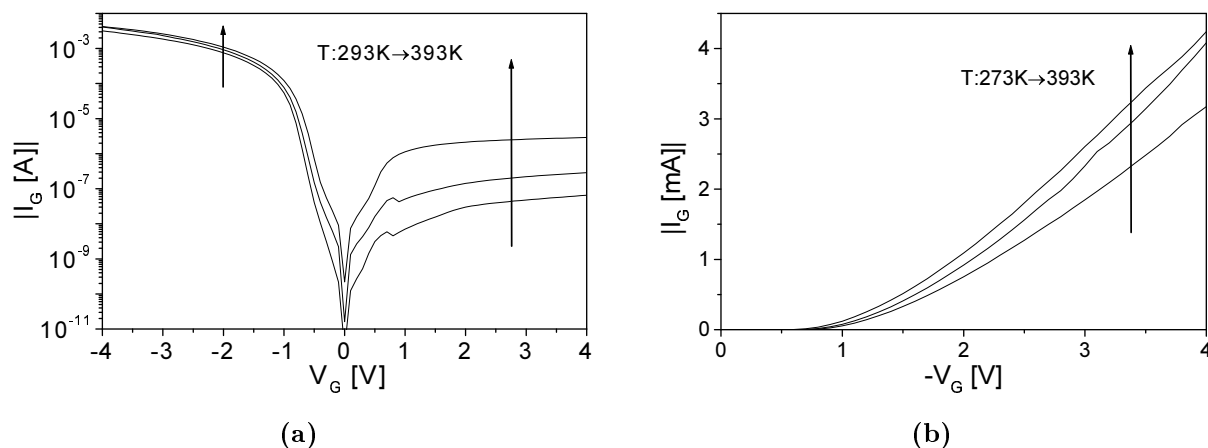


Bild 3.8: Gemessene I-U-Kennlinien nach Durchbruch bei nMOS-Kapazitäten mit Aluminiumoxid als Dielektrikum in (a) logarithmischer und (b) linearer Darstellung. Daten der Kapazität: Schichtdicke des Aluminiumoxids 4.5nm, Schichtdicke des parasitären Siliziumoxids 1-2nm, $N_{Substrat} = 3 \cdot 10^{16} \text{cm}^{-3}$, $N_{Poly} = 5 \cdot 10^{20} \text{cm}^{-3}$, $A_G = 0.1 \text{mm}^2$.

mit der Unsymmetrie um $V_G = 0\text{V}$. Bild 3.8(b) zeigt jedoch in Akkumulation eine rein positive Temperaturabhängigkeit, die zudem etwas höher ist als bei den Siliziumoxidproben. Dies legt die Vermutung nahe, daß in diesem Fall ein weiterer Effekt zum Tragen kommt.

An Aluminiumoxidproben wurde eine Ausdiffusion von reinem Aluminium bei hohen Temperaturen beobachtet [42]. Im Moment des Durchbruchs ergibt sich durch die anliegende Spannung von ca. 6V und einem Strom von ca. 50µA eine Leistung von einigen hundert Mikrowatt in dem sehr kleinen Durchbruchsvolumen von ungefähr 10^{-25}m^3 . Eine hohe Temperatur im Durchbruchskanal ist also sehr wahrscheinlich. Dadurch könnte eine starke Ausdiffusion von Metallatomen in den Durchbruchspfad erfolgen. Im Gegensatz zu den Siliziumoxidproben ergibt sich dann ein Metall-Halbleiter-Übergang statt einem pn-Übergang. Der Übergang vom Polysilizium-Gate zum Metall des Durchbruchspfads ist dabei wegen der hohen Dotierung des Gates ohmsch, während der Übergang vom Metall zum Substrat einen Schottky-Kontakt darstellt. In Bild 3.9 sind Simulationen mit Atlas gezeigt, bei denen der Durchbruchspfad aus Aluminium besteht. Die allgemeine Temperaturabhängigkeit wird in Übereinstimmung mit den Meßdaten wiedergegeben. Weitere Simulationen zeig-

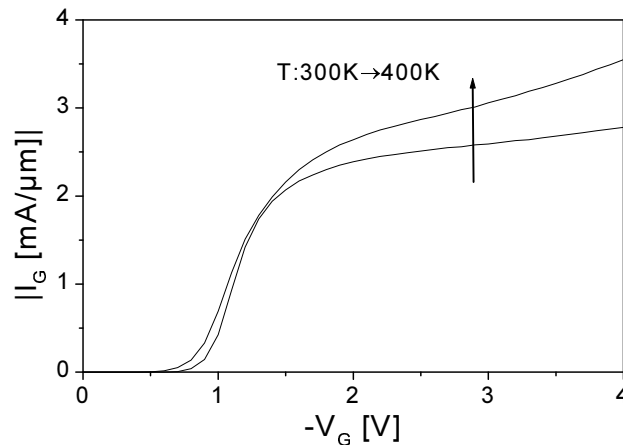


Bild 3.9: Atlas-Simulation in kartesischen Koordinaten der Struktur aus Bild 3.5(a) mit Aluminium im Durchbruchspfad. Die Austrittsarbeit des Aluminiums wurde mit 4.15eV, die Dielektrizitätskonstante des Isolators mit 7.5 angesetzt, welches sich aus der Schichtfolge des Aluminiumoxids mit dem Siliziumoxid ergibt, der Durchbruchspfad besteht vollständig aus Aluminium, $t_{ox} = 4\text{nm}$, $N_{Substrat} = 3 \cdot 10^{16}\text{cm}^{-3}$, $N_{Poly} = 5 \cdot 10^{20}\text{cm}^{-3}$, verwendete Modelle: mos, impact selb [34]).

ten, daß das Resultat nicht die Folge der unterschiedlichen Dotierungskonzentration oder Dielektrizitätskonstante ist. Die Simulation darf allerdings nur als Indiz gewertet werden, um die Annahme eines Metall-Halbleiter-Übergangs zu erhärten. Im realen Bauelement muß davon ausgegangen werden, daß der Durchbruchkanal aus einem Materialgemisch besteht und eine hohe Störstellendichte vorhanden ist. Die Simulation erfolgte überdies mit dem Drift-Diffusionsmodell und dem einfachen Schottky-Modell [34]. Trotzdem zeigt das Ergebnis die Tendenz, vor allem weil die Simulation eines Schottky-Übergangs ohne angrenzende MOS-Kapazität einen Kreuzungspunkt der temperaturabhängigen Kennlinien wie beim pn-Übergang ergibt. Die beim Durchbruch beobachtete positive Temperaturabhängigkeit ist also nicht Produkt des reinen Schottky-Übergangs, sondern der Kombination aus Schottky-Übergang mit paralleler MOS-Kapazität. In der Literatur sind zahlreiche Bauelemente vorgestellt worden, die die Vorteile von parallelgeschalteten Elementen und dem daraus folgenden Einfluß auf die Ladungsträgerkonzentration und den Leitungsmechanismus ausnutzen [43–45]. Auch im vorliegenden Fall muß von einer solchen Beeinflussung ausgegangen werden. Die Situation ist dem MOS-gesteuerten pn-Übergang ähnlich [46], im hier betrachteten Fall ist jedoch die Kathode der Diode mit dem Gate der MOS-Kapazität kurzgeschlossen.

Die I-U-Kennlinie und deren Temperaturabhängigkeit sind stark abhängig von der Austrittsarbeit des Metalls. Dieser Parameter bestimmt die Barrierenhöhe am Metall-Halbleiter-Übergang und die Ladungsträgerverteilung an der Grenzfläche. Da der größte Unterschied zwischen Schottky- und pn-Übergang bei hohen negativen Spannungen auftritt, wird die Situation bei $V_G = -4V$ betrachtet. Bild 3.10 zeigt die Ladungsträgerkonzentration und die Raumladung im Halbleiter bei $y=1\text{nm}$ horizontal entlang der Grenzfläche (s. Bild 3.5(a)). Weit entfernt vom Durchbruch befindet sich das Substrat bei negati-

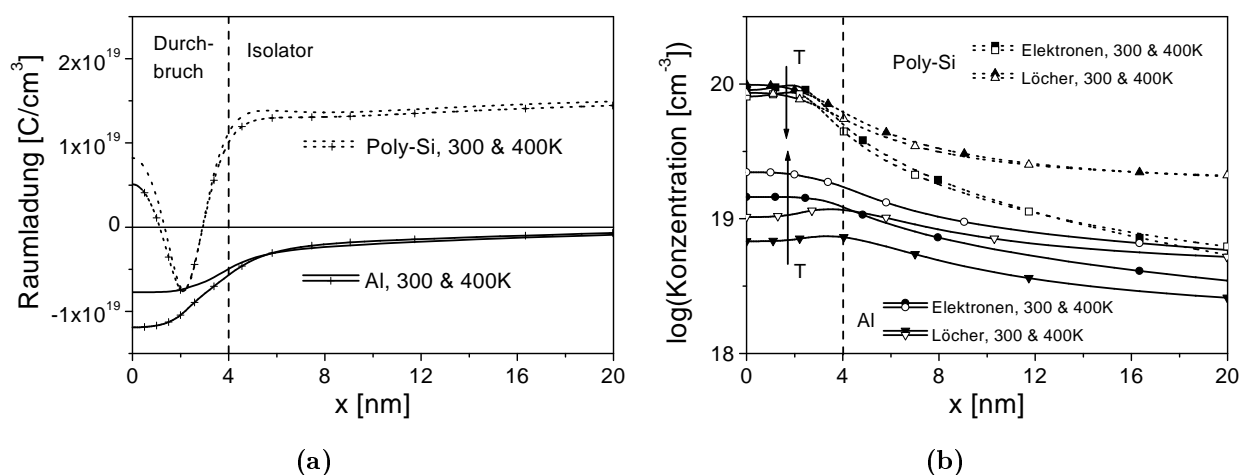


Bild 3.10: (a) Simulierte Raumladung und (b) Elektronen- und Löcherkonzentration bei $y=1\text{nm}$ (s. Bild 3.5(a)), $V_G=-4V$, $T=300\text{K}$ und $T=400\text{K}$ bei einem aus Aluminium oder Polysilizium bestehenden Durchbruch.

ven Gatespannungen in Akkumulation, es herrscht eine hohe Löcherkonzentration vor. Beim pn-Übergang ist dieser Zustand fast ganz bis zum Durchbruch unverändert. Beim Schottky-Übergang entsteht jedoch durch die Austrittsarbeitdifferenz zwischen Metall und Halbleiter eine Inversionsschicht am und um den Durchbruch, die weit in die ungeschädigte MOS-Kapazität hineinreicht. Diese unterschiedliche Ladungsträgerverteilung ist wahrscheinlich verantwortlich für die leicht unterschiedlichen Temperaturabhängigkeiten. Wie Bild 3.10(b) zeigt, nimmt die Ladungsträgerkonzentration beim Schottky-Übergang mit der Temperatur zu, so daß auch ein höherer Stromfluß zustandekommt, während die Konzentration beim pn-Übergang abnimmt und so ein niedrigerer Strom bei höheren Temperaturen fließt. Wird die Austrittsarbeit des Metalls erhöht, so nimmt die Inversionsschicht unterhalb des Metalls ab, bis ähnlich wie beim pn-Übergang eine Akkumulationsschicht entsteht. Das Temperaturverhalten entspricht dann dem des pn-Übergangs. Wird die Austrittsarbeit erniedrigt, so wird die Inversionsschicht verstärkt und ein Zwischenbereich

mit zwei Kreuzungspunkten in der Temperaturabhängigkeit entsteht. Aufgrund dieser starken Abhängigkeit von der Austrittsarbeit könnten sich möglicherweise experimentell unterschiedliche I-U-Kennlinien ergeben.

Zusammenfassung

Das Gesamtverhalten einer MOS-Kapazität nach hartem Durchbruch kann durch die Annahme eines dünnen leitenden Pfades zwischen Gateelektrode und Substrat verstanden werden. Die I-U-Kennlinie wird durch die Kontaktierung der beiden Elektroden bestimmt, der Durchbruchskanal selber beeinflusst die Kennlinie nur sekundär. Dies erfolgt einerseits durch die Durchbruchfläche und andererseits durch das Material. Je nach Materialart ergibt sich ein Schottky-Kontakt, ein pn-Übergang oder ein ohmscher Kontakt vom Durchbruchskanal zum Substrat. Dieser Übergang, parallel zur ungeschädigten MOS-Kapazität, ergibt letztendlich die Kennlinie. Die Superposition aus den durch die MOS-Kapazität bereitgestellten Ladungsträgern und dem Transportmechanismus am Durchbruch ergeben das Gesamtverhalten. Die Durchbruchfläche kann mit 10^{-17} bis 10^{-16}m^2 angesetzt werden, was einem Durchbruchsdurchmesser von einigen Nanometern entspricht. Dieser Wert stimmt mit Literaturangaben überein [47, 48].

3.1.2 Weicher Durchbruch

Der weiche Durchbruch wird bei Oxiddicken unter 5nm beobachtet. Er stellt sowohl hinsichtlich seiner Spannungsabhängigkeit als auch seiner Stromhöhe eine Zwischenstufe zwischen frischen bzw. leicht gestreßten und hart durchgebrochenen Oxiden dar. Unterschiedliche Modelle sind zu seiner Erklärung herangezogen worden, die bekanntesten sind das Variable Range Hopping Modell [49], die Perkolationstheorie [50, 51] und das Quantenpunktkontaktmodell [52]. Bild 3.11 zeigt gemessene typische I-U-Kennlinien von MOS-Kapazitäten in Akkumulation. Da der Durchbruch zufällig auftritt, unterliegt auch die Stromhöhe erheblichen statistischen Schwankungen von Bauelement zu Bauelement. Deutlich wird aus Bild 3.11, daß die Spannungs- und Temperaturabhängigkeit bei höheren Spannungen für die beiden Substrattypen gleich ist. Bei niedrigen Spannungen sieht man einen Verlauf nach einer Potenzfunktion, der zu hohen Spannungen hin exponentialförmig wird. Bei Verschiebung der temperaturabhängigen Stromkurven auf der Stromachse liegen diese übereinander (s. Bild 3.12(a)), die Form der Kurve ist also unverändert und wird allein von der Transmissionswahrscheinlichkeit durch den Durchbruch festgelegt. Bei Verschiebung auf der Spannungsachse hingegen (Bild 3.12(b)) liegen die Kurven nicht mehr

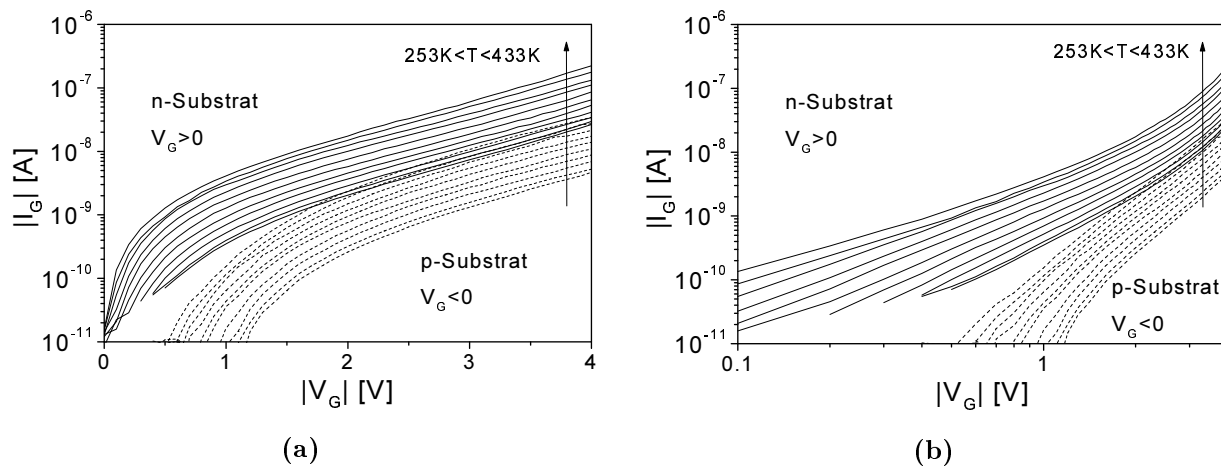


Bild 3.11: Gemessene I-U-Kennlinien von MOS-Kapazitäten in Akkumulation nach weichem Durchbruch bei unterschiedlichen Temperaturen. Strukturen aus einem Infineon $0.25\mu\text{m}$ dual-gate Prozeß mit n+Poly-Gate auf p-Substrat bzw. p+Poly-Gate auf n-Substrat, $t_{ox} = 4.4\text{nm}$, $N_{Substrat} = 3 \cdot 10^{17}\text{cm}^{-3}$, $N_{Poly} = 7 \cdot 10^{19}\text{cm}^{-3}$, $A_G = 0.3\text{mm}^2$. (a) Halblogarithmische und (b) doppeltlogarithmische Darstellung.

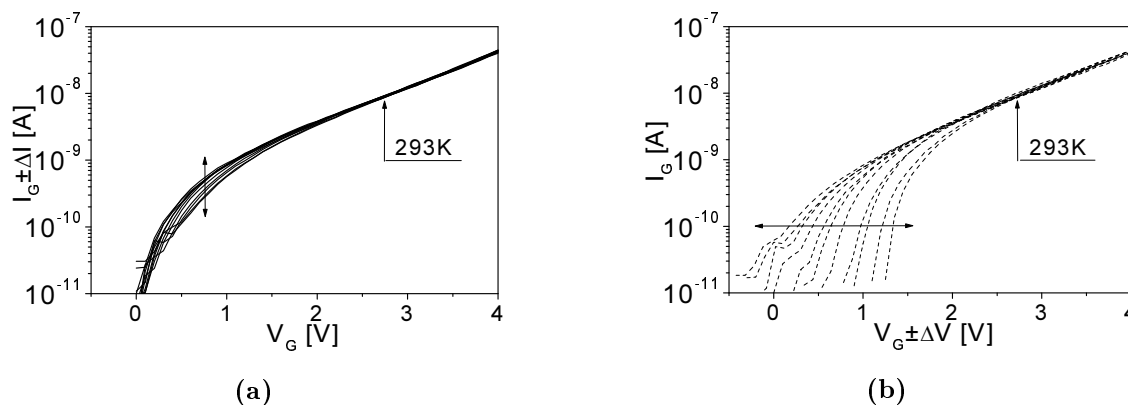


Bild 3.12: Verschiebung der I-U-Kennlinien der MOS-Kapazität mit n-Substrat aus Bild 3.11 auf der (a) Stromachse, (b) Spannungsachse, mit der Kennlinie bei 293K als Referenz.

übereinander, die Temperaturverschiebung ist also nicht gleichwertig mit einer Spannungsänderung. Der Stromtransport wird im Gegensatz zu den harten Durchbrüchen sehr stark durch die Eigenschaften des Durchbruchpfades bestimmt.

Beim Punktkontaktmodell wird der weiche Durchbruch, wie Bild 3.13 zeigt, als Effekt der Quantisierung in einem engen Potentialtopf erklärt, der sich aus dem leitenden Durchbruchspfad, umgeben von den hohen Wänden des Oxids, ergibt. Es bilden sich im

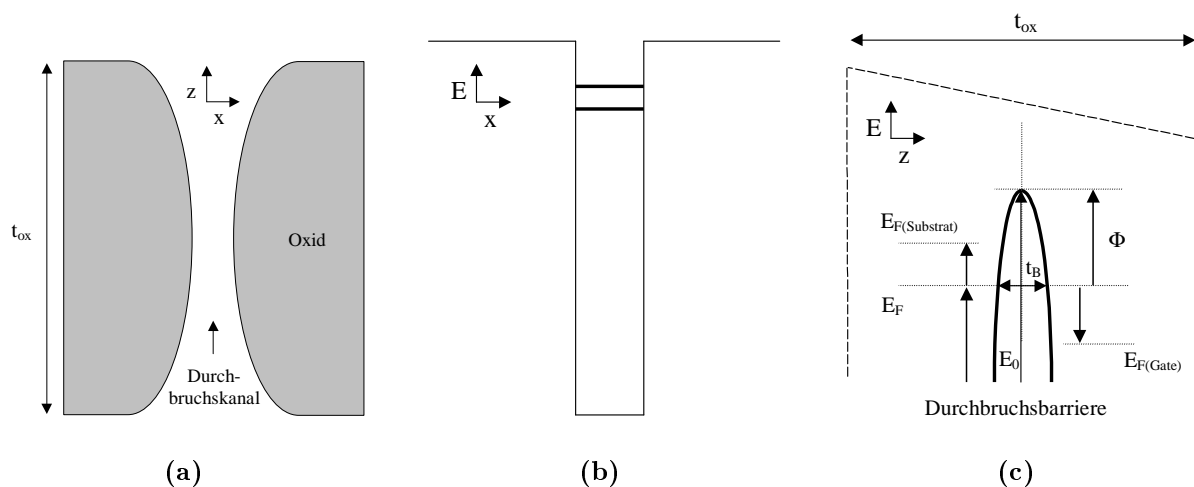


Bild 3.13: Weicher Durchbruch nach dem Punktkontaktmodell [52]: (a) Querschnitt mit umgebendem intaktem Oxid, Energiediagramm mit Durchbruchbarriere an der Verengung des leitenden Pfades (b) in x -Richtung mit zwei Energiebändern und (c) in z -Richtung bei Betrachtung des untersten Bandes für den Stromtransport.

Quantentopf Energiebänder aus. Elektronen mit niedrigerer Energie als dem niedrigsten Band im Quantentopf können dann die sich so einstellende Barriere nur durch einen Tunnelmechanismus überwinden. Im Unterschied zum direkten oder Fowler-Nordheim-Tunneln ergibt sich die Barriere nicht aus den Materialeigenschaften des Isolators sondern aus den geometrischen Eigenschaften des leitenden Pfades durch das umgebende Oxid. Die Barriere wird hierbei, wie in Bild 3.13(c) gezeigt, als parabelförmig angesehen.

Das Punktkontaktmodell wurde in Zusammenarbeit mit E. Miranda und D. Schröder um die Temperaturabhängigkeit und eine verbesserte Beschreibung der Potentialverteilung über dem Durchbruchkanal erweitert, die im Folgenden vorgestellt werden. Die Kenntnis der Temperaturabhängigkeit ist dabei für den Einsatz in Schaltungssimulatoren von großer Bedeutung.

Der Tunnelstrom kann analog zu Gleichung 3.1 geschrieben werden als

$$I = \frac{2q}{h} \cdot \int_{-\infty}^{\infty} D(E) \cdot [F(E - E_{\text{F(Substrat)}}) - F(E - E_{\text{F(Gate)}})] dE, \quad (3.10)$$

wobei sich die Quasi-Fermi-Energien auf Substrat- und Gateseite $E_{\text{F(Substrat)}}$ und $E_{\text{F(Gate)}}$ aus der Verteilung der angelegten Spannung V_G auf die beiden Seiten der Barriere ergeben. Beim Punktkontaktmodell wird die transversale Energie E_t vernachlässigt. Die Transmissionswahrscheinlichkeit durch die parabelförmige Barriere ist [52]

$$D(E) = \frac{1}{1 + \exp[-\alpha(E - E_0)]} \approx \exp[\alpha(E - E_0)], \quad (3.11)$$

wobei die Näherung für $E < E_0 - 3/\alpha$ gilt. Der Krümmungsparameter α ergibt sich aus der zweiten Ableitung der Barriere an ihrem Maximum E_0 als [53–55]

$$\alpha = \frac{\pi t_B}{\hbar} \sqrt{\frac{m}{2\Phi}}, \quad (3.12)$$

mit $\hbar = h/2\pi$, t_B der Barrierenbreite und $\Phi = E_0 - E_F$ der Barrierenhöhe. Für den Strom folgt

$$I = \frac{2q}{h} \cdot \int_{-\infty}^{\infty} \left[\frac{\exp[\alpha(E - E_0)]}{1 + \exp\left(\frac{E - E_{\text{F(Substrat)}}}{kT}\right)} - \frac{\exp[\alpha(E - E_0)]}{1 + \exp\left(\frac{E - E_{\text{F(Gate)}}}{kT}\right)} \right] dE. \quad (3.13)$$

Gleichung 3.13 kann nach Variablentransformation $u = \exp[(E - E_X)/kT]$ mit Hilfe des exakten Integrals [56]

$$\int_0^{\infty} \frac{u^{\alpha kT - 1}}{u + 1} du = \frac{\pi}{\sin(\pi kT \alpha)} \quad (\text{für } \alpha < 1/kT) \quad (3.14)$$

aufgelöst werden und ergibt

$$I = \frac{2q}{h} \cdot \exp[-\alpha E_0] \cdot \frac{\pi kT}{\sin(\pi kT \alpha)} \cdot \left\{ \exp[\alpha E_{\text{F(Substrat)}}] - \exp[\alpha E_{\text{F(Gate)}}] \right\}. \quad (3.15)$$

Diese Lösung gilt aufgrund der in Gleichung 3.11 angewendeten Näherung für $E_0 > E_{\text{F(Substrat)}}$ bzw. $E_{\text{F(Gate)}}$, also solange die Spitze der Barriere über den Fermineaus auf Kathoden- und Anodenseite liegt.

Modellierung der Potentialverteilung über den Elektroden

Im einfachsten Fall kann davon ausgegangen werden, daß sich die angelegte Spannung symmetrisch auf die beiden Seiten der Barriere verteilt. In der Tat ist es schwierig, die tatsächliche Potentialverteilung über dem Durchbruch zu kennen. Im Gegensatz zum Fowler-Nordheim-Tunnelstrom ist der weiche Durchbruchstrom um $V_G = 0\text{V}$ fast symmetrisch, wenn z. B. durch Beleuchtung [57] oder durch Verwendung von Transistoren (s. auch Bild 3.27 in Abschnitt 3.2.2) ausreichend Ladungsträger im Substrat zur Verfügung stehen. Eine Verschiebung um die Flachbandspannung wird nicht beobachtet, ebenso ist der Verlauf bei p- und n-dotierten Substraten ähnlich [57, 58]. Diese Unabhängigkeit von der Substratdotierung und Polarität der angelegten Spannung ist überraschend und führte bisher zu der vereinfachten symmetrischen Darstellung.

Um die Potentialverteilung genauer zu modellieren, muß die Lage des Durchbruchs [59] und das umliegende MOS-System berücksichtigt werden. Hierfür wird die Bandverbiegung im Substrat zur Berechnung der Oxidspannung nach

$$V_{ox} = V_G - \psi_s - \phi_{MS}, \quad (3.16)$$

miteinbezogen, wobei ϕ_{MS} die Austrittsarbeitsdifferenz zwischen Gate und Substrat ist. Die Verarmung im Polysilizium-Gate wird vernachlässigt. Unter der Voraussetzung, daß das Oberflächenpotential des MOS-Systems durch den weichen Durchbruch nicht signifikant verändert wird, ergibt sich Bild 3.14.

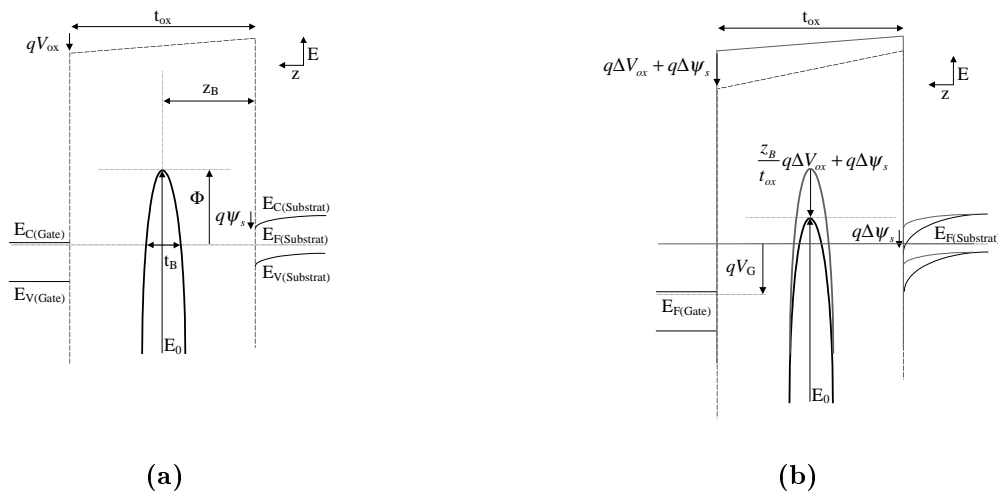


Bild 3.14: (a) Energiediagramm eines n+/p-MOS-Systems mit weichem Durchbruch bei 0V, (b) Verschiebung bei Anlegen einer positiven Gatespannung.

Es lassen sich unter Bezug auf $E_{\text{F(Substrat)}}$ folgende Energiegleichungen aufstellen:

$$E_{\text{F(Substrat)}} = 0 \quad (3.17)$$

$$E_{\text{F(Gate)}} = qV_G \quad (3.18)$$

$$E_0 = \Phi + q\Delta\psi_s + \frac{z_B}{t_{ox}}q\Delta V_{ox} = \Phi + \frac{z_B}{t_{ox}}qV_G + \left(1 - \frac{z_B}{t_{ox}}\right)q\Delta\psi_s. \quad (3.19)$$

Wenn ballistischer Transport ohne Energieverlust von Substrat bzw. Gate bis zur Barriere vorausgesetzt wird, können diese Energiegleichungen zur Berechnung des Tunnelstroms verwendet werden. Durch Einsetzen in Gleichung 3.15 ergibt sich mit $\Phi = E_0 - E_{\text{F(Substrat)}}$ für den Durchbruchstrom

$$I = \frac{2q}{h} \cdot \frac{\pi kT}{\sin(\pi kT\alpha)} \cdot \exp[-\alpha\Phi(T)] \cdot \exp[-\alpha\beta q\Delta\psi_s] \cdot \{\exp[\alpha\beta qV_G] - \exp[-\alpha(1-\beta)qV_G]\} \quad (3.20)$$

mit $\beta = 1 - z_B/t_{ox}$. Über den Parameter β wird somit die Entfernung der Barriere von der Silizium-Siliziumdioxid-Grenzfläche berücksichtigt. Bei Vernachlässigung der Bandverbiegung ($\Delta\psi_s = 0$) folgt die vereinfachte Form

$$I = \frac{2q}{h} \cdot \frac{\pi kT}{\sin(\pi kT\alpha)} \cdot \exp[-\alpha\Phi(T)] \cdot \{\exp[\alpha\beta qV_G] - \exp[-\alpha(1-\beta)qV_G]\}. \quad (3.21)$$

Es ergeben sich dann allerdings unterschiedliche Parametersätze, wobei Gleichung 3.21 ohne Bandverbiegung einfacher einsetzbar ist.

Temperaturabhängigkeit

In den Gleichungen 3.20 und 3.21 ist die Temperaturabhängigkeit über den Term $\frac{\pi kT}{\sin(\pi kT\alpha)}$ in der Verschmierung der Fermi-Dirac-Verteilung enthalten, die aber nur bei niedriger Barrierenhöhe eine Rolle spielt und ansonsten vernachlässigbar ist. Zusätzlich muß die Temperaturabhängigkeit der Barrierenhöhe selber betrachtet werden. Dieses Phänomen kann aus der Vibration der Atome und dem resultierenden durchschnittlichen makroskopisch gemessenen Wert der Barriere erklärt werden [60]. Wird das Potential entlang einer intakten Si-SiO₂-Grenzschicht mikroskopisch betrachtet, so entsteht aus der Superposition der Atompotentiale, wie in Bild 3.15 gezeigt, eine ortsabhängige Barrierenhöhe. Es ergeben sich also lokale Maxima und Minima, die abhängig von der Entfernung der Atomrümpfe zueinander sind. Der größte Anteil des Stroms fließt durch die Minima hindurch. Dies

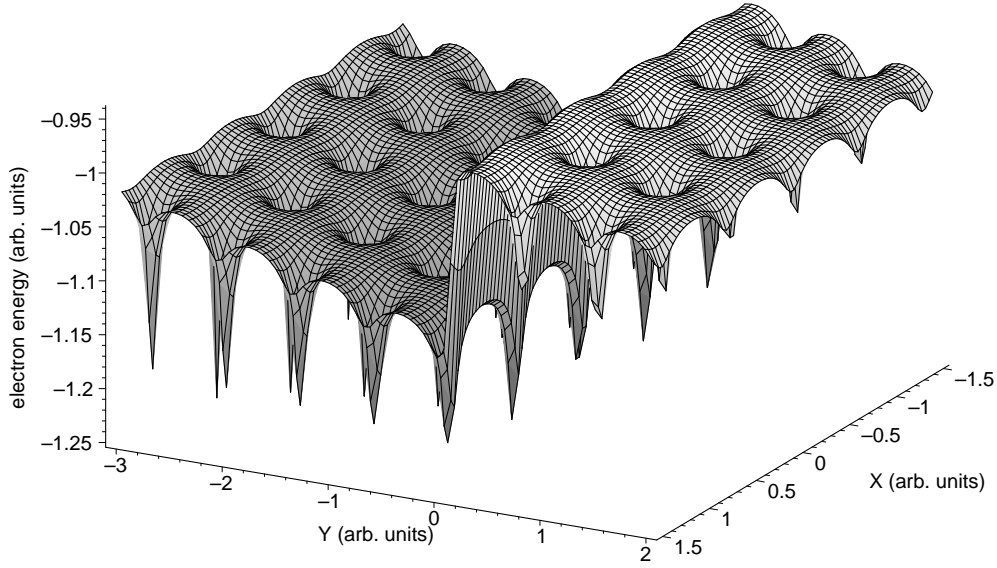


Bild 3.15: Potentialtopographie an der Si-SiO₂-Grenzschicht [60].

ist also der Wert, der makroskopisch durch die Strommessung zugänglich ist. Wird das Gitter erwärmt, beginnt eine thermische Bewegung der Atomkerne. Durch diese Bewegung verschieben sich die Maxima und Minima in ihrer Höhe, die Barrierenhöhe fluktuiert um einen Mittelwert Φ_0 , der den Wert am absoluten Nullpunkt darstellt. Wegen der exponentiellen Abhängigkeit der Stromhöhe von der Barrierenhöhe ist die Zunahme des Stroms, während die Barriere gerade niedrig ist, höher als die Abnahme des Stroms bei großer Höhe. Es ergibt sich also durch die thermische Bewegung im Mittel ein höherer Strom. Die gleiche Argumentation kann auch für die Barrierenhöhe des weichen Durchbruchpfades angewandt werden. Die mikroskopische Barrierenhöhe Φ_B und Krümmungsparameter α_B ergeben sich aus der Position der einzelnen Atome am Ort x_i , die den Durchbruchpfad bestimmen. Wird davon ausgegangen, daß die normierte Auslenkung der Atome Δx_i um den Mittelwert klein ist, so kann das Produkt $\alpha_B \Phi_B$ durch eine Taylorentwicklung um die Werte Φ_0 und α am absoluten Nullpunkt angenähert werden:

$$\begin{aligned} \alpha_B \Phi_B &= \left(\alpha - \sum_i \chi_i \Delta x_i \right) \cdot \left(\Phi_0 - \sum_i \zeta_i \Delta x_i \right) \\ &\approx \alpha \Phi_0 - \alpha \sum_i \zeta_i \Delta x_i - \Phi_0 \sum_i \chi_i \Delta x_i = \alpha \Phi_0 - \sum_i \Delta x_i (\alpha \zeta_i + \Phi_0 \chi_i), \end{aligned} \quad (3.22)$$

wobei ζ_i und χ_i Konstanten sind und Terme höherer Ordnung vernachlässigt werden. Für

die Auslenkung Δx_i kann die Boltzmann-Verteilung um den Wert Null angenommen werden,

$$p(\Delta x_i) = \sqrt{\frac{A_i}{2\pi T}} \exp\left(-\frac{A_i}{2T} \Delta x_i^2\right), \quad (3.23)$$

hierbei ist T/A_i die Standardabweichung der Auslenkung. Der mittlere Strom durch die Barriere in den Gleichungen 3.20 und 3.21 wird durch den Exponentialterm nach $I(T) \sim \exp[-\alpha\Phi(T)]$ bestimmt. Dabei ist $\Phi(T)$ die durch Messung makroskopisch zugängliche mittlere Barrierenhöhe bei Temperatur T im Gegensatz zu den fluktuierenden mikroskopischen Barrierenparametern Φ_B und α_B . Bei diesem Ansatz wird also der höhere Strom, der sich bei Erwärmung durch die Bewegung der Barriere ergibt, auf eine makroskopisch konstante, aber temperaturabhängige Barrierenhöhe $\Phi(T)$ zurückgeführt. Der mittlere Strom ist dann

$$\begin{aligned} I(T) &\sim \exp[-\alpha\Phi(T)] = \overline{\exp(-\alpha_B\Phi_B)} \\ &= \prod_i \int_{-\infty}^{\infty} \exp[-\alpha\Phi_0 + \Delta x_i(\alpha\zeta_i + \Phi_0\chi_i)] \cdot \sqrt{\frac{A_i}{2\pi T}} \exp\left(-\frac{A_i}{2T} \Delta x_i^2\right) d^3 \Delta x_i. \end{aligned} \quad (3.24)$$

Das Integral kann mit Hilfe von Maple wie folgt gelöst werden:

$$\sqrt{\frac{A_i}{2\pi T}} \cdot \int_{-\infty}^{\infty} \exp\left(\Delta x_i(\alpha\zeta_i + \Phi_0\chi_i) - \frac{A_i}{2T} \Delta x_i^2\right) d^3 \Delta x_i = \exp\left(\frac{\alpha^2\zeta_i^2 T}{2A_i} + \frac{\Phi_0^2\chi_i^2 T}{2A_i}\right) \quad (3.25)$$

Da das Integral für alle Raumrichtungen durchgeführt werden muß, ergibt sich als endgültiges Resultat der dreifache Wert und als Endergebnis [60]

$$I(T) \sim \exp[-\alpha\Phi(T)] = \exp\left[-\alpha\left(\Phi_0 - T \cdot \frac{3}{\alpha} \sum_i \frac{\alpha^2\zeta_i^2 + \Phi_0^2\chi_i^2}{2A_i}\right)\right]. \quad (3.26)$$

Die Temperaturabhängigkeit der makroskopischen Barriere für die Anwendung in dem Modell ergibt sich somit zu

$$\Phi(T) = \Phi_0 + \gamma T, \quad (3.27)$$

mit dem Temperaturkoeffizienten $\gamma \approx -10^{-3} \text{eVK}^{-1}$.

Vergleich mit Meßdaten

Die Beziehungen zwischen den Parametern des Modells können durch Ableitung der Meßdaten gewonnen werden. Bei hohen Spannungen ist der Term $\exp[-\alpha(1-\beta)qV_G]$ sowie

die Spannungsabhängigkeit der Bandverbiegung vernachlässigbar. Wird auch der Term $\frac{\pi}{\sin(\pi kT\alpha)}$ vernachlässigt, ergeben sich die Ableitungen der Gleichungen 3.20 und 3.21 nach der Spannung und nach der Temperatur bei hohen Spannungen zu:

$$\frac{d(\log[I/1A])}{dV_G} \cdot \ln(10) = \alpha\beta q \quad (3.28)$$

$$\frac{d(\log[I/1A])}{dT} \cdot \ln(10) = -\alpha\gamma \quad (3.29)$$

Bild 3.16 zeigt die Anwendung auf die Meßkurven. Die konstante Spannungs- und Temperaturabhängigkeit des Stroms bei hohen Gatespannungen wird in den Meßdaten wiedergefunden und somit die Beziehung zwischen den Parametern festgelegt. Bild 3.16(b)

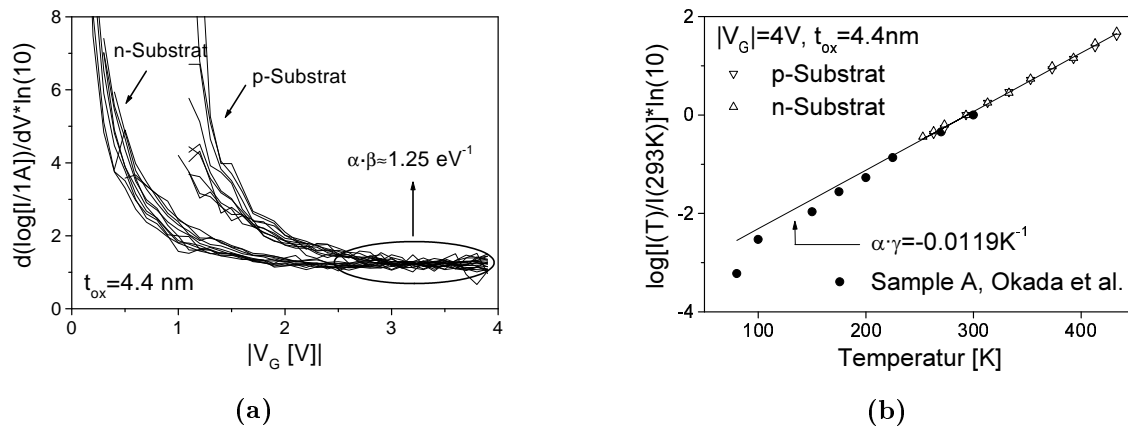


Bild 3.16: Ableitung der Meßdaten nach der (a) Spannung und (b) Temperatur und Vergleich mit Messungen von Okada et al. bei niedrigen Temperaturen [49].

illustriert sehr deutlich die exponentielle Abhängigkeit von der Temperatur bei hohen Spannungen, die Steigung ist sowohl bei unterschiedlichen Oxiddicken (2.9 und 4.4nm), als auch bei unterschiedlichen Substrattypen gleich und deckt sich mit Messungen von Okada et al. bei niedrigen Temperaturen [49, 61].

Zur Berechnung der Bandverbiegung und Anwendung von Gleichung 3.20 wird das Oberflächenpotential im Substrat in Inversion mit einer Näherung von van Langevelde und Klaassen [62] und in Akkumulation nach Pavan, Tsvividis und Nagaraj bestimmt [63]. Ebenfalls werden temperaturabhängige Näherungen für E_g und n_i nach Sze [38] bzw. Nicollian und Brews [36] verwendet, diese spielen jedoch nur eine untergeordnete Rolle.

Mit den in Tabelle 3.1 gezeigten Parametersätzen ergeben sich die in Bild 3.17 dargestellten Verläufe. Gute Übereinstimmung ist für beide Modellgleichungen zu sehen. Im

Niederspannungsbereich ergeben sich in beiden Fällen leichte Abweichungen von den Meßdaten. Dies deutet darauf hin, daß die Potentialverteilung hier wahrscheinlich noch schwieriger zu modellieren ist als angenommen.

Tabelle 3.1: Parameter für die Ergebnisse von Bild 3.17.

Parameter	Mit Bandverbiegung Gleichung 3.20	Ohne Bandverbiegung Gleichung 3.21
α [eV ⁻¹]	5.6	4.1
β	0.8	0.3
γ [meVK ⁻¹]	2.4	2.9
$\Phi(293\text{K})$ [eV]	2.8	2.7
$N_{Substrat}$ [cm ⁻³]	$3 \cdot 10^{17}$	-
N_{Poly} [cm ⁻³]	$7 \cdot 10^{19}$	-
t_{ox} [nm]	4.4	-

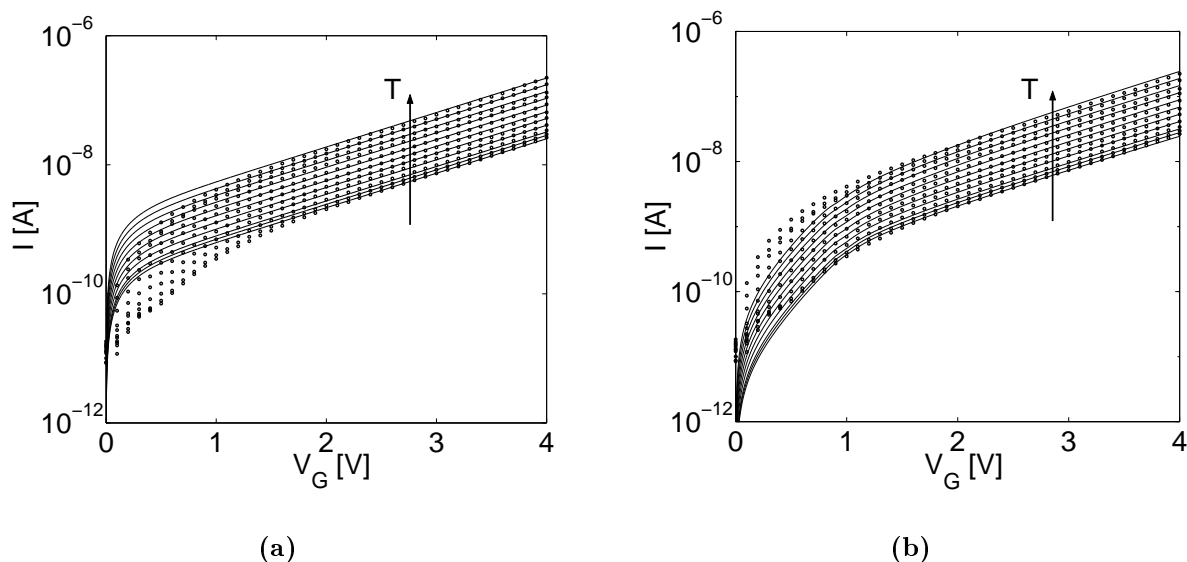


Bild 3.17: Vergleich der Ergebnisse von (a) Gleichung 3.21 ohne Bandverbiegung und (b) Gleichung 3.20 mit Bandverbiegung mit den Parametern von Tabelle 3.1 für die p+/n-MOS-Kapazität. Messung (Punkte) und Modell (durchgezogene Linien) bei $253\text{K} < T < 433\text{K}$.

3.1.3 Fazit aus den Messungen an MOS-Kapazitäten

Zusammenfassend ergeben sich die in Tabelle 3.2 gezeigten Unterscheidungsmerkmale für die Durchbrüche an den MOS-Kapazitäten. Insgesamt läßt sich sagen, daß der harte Durchbruch mit großer Sicherheit als Kontakt zwischen den beiden Elektroden der Kapazität

Tabelle 3.2: Überblick der Durchbrüche bei MOS-Kapazitäten

Harter Durchbruch		Weicher Durchbruch
Strom ist <i>elektrodenbegrenzt</i> Hohe Stromdichte		Strom ist <i>durchbruchsbegrenzt</i> Mittlere Stromdichte
Bei gleicher Dotierung der Elektroden:	Bei entgegengesetzter Dotierung der Elektroden:	Wenig abhängig von Elektrodendotierung und Polarität der Gatespannung
linear (Potenzfunktion 1. Ordnung)	Exponentialförmig bei niedrigen, linear bei höheren Spannungen	Potenzfunktion bei niedrigen, Exponentialfunktion bei höheren Spannungen
↓ Ohmscher Verlauf	↓ Diodenartiger Verlauf (mit Reihenwiderstand)	↓ Quantenpunktkontaktmodell

betrachtet werden kann, während der weiche Durchbruch noch eine Barriere zwischen den Elektroden darstellt. In beiden Fällen handelt es sich um ein lokalen Durchbruchspfad inmitten einer weitgehend intakten MOS-Struktur. Für MOS-Kapazitäten können somit je nach Durchbruchstyp die in Bild 3.18 gezeigten Ersatzschaltbilder definiert werden. Das Quantenpunktkontaktmodell wurde erfolgreich erweitert, so daß eine physikalisch fundierte Beschreibung des Durchbruchstroms in Abhängigkeit von Spannung und Temperatur möglich ist.

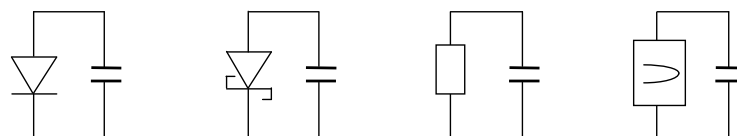


Bild 3.18: Ersatzschaltbilder für durchgebrochene MOS-Kapazitäten. Der Durchbruch liegt parallel zur restlichen intakten MOS-Kapazität und stellt einen pn-Übergang, Schottky-Übergang, ohmschen Kontakt oder Quantenpunktkontakt dar.

3.2 Durchbrüche an MOS-Transistoren

Die Vierpolnatur des Transistors macht eine Auswertung der Messergebnisse sehr viel komplexer, da die Einflüsse aller Elektroden und deren jeweilige Dotierung berücksichtigt werden müssen. Im Folgenden werden nur dual-gate Transistoren betrachtet, deren Gate-Dotierung stets der Bulk-Dotierung entgegengesetzt und somit gleich der Dotierung der Source-/Drain-Gebiete ist (s. auch Bild 3.23(a) auf S. 56).

3.2.1 Harter Durchbruch

Drei mögliche markante Orte wurden ausgewählt, an denen, wie in Bild 3.19 gezeigt, ein harter Durchbruch durch Wegnahme des Oxids mit Atlas simuliert wurde. Der Durchbruch

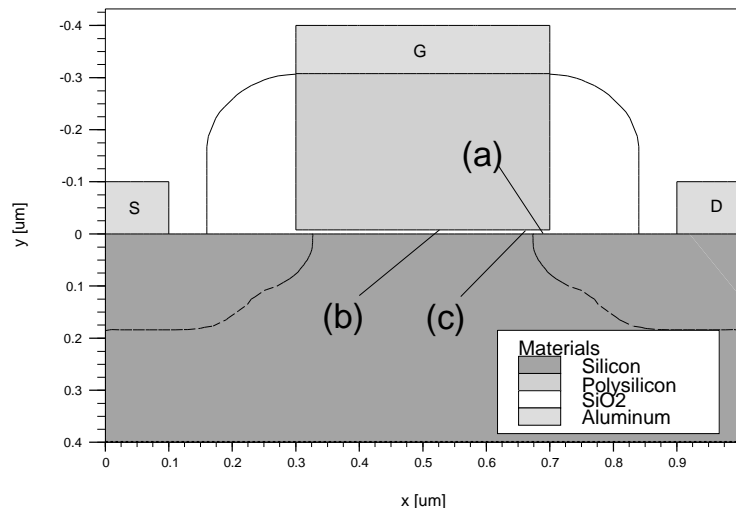


Bild 3.19: Mit Atlas simulierte Durchbruchsstellen in einem n-Kanal-Transistor, $t_{ox} = 4\text{nm}$, Länge des Durchbruchs 1nm , gleiche Dotierung wie das Substrat (Modelle mos, impact selv [34]).

im pMOS-Transistor sowie zur Source-Seite des Transistors wird als äquivalent angesehen. Da es sich hierbei nur um eine zweidimensionale Simulation handelt, wird der Durchbruch wie der Rest des Bauelements mit einer Weite von $1\mu\text{m}$ berechnet und ist auf der gesamten Weite des Bauelements durchgezogen. Die eingesetzten Drift-Diffusions-Gleichungen sind ebenfalls nur eingeschränkt auf den vorliegenden Fall anwendbar, so daß die Simulationsergebnisse als Näherung zu verstehen sind. Die Tatsache, daß die unterschiedlichen Verläufe in den Messungen wiedergefunden werden konnten, zeigt aber, daß die Simulation die grundlegenden Abhängigkeiten wiedergibt. Im Folgenden werden die drei unterschiedlichen Fälle anhand der Simulation und der Messungen gezeigt. Die Messungen lassen sich

in dem Raster der drei gewählten Orte gut klassifizieren. Die gute Übereinstimmung der Ergebnisse deutet an, daß man auch hier wie bei den Kapazitäten von einem Durchwachsen der Elektrode ausgehen kann.

(a) Durchbruch zum Drain-/Source-Überlappbereich

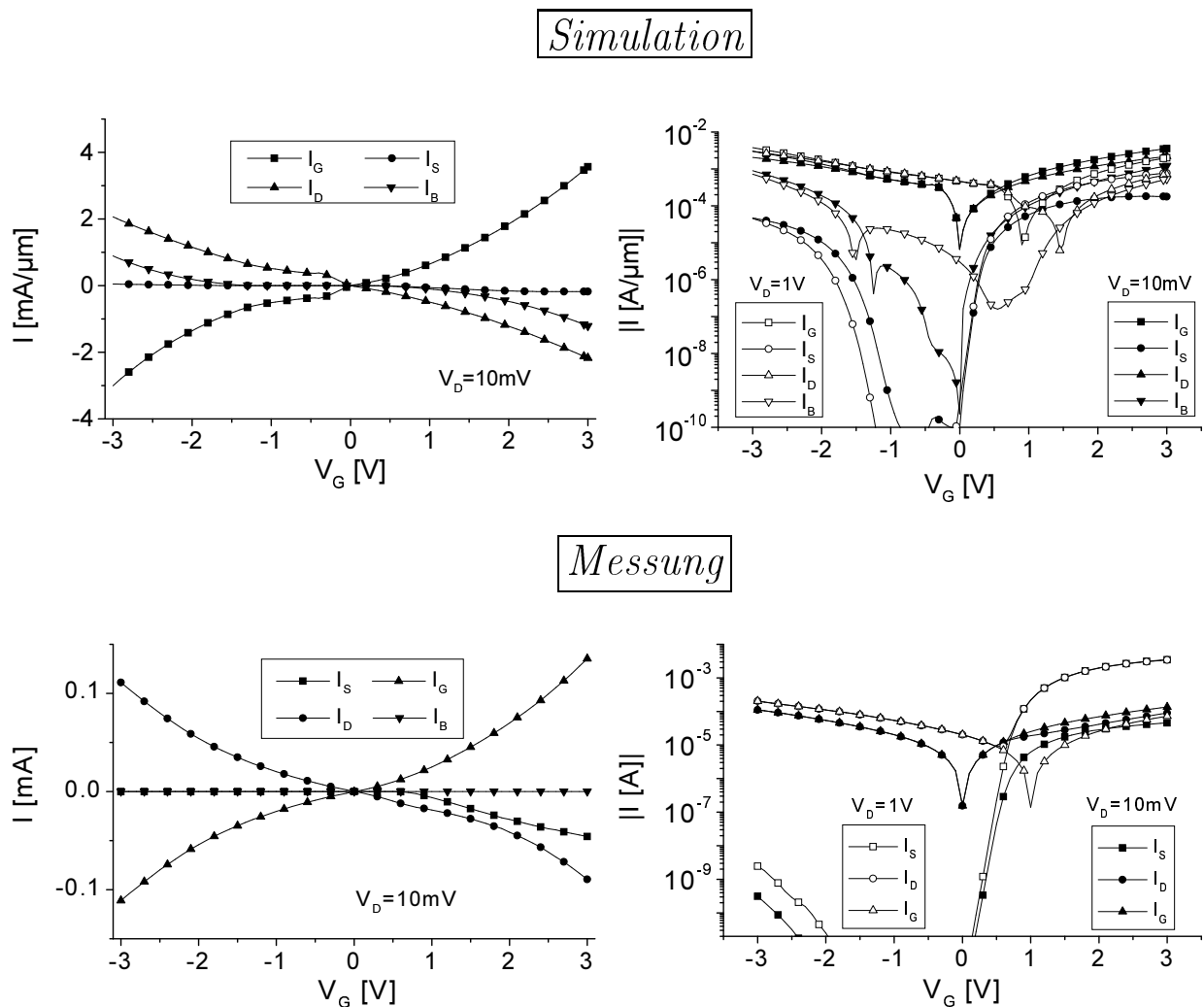


Bild 3.20: Simulation und Messung (nMOS-Transistor 03_1_T3, $L=0.35\mu\text{m}$, $W=10\mu\text{m}$, $t_{ox} = 4.4\text{nm}$) in linearer (links) und halblogarithmischer (rechts) Darstellung.

Durch die gleichartige Dotierung der Gate- und Drain-/Source-Elektroden ergibt sich hier der in Bild 3.20 gezeigte ohmsche Verlauf des Gatestroms, der sich vollständig im Drainstrom wiederfindet. Eine starke Abhängigkeit von der am Drain angelegten Spannung zeigt sich in der halblogarithmischen Darstellung. Das Minimum des Gatestroms liegt

bei $V_G = V_D$. Daß ein Teil des Transistors trotz der erheblichen parasitären Ströme funktioniert, ist am Source-Strom zu sehen, der weiterhin ein „Abschaltverhalten“ bei negativen Spannungen, und im aktiven Bereich bei hohen positiven Gatespannungen einen Feldefektstrom zeigt. In den Messungen wurden unterschiedliche Verläufe des Bulkstroms festgestellt, jedoch immer auf deutlich niedrigerem Niveau als dem Rest der Ströme. Je nach Nähe zum Kanal scheint ein geringer Teil des Stroms, möglicherweise durch Stoßionisation, zum Bulk abzufießen.

(b) Durchbruch zur Mitte des Kanals (zum Substrat)

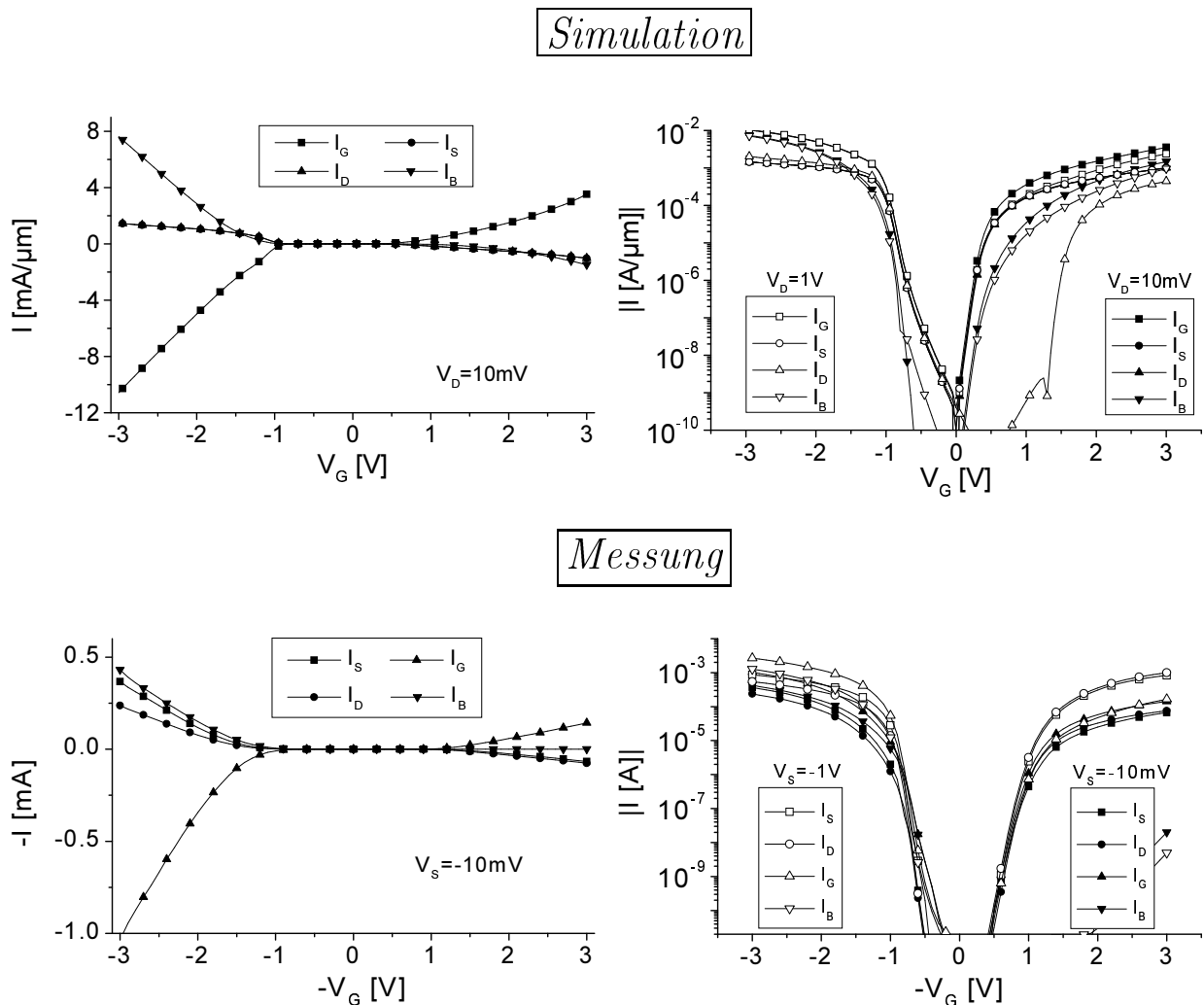


Bild 3.21: Simulation und Messung (pMOS-Transistor 03_2_T2p, $L=0.4\mu\text{m}$, $W=10\mu\text{m}$, $t_{ox} = 4.4\text{nm}$) in linearer (links) und halblogarithmischer (rechts) Darstellung.

Hier ist in Bild 3.21 ein diodenartiger Verlauf zu erkennen. Der Gatestrom fließt je nach Polarität der angelegten Gatespannung entweder zum Bulk bei Durchlaßbetrieb der Gate-Bulk-Diode oder aber Richtung Source und Drain bei Sperrbetrieb bzw. in Inversion des noch funktionierenden Resttransistors. Es ist praktisch keine Abhängigkeit von V_D oder V_S festzustellen, wie die halblogarithmische Darstellung zeigt. Die Messung zeigt einen pMOS-Transistor, der bis auf das Vorzeichen der Spannungen und Ströme gleiches Verhalten wie der simulierte nMOS-Transistor zeigt. Bei gesperrtem Transistor verhält sich der Gatestrom praktisch wie bei einer MOS-Kapazität, während bei eingeschaltetem Transistor durch die schnell bereitgestellte Inversionsschicht das Substrat vom Durchbruch abgeschirmt wird. Am Durchbruchspfad werden somit ausreichend Minoritätsladungsträger bereitgestellt und es kann ein im Vergleich zur reinen MOS-Kapazität hoher Strom zu den Source-/Drain-Gebieten fließen. Dadurch ergibt sich die um $V_G = 0$ praktisch symmetrische Gatestromkennlinie.

(c) Durchbruch zum Rand des Kanals (zum Substrat nahe der Drain-Elektrode)

Dieser in Bild 3.22 gezeigte Fall ist im Prinzip ähnlich dem vorhergehenden, allerdings ergibt sich durch die räumliche Nähe zu einem Minoritätsladungsträgerreservoir und durch die nicht abrupten pn-Übergänge an Source und Drain ein Zwischenverhalten. Auch beim gesperrten Transistor zeigt sich im Unterschied zum Durchbruch zur Kanalmitte ein Stromfluß nicht nur zum Substrat sondern auch zum Drain. Obwohl aber ein Teil des Stroms zum gleichartig dotierten Drain fließt, herrscht weiterhin diodenförmiges Verhalten vor. Das Minimum des Gatestroms wird bei $|V_G| < |V_D|$ erreicht, in gewisser Weise ist dies ein Indikator für die Entfernung vom Drain-Gebiet.

Zusammenfassung

Bei einem harten Durchbruch erfolgt eine sehr starke Änderung der Funktion des MOS-Transistors. Zusätzlich zum einsetzenden Gatestrom als dominierendem Kennzeichen verschlechtern sich Daten wie Schwellspannung und Steilheit. Das Bauelement hat jedoch seine Funktion nicht vollständig eingebüßt. Im aktiven Betrieb bei hohen Drainspannungen treibt der Transistor immer noch einen um Größenordnungen höheren Drainstrom als den parasitären Gatestrom (s. hierzu z. B. die Meßkurven bei $V_D = 1V$ in Bild 3.20). Tabelle 3.3 zeigt die wichtigsten Eigenschaften des Gateleckstroms je nach Ort des Durchbruchs. Die Kurvenform wird durch die Dotierung der durch den Durchbruch verbundenen Elektroden und durch den Betrieb in Akkumulation oder Inversion bestimmt. Bei Durchbruch

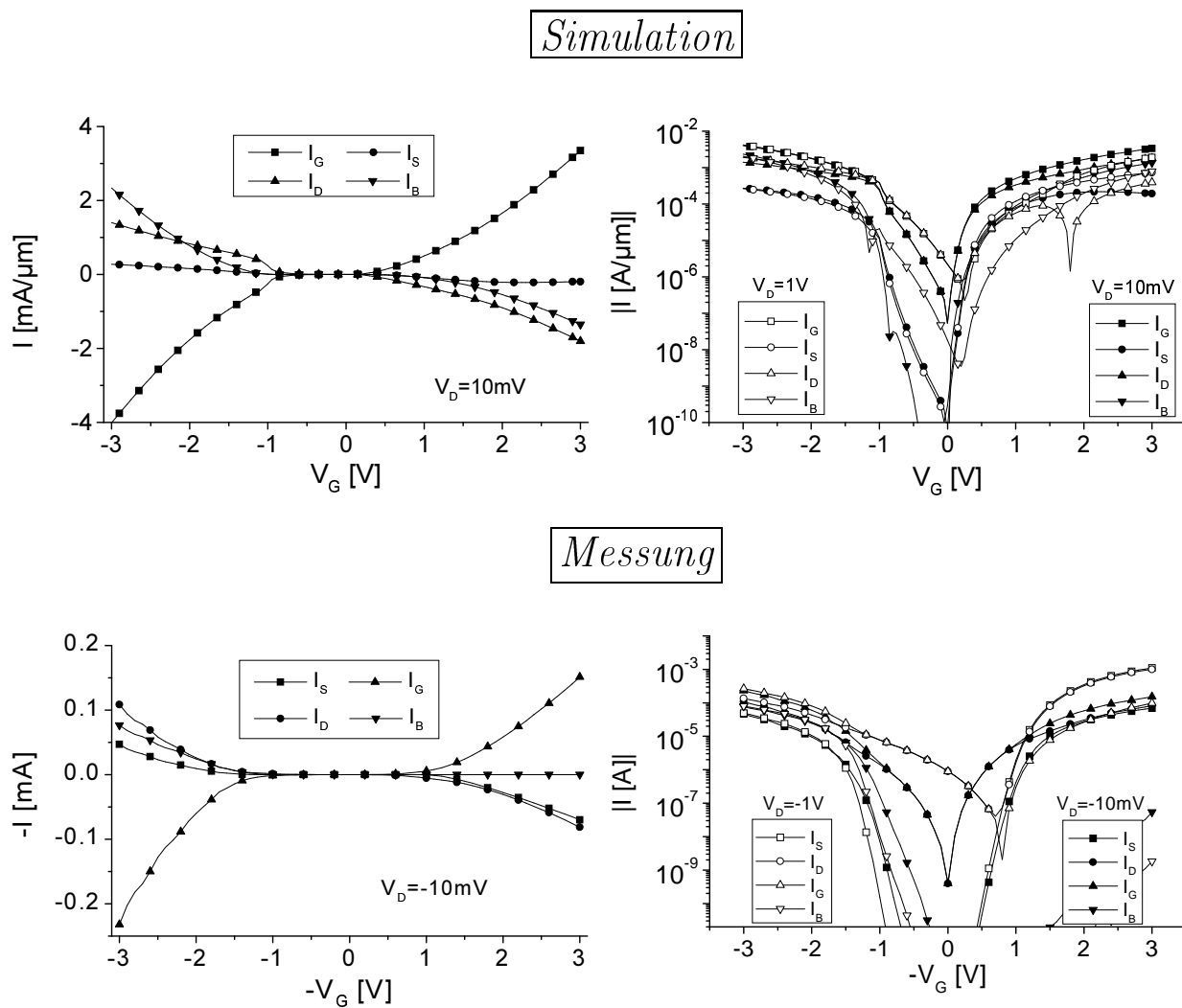


Bild 3.22: Simulation und Messung (pMOS-Transistor 03_2_T4p, $L=0.3\mu\text{m}$, $W=10\mu\text{m}$, $t_{ox} = 4.4\text{nm}$) in linearer (links) und halblogarithmischer (rechts) Darstellung.

Tabelle 3.3: Zusammenstellung der Merkmale von I_G bei hartem Durchbruch im Transistor

	(a) Durchbruch zum (Drain-) Überlappbereich	(b) Durchbruch zur Kanalmitte	(c) Durchbruch zum Kanalrand (Drain)
Allgemeine Merkmale von I_G	Ohmscher Verlauf	Diodenartiger Verlauf	Diodenartiger Verlauf
I_G fließt zum...	...Drain	...Bulk (Akk.) ...Drain & Source (Inv.)	...Bulk (Akk.) ...Drain (Inv.)
Abhängigkeit von V_D	I_G hat Minimum bei $V_G = V_D$	gering	I_G hat Minimum bei $ V_G < V_D $
Abhängigkeit von V_S	keine	gering	gering

zum Substrat erfolgt in Akkumulation Diodenbetrieb zum Bulk und in Inversion zu den Source- und Drain-Gebieten. Die Aufteilung des Stroms auf die Source-Drain-Gebiete und die Abhängigkeit vom dort anliegenden Potential hängt von der Lage des Durchbruchs im Kanal ab. Die erhaltenen Ergebnisse können zur Abschätzung der Fläche des Durchbruchs verwendet werden. Der simulierte Durchbruchstrom zum Gate liegt etwa eine Größenordnung über den gemessenen Werten. In der Simulation ergibt sich die Durchbruchfläche als $1\text{nm} \cdot 1\mu\text{m} = 10^{-15}\text{m}^2$. Hieraus kann der Schluß gezogen werden, daß die tatsächliche Fläche $\sim 10^{-16}\text{m}^2$ beträgt. Wie bereits erwähnt, müssen die Simulationsergebnisse aber wegen der verwendeten Modelle und der Dreidimensionalität des Problems skeptisch betrachtet werden.

Ersatzschaltbilder

Ziel der Durchbruchmodelle ist die Untersuchung der Auswirkung auf ganze Schaltungen. Hierdurch können Aufschlüsse über die Fehlertoleranz gewonnen und gewisse Voraussagen zum allgemeinen Verhalten bei signifikanten Gateströmen gemacht werden. Von Yeoh et al. wurden erste einfache Ersatzschaltbilder für den harten Durchbruch und Simulationen von Invertern gezeigt [64,65]. Kaczer et al. untersuchten ebenfalls die Auswirkung auf Schaltkreise [66] und stellten ein einheitliches Ersatzschaltbild für den harten Durchbruch bei dual-gate Transistoren vor, das für den Durchbruch zu allen Orten gilt [67,68]. Bild 3.23 zeigt das Prinzip. Der Durchbruch erzeugt eine Parallelschaltung aus zwei Bipolartransistoren und zwei Feldeffekttransistoren, der dritte Feldeffekttransistor repräsentiert den

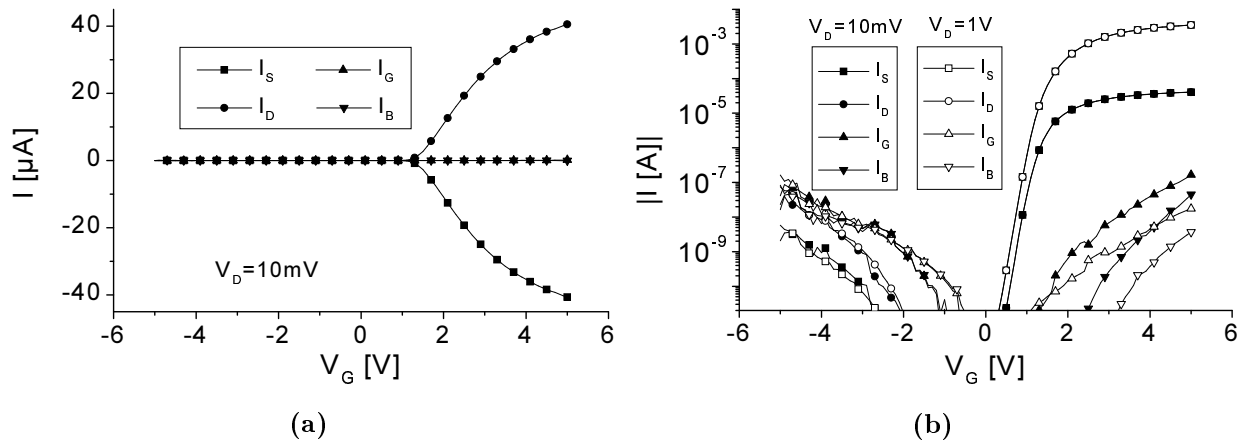


Bild 3.24: Messung (nMOS-Transistor 03_5_T2, $L=0.4\mu\text{m}$, $W=10\mu\text{m}$, $t_{ox} = 4.4\text{nm}$) in linearer (links) und halblogarithmischer (rechts) Darstellung.

Drainströme. Nur bei negativen Spannungen wird der Gatestrom dominierend. Dies darf jedoch nicht darüber hinwegtäuschen, daß sich die Kennlinien des Transistors sehr wohl verschieben, wie der Vergleich mit dem frischen Bauelement in Bild 3.25 zeigt. Die starke Veränderung der Kennlinien wurde sowohl an n- als auch pMOSFETs beobachtet. Diese Verschiebung kann durch die als Nebenprodukt des Durchbruchs entstandenen Grenzflächenzustände und durch die Veränderung des lokalen Potentials am Durchbruchsort entstehen. Bild 3.26 zeigt zur Illustration die mit Atlas simulierte Potentialverteilung im Transistor bei hartem Durchbruch. Im Durchbruchskanal fällt das Potential vom Wert im Polysilizium des Gates auf das Niveau im invertierten Kanal ab, das sich durch das um-

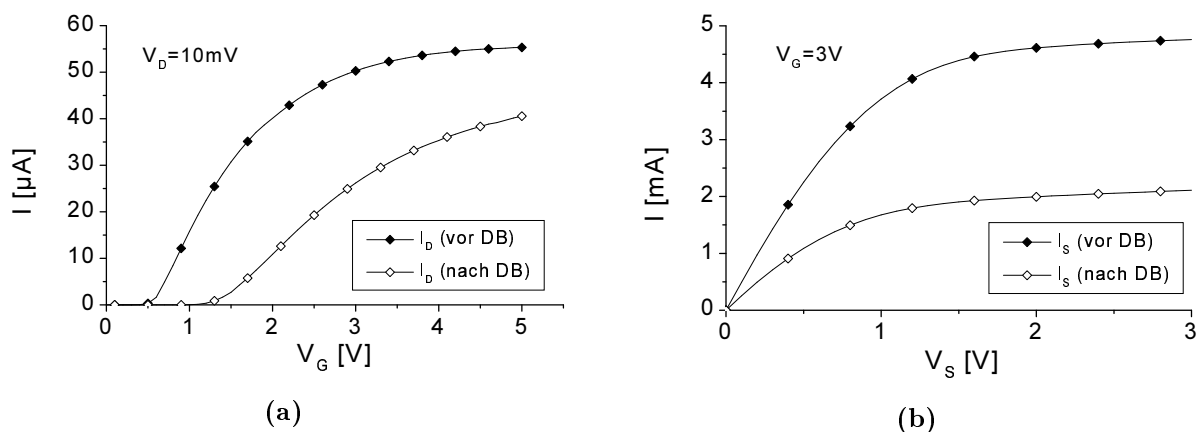


Bild 3.25: Vergleich der Transistorkennlinien vor und nach weichem Durchbruch (DB).

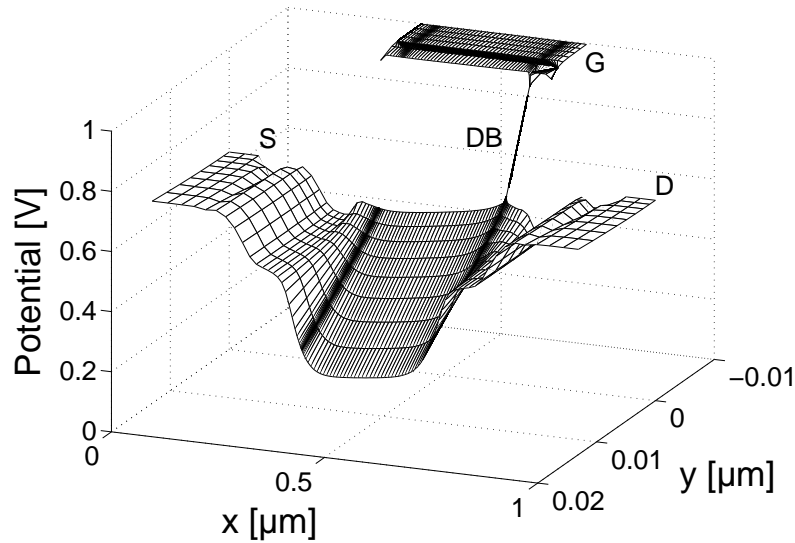


Bild 3.26: Von Atlas simulierte Potentialverteilung der Struktur aus Bild 3.19 mit hartem Durchbruch bei $V_G = 0.4\text{V}$ und $V_D = 10\text{mV}$ (Modelle mos, impact selv [34], der Durchbruchskanal hat die gleiche Dotierung wie das Substrat).

liegende ungeschädigte MOS-System ergibt. Dadurch erfolgt eine Potentialanhebung am Durchbruchsort. Die Simulationen beim harten Durchbruch ergaben eine recht starke Abhängigkeit dieser Potentialanhebung von den verwendeten Modellen (hydrodynamisch oder Drift-Diffusion) und der Dotierung des Durchbruchskanals. Bei einem weichen Durchbruch ist deswegen schwer abzuschätzen, inwieweit das Oberflächenpotential beeinflusst wird.

In der Literatur finden sich widersprüchliche Aussagen zur Veränderung der Transistorkennlinien nach weichem Durchbruch. Einige Autoren stellten keine Veränderungen fest [69, 70], andere Autoren zeigten eine Abhängigkeit von der Kanallänge [71] oder Kanalweite [48]. Cester et al. fanden eine Übereinstimmung der Kennlinienverschiebung mit der von Bauelementen nach Bestrahlung, bei denen eine lokale Schädigung des Oxids ohne Durchbruch entsteht [48]. Daraus kann geschlossen werden, daß beim Durchbruch tatsächlich eine sehr viel größere Region beschädigt wird als die Fläche des reinen Durchbruchspfads. Der Einfluß der Lage des Durchbruchs im Überlapp- oder Kanalbereich ist auch hier sehr groß. Die Änderung von Schwellspannung und Steilheit muß in den Durchbruchmodellen zusätzlich zum eigentlichen Gateleckstrom berücksichtigt werden.

Bild 3.27 zeigt die Meßdaten im Vergleich zu dem in Abschnitt 3.1.2 abgeleiteten Quantenpunktmodell. Gute Übereinstimmung in einem weiten Spannungsbereich ist gewährleistet. Es zeigt sich auch hier wie beim harten Durchbruch, daß der Stromfluß in Akkumulation zum Bulk, und in Inversion zu den Source-/Drain-Gebieten erfolgt. Bild 3.27

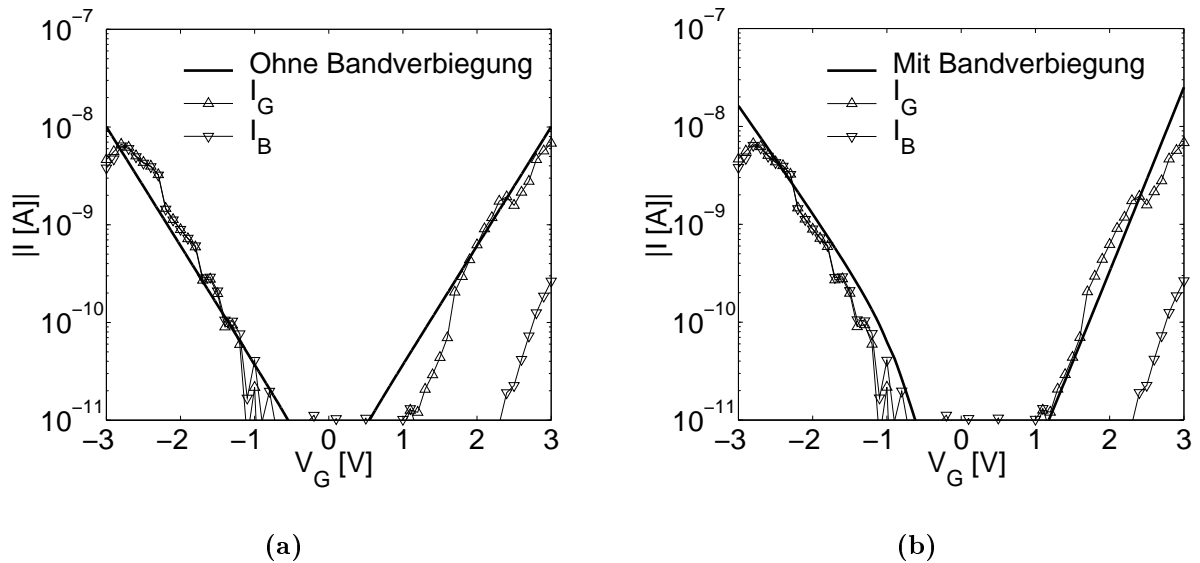


Bild 3.27: Gemessene Gate- und Bulkströme aus Bild 3.24 im Vergleich zum Quantenpunktkontaktmodell aus Abschnitt 3.1.2: (a) ohne Bandverbiegung ($\alpha = 5.6\text{eV}^{-1}$, $\beta = 0.5$, $\gamma = 2.4\text{meVK}^{-1}$, $\Phi(293\text{K}) = 2.8\text{eV}$), (b) mit Bandverbiegung ($\alpha = 6.8\text{eV}^{-1}$, $\beta = 0.7$, $\gamma = 2.4\text{meVK}^{-1}$, $\Phi(293\text{K}) = 2.8\text{eV}$).

zeigt auch das hohe Rauschen im Gatestrom (s. auch Abschnitt 4.2.1), was als Detektionshilfe für Durchbrüche bei dünnen Oxiden mit hohen direkten Tunnelströmen verwendet wird [72, 73].

3.3 Fazit

Es wurden die physikalischen Effekte des Stromtransports durch durchgebrochene Gateoxide von MOS-Strukturen untersucht und modelliert. Dabei wurde eine vollständige Charakterisierung der I-U-Kennlinien von MOS-Kapazitäten und Transistoren nach Gateoxiddurchbruch vorgestellt. Ausgehend von der physikalischen Modellierung des Gateleckstroms an MOS-Kapazitäten wurde auf den weichen und harten Durchbruchmechanismus eingegangen und analytische Modelle zu dessen Beschreibung entwickelt. In MOS-Transistoren kann aufgrund der Stromverteilung auf die vier Elektroden eine Lokalisierung und Klassifizierung der Durchbrüche durchgeführt werden. Hieraus ergeben sich unterschiedliche Ersatzschaltbilder für die Kennlinien nach Durchbruch. Weitere Aufmerksamkeit verdient die Verschiebung der Transistorkennlinien zusätzlich zum einsetzenden Gateleckstrom.

4

Random Telegraph Signals und Rauschen

Als Rauschen werden stochastische, also zufällig verteilte Signale bezeichnet, die dazu führen, daß das eigentliche Nutzsignal verfälscht wird. Im Gegensatz zu deterministischen Signalen ist dieser Zufallsprozeß nicht durch eine endliche Parameterzahl vollständig beschreibbar und auch nicht durch eine Momentaufnahme charakterisiert, so daß auf die Bildung von Mittelwertfunktionen und Wahrscheinlichkeitsdichten zurückgegriffen werden muß. In den meisten Ingenieursanwendungen geht es darum, das Signal-Rausch-Verhältnis zu maximieren, ohne den Ursprung des Rauschens weiter analysieren zu müssen. Auch dann ist es aber notwendig, Abschätzungen und Modelle über die erwarteten Rauschquellen zur Verfügung zu haben. Aus der Forschung haben sich innerhalb der zunächst zufälligen unsystematischen Signale Kategorien entwickeln lassen, die physikalisch begründet und durch bestimmte Merkmale unterscheidbar sind [74, 75]. So sind zum Beispiel das thermische Rauschen aufgrund der zufälligen thermischen Bewegungen der Ladungsträger, das Schrotrauschen sowie das Generations- und Rekombinationsrauschen bekannt. Das Studium der Fluktuationsphänomene begründet sich also aus der Herleitung der grundlegenden Mechanismen und der Verwendung des Rauschens als Untersuchungsgröße einerseits, führt dann andererseits zur Charakterisierung und später zur technologischen Optimierung von Strukturen und weiterhin zur Entwicklung von leistungsfähigen und empfindlichen Meßmethoden.

Schwankungen einer Meßgröße entstehen durch Auslenkungen eines oder mehrerer bestimmender Parameter um einen Mittelwert, in der Elektrotechnik sind die Meßgrößen Ströme und Spannungen. Der elektrische Strom I ist gegeben durch das Produkt von

Ladungsträgerdichte n , ihrer Ladung q , ihrer mittleren Geschwindigkeit v und dem Querschnitt A der Probe. Schwankungen im Strom können auf die veränderlichen Parameter, also Ladungsträgerdichte- oder Beweglichkeitsfluktuationen zurückgeführt und kategorisiert werden.

Bei niedrigen Frequenzen ist die in MOS-Transistoren vorherrschende Rauschform das Funkelrauschen, wegen des $1/f$ -förmigen Verlaufs seines Leistungsspektrums meist $1/f$ Rauschen genannt. Es kommt außer in MOS-Transistoren in einer Vielzahl unterschiedlichster Systeme und Materialien vor, weswegen noch immer nach einer einheitlichen Erklärung des Phänomens gesucht wird [76–78]. Unklar ist, ob es sich um einen Volumen- oder einen Grenz- bzw. Oberflächeneffekt handelt. Bei MOS-Transistoren, bei denen die Silizium-Siliziumdioxid-Grenzfläche eine entscheidende Rolle für die gesamte Bauteilcharakteristik spielt, ist diese Frage besonders schwer zu beantworten. Die konkurrierenden Theorien der letzten dreißig bis vierzig Jahre sind die folgenden:

1. Modellierung über Ladungsträgerzahlschwankungen: Oberflächenmodell

Bei dieser auf McWorther zurückgehenden Theorie [79] wird das $1/f$ Spektrum als Überlagerung der Lorentz-Spektren vieler Generations- und Rekombinationszentren, die jeweils einen Verlauf mit $1/f^2$ haben, interpretiert. Die unterschiedlichen Zeitkonstanten ergeben sich aus der energetischen Verteilung der Störstellen und der von der Entfernung zur Grenzfläche abhängigen Tunnelwahrscheinlichkeit in das Oxid.

2. Modellierung über Beweglichkeitsfluktuationen: Volumenmodell

Diese Theorie von Hooge [80] zieht zur Erklärung des Funkelrauschens die Änderung der Beweglichkeit μ durch Streuung im Volumen heran. Die empirische Beziehung für die spektrale Rauschleistung

$$S = \frac{\alpha_H \cdot I^2}{N f}, \quad (4.1)$$

mit N der Anzahl der freien Ladungsträger und α_H dem empirisch zu ermittelndem sogenannten Hooge-Parameter, hat sich als sehr universell herausgestellt. Die Beweglichkeitsänderungen werden durch Gitterschwingungen oder Störstellen erklärt.

3. Quanten- $1/f$ -Rauschen

Bei diesem von Handel entwickelten Modell [81] werden die Ladungsträger als elektromagnetische Welle und das $1/f$ Rauschen als Effekt der durch Streuung auftretenden Bremsstrahlung interpretiert. In Festkörpern kann mit Hilfe dieser Theorie der

Hooge-Parameter aus der Streuung am Gitter analytisch hergeleitet werden.

Bild 4.1 zeigt die gemessene spektrale Rauschleistungsdichte der Drain-Source-Spannung eines nMOS-Transistors bei zwei unterschiedlichen eingepprägten Drain-Strömen. Durch

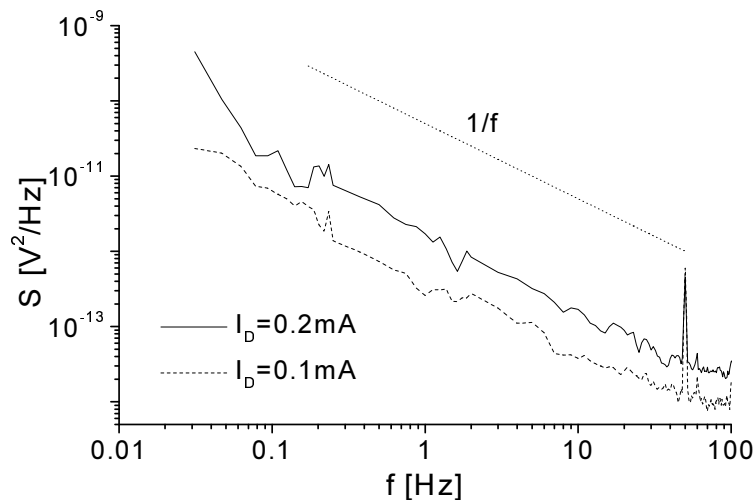


Bild 4.1: Spektrale Rauschleistungsdichte der Drain-Source-Spannung eines nMOS-Transistors bei zwei unterschiedlichen eingepprägten Drain-Strömen (Daten des Transistors: AMS 0.35 μm Prozeß, $L=0.3\mu\text{m}$, $W=0.6\mu\text{m}$, $V_G = 3\text{V}$).

die quadratische Abhängigkeit der Rauschleistung von der Stromstärke (s. Gleichung 4.1) nimmt bei einer Verdoppelung des Stroms das Rauschen um den Faktor vier zu.

Trotz intensiver Bemühungen ist das Phänomen Funkelrauschen noch nicht geklärt. So muß sich ein allgemeingültiges Modell nicht nur bei unterschiedlichen Materialien bewähren, sondern zum Beispiel auch Temperatur- und Geometriebhängigkeit richtig erfassen. Auch hier ist wohl wie so oft unser unzureichendes Verständnis der Phänomene der Natur ausschlaggebend für die noch unbefriedigenden Deutungsversuche.

4.1 Einführung in die Random Telegraph Signals

Random Telegraph Signals (RTS) sind eine faszinierende Unterform des Rauschens. Sie wurden ebenfalls in einer Vielzahl von Systemen, JFETs, MOS- und Bipolartransistoren beobachtet, trotzdem bleiben sie ein bis heute schwer zu greifendes Phänomen. Ein typischer Verlauf eines RT-Signals ist in Bild 4.2 gezeigt. Das Signal zeigt zwei definierte Zustände, zwischen denen es hin- und herwechselt, woraus sich die Bezeichnung Telegraphensignal ableitet. Die Zeiten im Zustand „oben“ und „unten“ sind exponential-(Poisson)-verteilt,

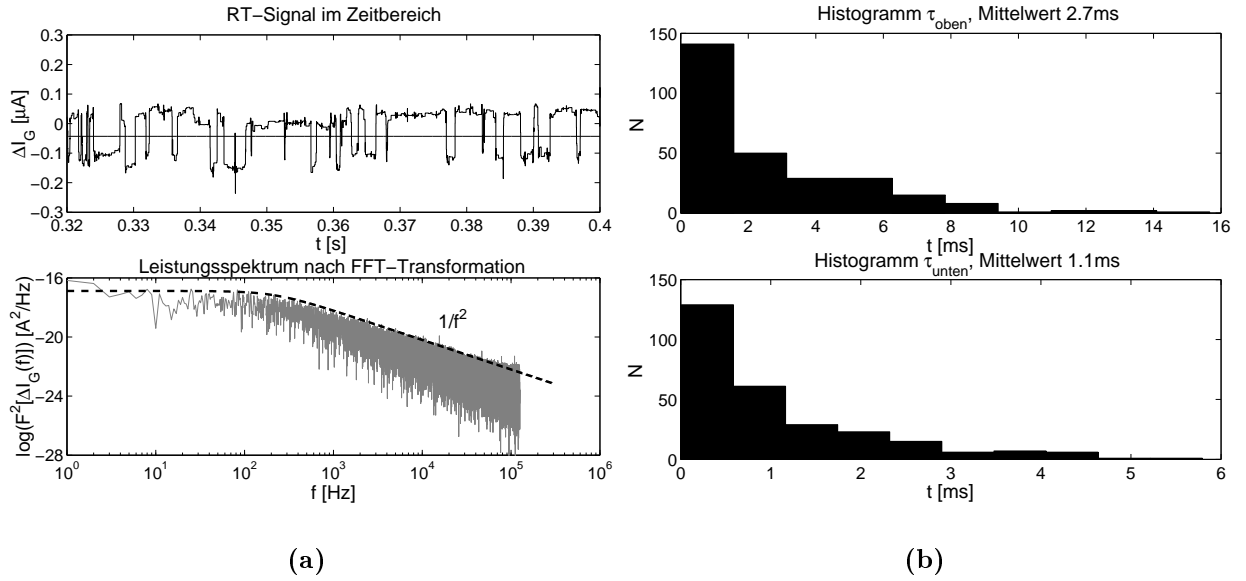


Bild 4.2: (a) Typisches RT-Signal, gemessen im Gatestrom eines durchgebrochenen nMOS-Transistors (08_8_1/2/rtva16) mit $t_{ox} = 2.7\text{nm}$, $L=0.5\mu\text{m}$, $W=10\mu\text{m}$ bei $V_{GS} = 0.8\text{V}$, $V_{DS} = 10\text{mV}$, $I_D = 8\mu\text{A}$ und zugehöriges Leistungsspektrum nach Fourier-Transformation im Vergleich zum mit Gl. 4.2 berechneten Spektrum ($\Delta I = 0.15\mu\text{A}$, $\overline{\tau_{oben}} = 2.6\text{ms}$, $\overline{\tau_{unten}} = 1\text{ms}$), (b) Histogramme der Zeitkonstanten ober- und unterhalb der eingezeichneten Schwelle.

woraus folgt, daß die Einfang- und Emissionswahrscheinlichkeit zeitunabhängig bzw. unabhängig von der Vorgeschichte ist. Die Varianz, welche für eine Exponentialverteilung gleich dem Mittelwert ist, kann zur Überprüfung der Validität von Meßdaten verwendet werden. Im Zeitbereich werden RT-Signale durch ihre Amplitude ΔI und mittleren Zeitkonstanten $\overline{\tau_{oben}}$ und $\overline{\tau_{unten}}$ in Abhängigkeit von Spannung und Temperatur spezifiziert. Im Frequenzbereich ergibt sich ein Lorentz-Spektrum, welches bis zur charakteristischen Grenzfrequenz konstant ist und dann mit $1/f^2$ abfällt. Der Zusammenhang zwischen Spektral- und Zeitbereich ist nach Machlup [82] gegeben durch

$$S(f) = \frac{4(\Delta I)^2}{(\overline{\tau_{oben}} + \overline{\tau_{unten}}) \cdot [(1/\overline{\tau_{oben}} + 1/\overline{\tau_{unten}}) + (2\pi f)^2]}. \quad (4.2)$$

Unter den herausragenden Forschern und Forschergruppen, die sich mit RT-Signalen beschäftigt haben, sind sicher zu nennen Kirton und Uren [83–85] und Schulz [86–89]. Viele weitere Forscher haben sich experimentell [90–93] und per Simulation [94, 95] mit dem Phänomen auseinandergesetzt und Modelle zu dessen Beschreibung entwickelt. Immer wieder ist zudem der Zusammenhang von RTS und $1/f$ Rauschen diskutiert worden. RTS in

MOSFETs wird als Indiz für die Richtigkeit der Grenzflächentheorie angesehen. RT-Signale treten demnach bei genügend kleinen geometrischen Abmessungen der Bauelemente zutage, in denen immer weniger Störstellen das Verhalten beeinflussen und sich schließlich nur noch eine einzige im relevanten Energiebereich befindet. Das Spektrum bei großflächigen Bauelementen wird dann, wie in Bild 4.3 gezeigt, als Überlagerung der Lorentz-Spektren mehrerer RT-Signale interpretiert [83,85]. Hierfür muß allerdings eine Verteilungsfunktion der Zeitkonstanten $g(\tau) \sim 1/\tau$ angesetzt werden [79,96]. Folgt man dieser Anschauung, so

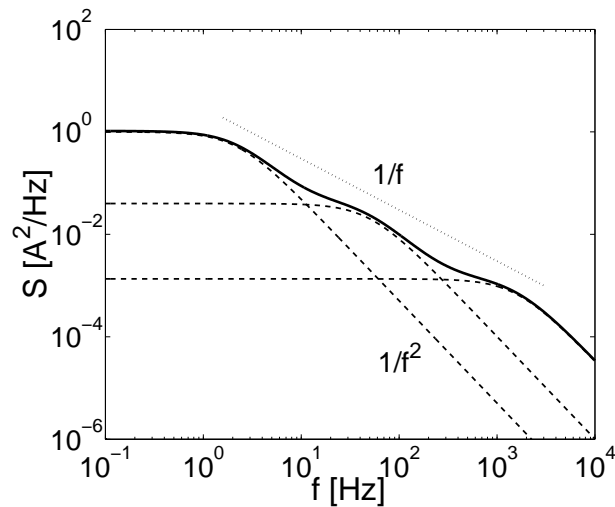


Bild 4.3: Schematische Darstellung der Überlagerung einzelner Lorentz-Spektren (gestrichelt) zu einem $1/f$ -förmigen Spektrum (durchgezogene Linie).

ist mit der Erforschung des Rauschens die Hoffnung verbunden, Aussagen über die Qualität des Oxids und der Grenzfläche treffen zu können und somit ein weiteres Werkzeug zur Lebensdauervorhersage von Bauelementen zur Verfügung zu haben [97]. Die Vorstellung von RTS als Ursprung des $1/f$ Rauschens wird aber auch kritisiert [78].

Unabhängig von der Interpretation des Funkelrauschens scheinen RT-Signale tatsächlich die einmalige Möglichkeit zu bieten, einzelne Störstellen genauestens bezüglich energetischer und geometrischer Lage und ihrem Einfang- und Aussendeverhalten zu charakterisieren. Zunächst soll hier eine Betrachtung der Historie und des Forschungsstandes erfolgen, bevor in Abschnitt 4.2 die im Rahmen dieser Arbeit durchgeführten RTS-Messungen in Gate- und Drainströmen von MOSFETs mit durchbrochenen Oxiden und in Abschnitt 4.3 ein neues Modell zur Erklärung der experimentellen Ergebnisse vorgestellt werden.

4.1.1 Grundlagen von RTS im Drainstrom von MOSFETs

Am intensivsten untersucht und modelliert ist wahrscheinlich das RTS-Rauschen im Drainstrom des eingeschalteten MOS-Transistors. Hier wird eine geringe Drain-Source-Spannung V_{DS} angelegt, so daß man näherungsweise von thermischem Gleichgewicht ausgehen kann. Die beobachteten Fluktuationen im Drainstrom werden, wie Bild 4.4 schematisch zeigt, auf eine lokale Modulation des Leitwerts des invertierten Kanals zurückgeführt. Dies führt zu diskreten Stromsprüngen, die dem Einfang und der Emission der Ladungsträger zugeordnet werden können. Im Standardwerk von Kirton und Uren [84] sind die meisten

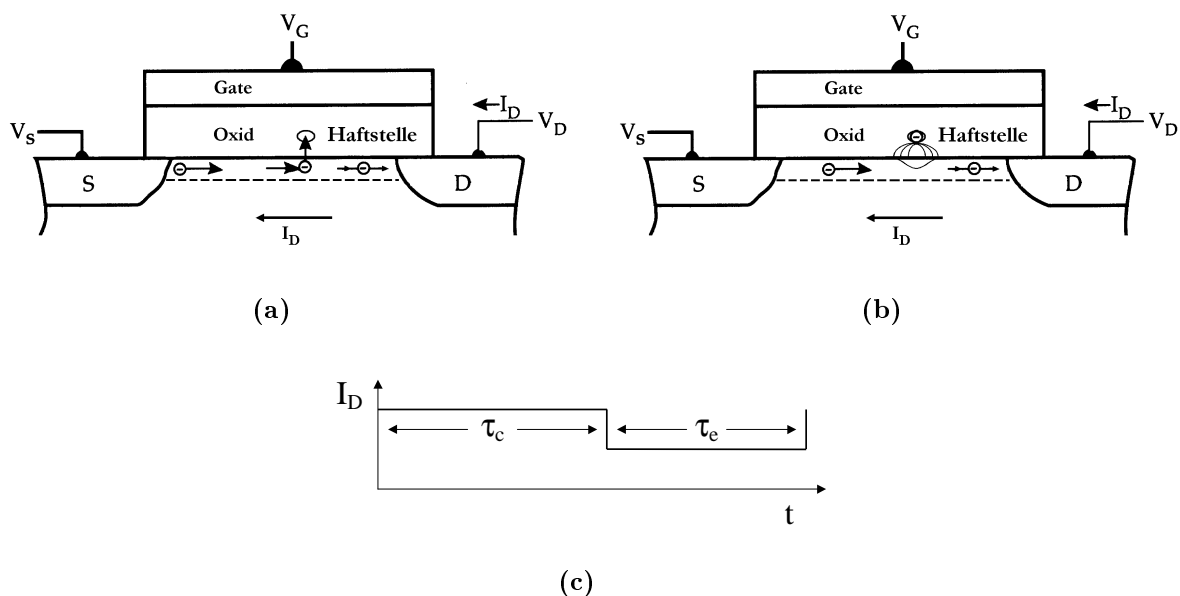


Bild 4.4: Prinzip der Entstehung von Random Telegraph Signals. (a) Ladungsträgereinfang, (b) lokale Störung des Kanals und (c) makroskopische Auswirkung auf den Drainstrom.

Grundlagen für die Modellierung der Signale gelegt worden. Obwohl es manchmal schwer ist nachzuvollziehen, daß die Wirkung eines einzelnen Ladungsträgers zur Deutung der Meßdaten ausreicht, besteht eine gut ausgearbeitete Theorie, die ausschließlich auf Einzelladungseffekten beruht.

Modellierung der Amplitude

Eine weit verbreitete Herleitung für die Fluktuationsamplitude erwächst als Sonderfall aus dem von Hung et al. vorgestellten sogenannten vereinigten Modell für das Funkelrauschen [98], bei dem davon ausgegangen wird, daß sich durch Einfang an Störstellen

eine korrelierte Fluktuation in der Dichte n und Beweglichkeit μ der freien Ladungsträger ergibt. Wird der Drainstrom I_D vereinfacht geschrieben als

$$I_D = \mu q n A E_x, \quad (4.3)$$

wobei E_x das Feld in horizontaler Richtung und A die Querschnittsfläche des Kanals ist, dann folgt für die Änderung des Drainstroms δI_D in einem Teilstück δx des Kanals:

$$\delta I_D = \mu q \cdot \delta n \cdot A E_x + \delta \mu \cdot q n A E_x \quad (4.4)$$

Hieraus ergibt sich die relative Änderung des Drainstroms mit Umformung nach der Variation der Anzahl besetzter Störstellen δN_T zu

$$\frac{\delta I_D}{I_D} = \frac{\delta n}{n} + \frac{\delta \mu}{\mu} = \frac{\delta N}{N} + \frac{\delta \mu}{\mu} = \left(\frac{1}{N} \frac{\delta N}{\delta N_T} + \frac{1}{\mu} \frac{\delta \mu}{\delta N_T} \right) \delta N_T. \quad (4.5)$$

Der Faktor $\delta N/\delta N_T$ gibt die Kopplung der Besetzung der Störstellen mit der Anzahl der Ladungsträger im Kanal an. Im einfachsten Fall führt die Emission eines Ladungsträgers aus einer Störstelle auch zu einem freien Ladungsträger und dieser Faktor wird zu -1 gesetzt. Zur Modellierung der Beweglichkeit werden im wesentlichen zwei Streumechanismen betrachtet, Coulomb-Streuung an Ladungen und Gitter- bzw. Grenzflächenstreuung. Mit μ_c der Coulomb-Beweglichkeit und μ_n der Grenzflächenbeweglichkeit ergibt sich die Gesamtbeweglichkeit

$$\frac{1}{\mu} = \frac{1}{\mu_n} + \frac{1}{\mu_c} = \frac{1}{\mu_n} \pm \alpha_s N_T. \quad (4.6)$$

Die Coulomb-Beweglichkeit wird über den Streufaktor α_s beschrieben, der wiederum abhängig ist von der Ladungsträgerdichte und der Entfernung der Störstelle von der Grenzfläche. Im Prinzip wird die Beweglichkeit also durch einen störstellenunabhängigen und einen abhängigen Term modelliert. Das Vorzeichen vor dem zweiten Term folgt daraus, ob die Störstelle im unbesetzten Zustand neutral oder geladen ist. Durch Ableitung von Gleichung 4.6 nach δN_T und Umrechnung einer Störstellenumladung flächenabhängig auf die globale Beweglichkeit vereinfacht sich Gleichung 4.5 zu

$$\frac{\delta I_D}{I_D} = - \left(\frac{1}{n} \pm \alpha_s \mu \right) \frac{\delta N_T}{WL}. \quad (4.7)$$

Das Rauschspektrum ergibt sich aus der Integration über die gesamte Fläche und der darin enthaltenen Störstellen.

Wird davon ausgegangen, daß RTS durch eine einzige Störstelle hervorgerufen wird, kann für die Betrachtung der Amplitude ΔI_D des Signals Gleichung 4.7 mit $\delta N_T = 1$ verwendet werden, und es folgt

$$\frac{\Delta I_D}{I_D} = \frac{\Delta N}{N} \pm \frac{\Delta \mu}{\mu} = - \left(\frac{1}{n} \pm \alpha_s \mu \right) \frac{1}{WL}, \quad (4.8)$$

wobei der Streufaktor über $\alpha_s = \alpha_0 + \alpha_1 \ln N$ mit den Fitparametern α_0 und α_1 modelliert werden kann [98].

Von Kirton und Uren wurde bei Vernachlässigung der Streuung und unter Berücksichtigung der Ladungserhaltung im MOS-System die folgende einfache Beziehung aufgestellt [84]

$$\frac{\Delta I_D}{I_D} = \frac{aq}{V_t \cdot (C_{ox} + C_{it} + C_D) - aQ_n}, \quad (4.9)$$

wobei a eine Konstante zwischen 0.5 und 1, V_t die Temperaturspannung, C_{ox} , C_D und C_{it} die Kapazität des Oxids, der Verarmungszone und der Grenzflächenzustände und Q_n die Ladung der Inversionsschicht sind.

Bei einem ebenfalls sehr einfachen Modell von Roux-dit-Buisson [99, 100] wird der Einfang einer Einzelladung q über die Steilheit g_m des Transistors als äquivalent zu einer Änderung der Flachbandspannung und damit indirekt der Gatespannung angesehen:

$$\Delta I_D = g_m \frac{q}{WL C_{ox}} \quad (4.10)$$

Die gemessenen Amplituden der Fluktuationen haben eine große Bandbreite, die mit den angeführten Modellen oft schwer zu erklären ist. Schulz führte die Vorstellung eines inhomogenen Kanals ein, in dem der Stromtransport hauptsächlich durch leitende perkolative Pfade stattfindet [88]. Daraus ergibt sich, daß strategisch günstig an einem solchen Pfad gelegene Oxidstörstellen den Kanal wesentlich stärker beeinflussen können als weiter abgelegene und somit die große Streubreite der beobachteten Amplituden zu erklären ist.

Modellierung der Zeitkonstanten

Einfang und Emission an Störstellen in einem Halbleiter wird üblicherweise nach der Shockley-Read-Hall-(SRH)-Statistik und dem Prinzip des detaillierten Gleichgewichts analysiert [9, 101, 102]. Hierbei muß unterschieden werden zwischen Generations-Rekombinationszentren, die mit ihrem Energieniveau nahe der Mitte der Bandlücke als Trittstein für das Band-zu-Band Tunneln dienen, und Haftstellen, bei denen der Ladungsträger in das

ursprüngliche Band zurückemittiert wird. Bei Elektronenhaftstellen liegt das Energieniveau näher am Leitungsband und bei Löcherhaftstellen näher am Valenzband. Obwohl die SRH-Theorie ursprünglich für Störstellen im Halbleitervolumen entwickelt wurde, findet sie auch für Einfang und Emission an Grenzflächenzuständen und sich nahe der Grenzfläche im Oxid befindende Störstellen Anwendung. Es wird hierbei von der selben Einfang- und Emissionskinetik ausgegangen, obwohl bei Störstellen im Oxid noch ein Tunnelvorgang stattfinden muß und die Ladungsträger sich nicht frei in einem Volumen bewegen.

Es werde nun der allgemeine Fall von Elektronen-(Akzeptor)-Störstellen der Dichte n_T betrachtet. Im thermischen Gleichgewicht ergibt sich der Besetzungsfaktor f_T einer Störstelle aus ihrer energetischen Lage E_T zum Fermi-Niveau E_F aus der Fermi-Dirac-Funktion:

$$f_T = \frac{1}{1 + \exp[-(E_F - E_T)/kT]} \quad (4.11)$$

Es folgt für die Dichte der unbesetzten Störstellen n_T^0 und besetzten Störstellen n_T^- :

$$n_T^0 = (1 - f_T) n_T \quad (4.12)$$

$$n_T^- = f_T n_T \quad (4.13)$$

Nun können Übergangsraten für den Elektronenübergang vom Leitungsband auf Störstellenniveau (Einfang) und den umgekehrten Prozeß (Emission) definiert werden. Die Einfangrate R_C folgt aus der Wahrscheinlichkeit, daß ein Elektron durch die zufällige thermische Bewegung im Volumen in den Einflußbereich einer freien Störstelle kommt. Sie ist also abhängig vom Produkt aus zur Verfügung stehenden Ladungsträgern n , unbesetzten Störstellen n_T^0 , und dem Einfangfaktor c_n , der sich aus mittlerer thermischer Geschwindigkeit v_{th} der Ladungsträger und dem Einfangquerschnitt σ der Störstellen zusammensetzt. Analog ergibt sich die Emissionsrate R_E als Produkt von besetzten Störstellen n_T^- , freien Zuständen im Leitungsband $(N_C - n)$ und dem Emissionsfaktor e_n :

$$R_C = c_n \cdot n \cdot n_T^0 = v_{th} \cdot \sigma \cdot n \cdot n_T^0 \quad (4.14)$$

$$R_E = e_n \cdot (N_C - n) \cdot n_T^- = e_n' \cdot n_T^- \quad (4.15)$$

Im Leitungsband ist in einem nicht entarteten Halbleiter eine große Anzahl besetzbarer Zustände vorhanden, da die effektive Zustandsdichte N_C sehr viel größer ist als n . Deswegen wird üblicherweise ein neuer Faktor e_n' definiert, in dem die effektive Zustandsdichte aufgeht. Im thermodynamischen Gleichgewicht müssen sich die gegenläufigen Prozesse

aufheben und R_E gleich R_C gelten, so daß im Mittel der Besetzungsfaktor f_T eingehalten wird. Hieraus folgt für den Faktor

$$e'_n = c_n \cdot n \cdot \frac{n_T^0}{n_T} = v_{th} \cdot \sigma \cdot n \cdot \exp\left(-\frac{E_F - E_T}{kT}\right) = v_{th} \cdot \sigma \cdot N_C \cdot \exp\left(-\frac{E_C - E_T}{kT}\right), \quad (4.16)$$

wenn zur Berechnung von n die Boltzmann-Näherung verwendet wird. Ähnliche Gleichungen lassen sich auch für den Einfang von Löchern aus dem Valenzband ableiten. Shockley und Read verwendeten die so erhaltenen Beziehungen für Elektronen und Löcher zur Berechnung der Rekombinationsrate bei Auslenkung eines Halbleiters aus seinem Gleichgewichtszustand [103]. Wird durch schwache Injektion von Minoritätsladungsträgern das Gleichgewicht gestört, so erfolgt Rekombination mit den Majoritätsladungsträgern über die Störstellen, wobei deren Besetzungsfaktor unverändert bleibt. Die Minoritätsträgerlebensdauer als charakteristische mittlere Dauer bis zur Rekombination der überschüssigen Ladungsträger zum Beispiel von Elektronen in einem p-Halbleiter ist dabei als $\tau_n = \frac{1}{c_n n_T}$ definiert.

Analog hierzu kann für die Betrachtung von RT-Signalen der Zustand der Störstellenbesetzung als Variable angenommen werden. Wird dieser aus dem Gleichgewicht ausgelenkt, kehrt das System mit der Zeitkonstanten $\tau_c = \frac{1}{c_n n}$ bzw. $\tau_e = \frac{1}{e'_n}$ zum Besetzungsfaktor $f_T = \frac{\tau_e}{\tau_c + \tau_e}$ zurück. Bei Betrachtung einer einzigen Störstelle ist τ_c die Einfangzeitkonstante, also die mittlere Zeit, während der die Störstelle unbesetzt ist. Ist das Fermi-Niveau im Inversionskanal bekannt, so kann aus dem Verhältnis der mittleren Zeitkonstanten das Energieniveau der Störstelle berechnet werden.

Die Spannungsabhängigkeit der Zeiten „oben“ und „unten“ wird zur Identifizierung als Einfang- bzw. Emissionszeit verwendet. Die Zeit τ_c bis zum Einfangen fällt beim nMOSFET mit zunehmender Gatespannung und somit zunehmendem n ab, während die Verweildauer τ_e , wie aus Gleichung 4.16 hervorgeht, praktisch spannungsunabhängig ist. Die Zunahme der Ladungsträgerdichte reicht aber zur Erklärung der starken Spannungsabhängigkeit von τ_c nicht aus. Vor allem bei starker Inversion wächst die Anzahl der Ladungsträger im Kanal fast nicht mehr an, eine starke Abnahme der Einfangzeitkonstante wird aber weiterhin beobachtet. Kirton und Uren wiesen auf diesen Widerspruch hin, den sie auf eine Zunahme des Einfangquerschnitts zurückführten, und schlugen einige Hypothesen zur Erklärung vor [84, Abschnitt 5.1.4]. Schulz führte das Konzept der zur Umladung der Störstelle notwendigen Coulomb-Energie ein [86]. Die anteiligen Spiegelladungen des eingefangenen Ladungsträgers auf der Gateelektrode, im Kanal und im Substrat müssen

entgegen dem elektrischen Feld aufgebaut werden, wofür Energie notwendig ist. Auch die Coulomb-Energie ändert sich aber in starker Inversion aufgrund der fast konstanten Ladungsträgerdichte kaum noch.

Ein neu entwickeltes, aus der Multiphonontheorie abgeleitetes Modell, welches die beobachteten Daten wiedergibt, wird in Abschnitt 4.3 vorgestellt.

4.1.2 Überblick zu RTS in MOS-Tunneldioden

Beim Großteil der in den 80er Jahren durchgeführten Studien wurden RT-Signale im Gatestrom von in Akkumulation befindlichen MOS-Dioden mit Tunneloxiden beobachtet [104–107]. Zur Interpretation der teilweise sehr hohen und sich unterschiedlich verhaltenden Fluktuationsamplituden wurde vorgeschlagen, daß interagierende Störstellen gleichzeitig ihre Konfiguration und Ladungszustand ändern, bzw. mehrere Ladungsträger gleichzeitig eingefangen werden [108]. Obwohl die Idee der Clusterbildung nicht von vornherein ausgeschlossen werden kann, entzieht sie sich ein wenig der analytischen Beschreibung und kann daher auch etwas willkürlich angewandt werden.

Bei den Zeitkonstanten wurde zunächst das Temperaturverhalten untersucht und eine thermische Aktivierung oberhalb von $T \sim 15\text{K}$ ähnlich wie beim Drainstrom von MOSFETs beobachtet [104, 105]. Die Spannungsabhängigkeit wies jedoch im Gegensatz zu den Daten im Drainstrom in den überwiegenden Fällen eine Abnahme beider Zeitkonstanten mit zunehmender Spannung auf. Teilweise wurde eine Überlagerung aus zwei unterschiedlichen Zeitkonstanten in Betracht gezogen [104]. Ebenfalls wurde versucht, die Verschiebung des Energieniveaus der Störstelle relativ zum Fermi-Niveau der Elektrode durch die Spannung V_G ähnlich wie im Drainstrom für die beobachtete Abhängigkeit der Zeitkonstanten verantwortlich zu machen. Diese Verschiebung ist abhängig vom Abstand z der Störstelle zur Elektrode. Für eine in Akkumulation befindliche MOS-Diode und einen Tunnelvorgang von und zur selben Elektrode folgt für die Zeitkonstanten [107]

$$\tau_{e/c} = \tau_0 \exp\left(\frac{E_A \pm qV_G \cdot z/t_{ox}}{kT}\right), \quad (4.17)$$

wobei das Vorzeichen in der Klammer für die Verweildauer τ_e bzw. Einfangzeitkonstante τ_c gilt und E_A die thermische Aktivierungsenergie bezeichnet. Gleichung 4.17 sagt also eine gleichartige Ab- und Zunahme der Zeitkonstanten mit der Spannung voraus und erklärt somit nicht die beobachtete gleichzeitige Abnahme. Innerhalb der RTS-Studien in MOS-Dioden am weitesten entwickelt ist ein numerischer Modellansatz, der auf der Vor-

stellung des störstellengestützten Tunnelns als zweistufigem Tunnelprozeß durch das Oxid beruht [109]. Die Störstelle kann hierbei einen Grund- und einen angeregten Zustand mit zwei unterschiedlichen Zeitkonstanten einnehmen. Der Stromtransport erfolgt über den angeregten Zustand, während bei Besetzung des Grundzustands der Stromfluß blockiert wird.

Im Vergleich zu der Theorie im Drainstrom von MOSFETs ist hier noch keine abschließende Beschreibung vorhanden.

4.2 Beobachtung von RTS in MOSFETs mit durchgebrochenen Gateoxiden

Im Rahmen dieser Arbeit wurden RTS-Analysen an MOS-Transistoren nach Gateoxid-durchbruch durchgeführt. Gemessen wurde mit Hilfe eines nach Hung et al. [110] konzipierten Meßverstärkers für AC-Signale der Verstärkung 10^6 plus einem LeCroy Oszilloskop LC584AM oder mit einem Keithley SCS-4200 Semiconductor Characterization System. Bei der ersten Variante wurden schnelle Signale bei einer Samplingrate von 250kHz aufgenommen, der DC-Anteil wurde mit dem Picoamperemeter HP4140B und einem Multimeter HP3478A oder einem Keithley 616 Electrometer aufgenommen, Bild 4.5 zeigt das Prinzip. Die Auflösung der DC-Ströme war durch die Meßplatine auf Werte zwischen 10nA

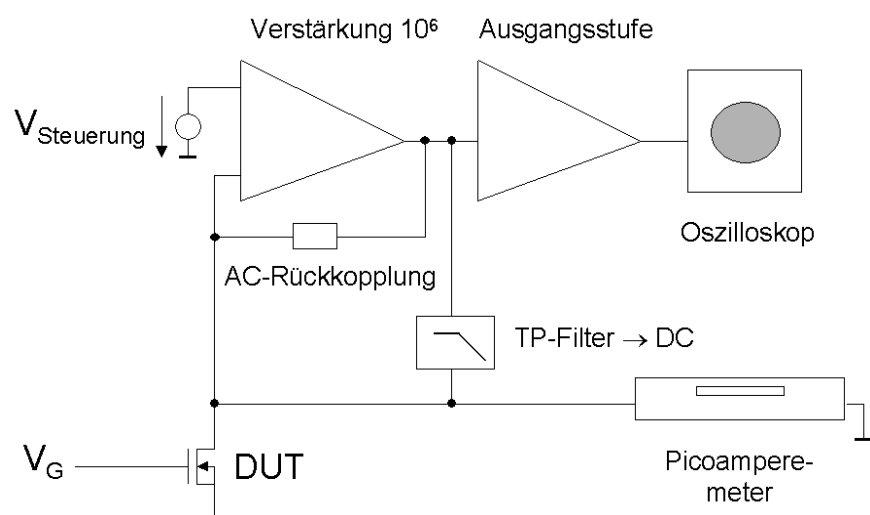


Bild 4.5: Prinzip des verwendeten Meßverstärkers [110]. Beispielhaft ist hier die Messung des Drainstroms des untersuchten Transistors „DUT“ gezeigt. Gleichzeitig kann für Korrelationsmessungen ein identischer Verstärker an das Gate angeschlossen werden.

und $100\mu\text{A}$ beschränkt, RT-Signale konnten bis 1nA aufgelöst werden. Die zweite Variante mit dem Keithley SCS-4200 bot sich bei langsameren Zeitkonstanten über $0,1\text{s}$ (100Hz) an. Die aufgenommenen Signale wurden unter Matlab weiterverarbeitet und ausgewertet. Höherfrequentes Rauschen wurde gegebenenfalls mit Hilfe der Wavelet-Denoising-Toolbox durch die automatische Rauschentfernung bei 16 Leveln und dem rechteckförmigen Haar-Wavelet biorthogonaler Basis „bior1.5“ entfernt [111, 112].

Es ist bekannt, daß nach Gateoxiddurchbruch eine deutliche Zunahme des Rauschens zu verzeichnen ist [73]. Die hohe Defektdichte im Durchbruchskanal führt zu einer hohen Wahrscheinlichkeit, daß auch RT-Signale zu beobachten sind. Da bei dünnen Oxiden mit hohen Tunnelströmen ein Durchbruch in der I-U-Kennlinie nur schwer zu detektieren ist, bietet sich die Rauschmessung im Gatestrom als alternative Meßmethode an. Die Zunahme des $1/f$ - oder $1/f^2$ -Rauschens kann somit als zusätzliches Kriterium bei der Bewertung der Transistorqualität verwendet werden. Bei Bauelementen mit geringen Abmessungen könnte dabei das $1/f^2$ -Rauschen in den Vordergrund gelangen.

4.2.1 Lokalisierung des Durchbruchs bei MOS-Transistoren

Durch die Vierpolnatur des MOS-Transistors ergibt sich die Frage, inwieweit sich bei dünnsten Oxiden in Gate- und Drainstrom Rauschanteile überlagern. Prinzipiell kann man von korreliertem Rauschen der beiden Ströme ausgehen, wenn das Rauschen des Gatestroms direkt in den Drain- oder Source-Strom einfließt oder über die Steilheit des Transistors Änderungen im Gatepotential eingekoppelt werden. Unkorreliertes Rauschen kann entstehen, wenn weitere Störstellen im Kanal vorhanden sind, die den Drain-Source-Strom unabhängig von einem eventuell vorhandenem Gatestrom beeinflussen. Neben DC-Messungen (s. Abschnitt 3.2) kann die Korrelation der RT-Signale an den Elektroden zur Bestimmung des Durchbruchortes verwendet werden [113]. Bild 4.6 mit den gleichzeitig aufgenommenen Signalen in Gate, Source und Drain ist ein Beispiel hierfür. Die DC-Analyse ergab hier einen harten Durchbruch vom Gate zum Drain-Überlappbereich, was durch die RT-Signale bestätigt wird: Die Korrelation der Fluktuationen in Gate- und Drainstrom ist -1 . Ein zusätzlicher Strompfad zwischen Gate und Drain ist entstanden, der Fluktuationen aufweist. In diesem Fall fließt der Gatestrom sofort zum Drain ab und findet sich nicht im Kanal wieder (s. auch Abschnitt 3.2), somit sind auch keine Fluktuationen im Source-Strom zu erkennen. Hier ist zwar ein unkorreliertes Rauschen zu sehen, auswertbare RT-Signale an der nicht durchgebrochenen Elektrode wurden im Rahmen dieser Arbeit allerdings nicht beobachtet.

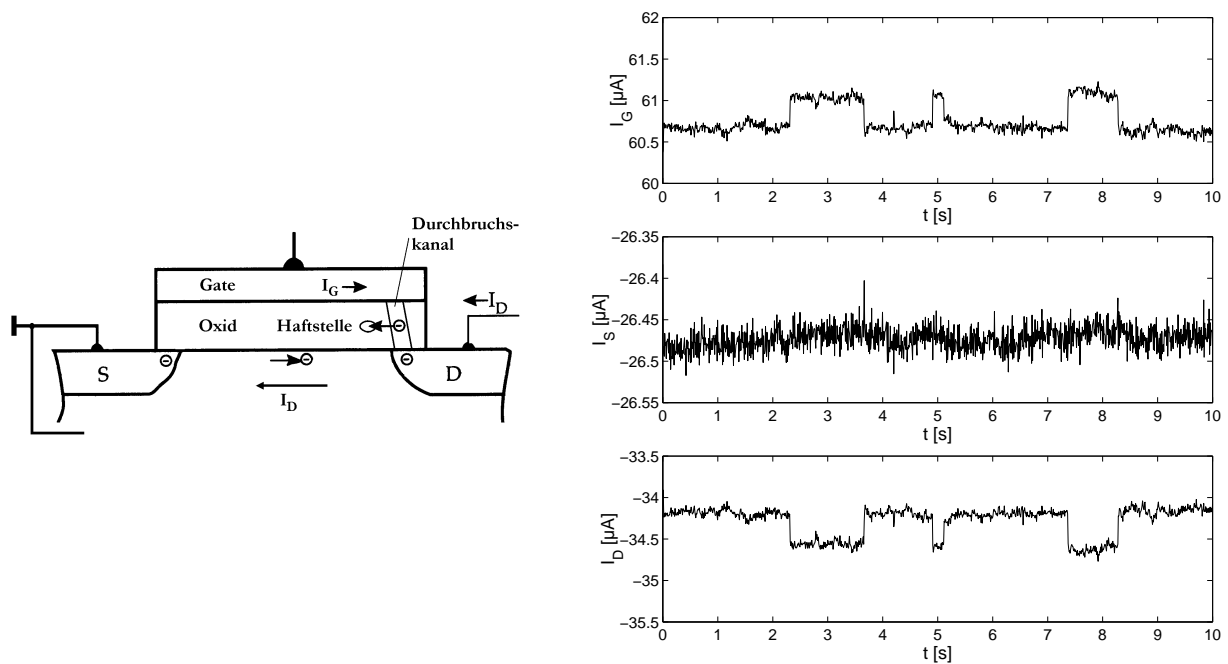


Bild 4.6: Gleichzeitig mit dem Keithley SCS-4200 (Samplingzeit 9ms) gemessene Ströme in Gate, Source und Drain eines durchgebrochenen nMOSFET (08_5_1) mit $L=0.5\mu\text{m}$, $W=10\mu\text{m}$, $t_{ox} = 2.7\text{nm}$ bei 20°C , $V_G = 0.6\text{V}$, $V_D = 50\text{mV}$ und dazugehöriges Prinzipbild.

4.2.2 Messung der Amplituden und Zeitkonstanten

Das Verhalten der Amplituden und Zeitkonstanten der RT-Signale im Gatestrom in Abhängigkeit der angelegten Gatespannung wurde an vier ausgewählten nMOS-Transistoren betrachtet. Im aufgenommenen Zeitintervall von einer Sekunde wiesen diese Transistoren einige hundert Sprünge zwischen den beiden Zuständen auf, so daß eine gute statistische Basis für die Auswertung der Signale gewährleistet war. In Bild 4.7(a) sind statischer mittlerer Gatestrom und Fluktuationsamplitude als Funktion der Gatespannung einander gegenübergestellt. Die DC-Analyse läßt auf einen harten Durchbruch zum Kanalrand in Drainnähe schließen (Fall (c) in Abschnitt 3.2.1). Der Verlauf mit der Gatespannung ist für Strom und Fluktuationsamplitude vergleichbar, die Höhe der Fluktuationen ist ebenfalls für alle vier Transistoren in ähnlicher Größenordnung. Wird, wie in Bild 4.7(b) gezeigt, der Quotient aus Fluktuationsamplitude und mittlerem Gatestromwert gebildet, so ergeben sich zwei unterschiedliche Abhängigkeiten: Ein starker Abfall der relativen Amplitude bei den mit „D“ bezeichneten, und ein eher konstanter bzw. leicht ansteigender Verlauf bei den mit „C“ bezeichneten Transistoren. Diese unterschiedlichen Verläufe finden sich auch bei Betrachtung der Zeitkonstanten in Bild 4.8 wieder. Bild 4.8(a) zeigt bei den Zeitkon-

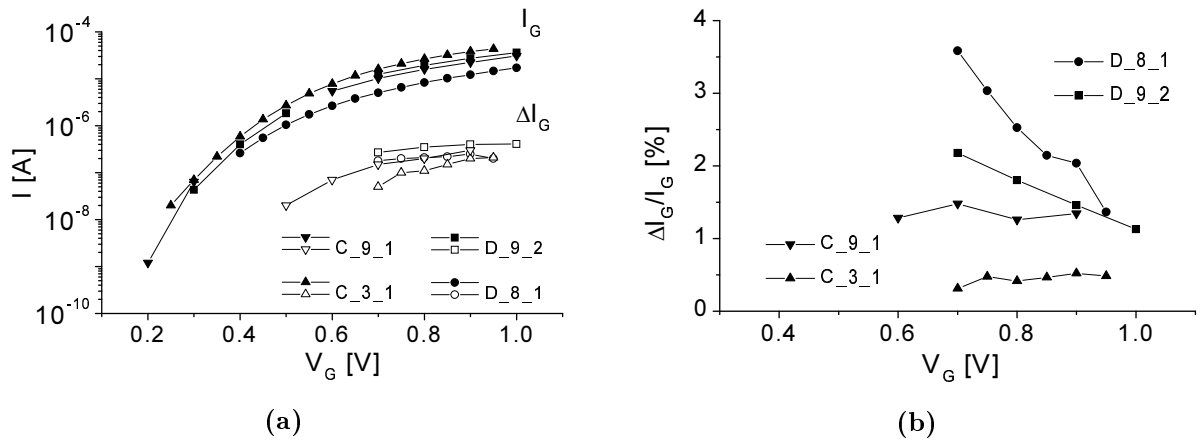


Bild 4.7: (a) Mittlerer DC-Gatestrom I_G und Fluktuationsamplitude ΔI_G in Abhängigkeit der Gatespannung von vier nMOS-Transistoren nach Gateoxiddurchbruch nahe am Drain, (b) Relative Fluktuationsamplitude. Daten der Transistoren: Infineon 0.25 μm Prozeß, $W=10\mu\text{m}$, $L=0.4\mu\text{m}$ bei D_9_2 und 0.5 μm bei den restlichen Transistoren, $t_{ox} = 2.7\text{nm}$, $V_D = 20\text{mV}$ für C_3_1 und 10mV für die restlichen Transistoren.

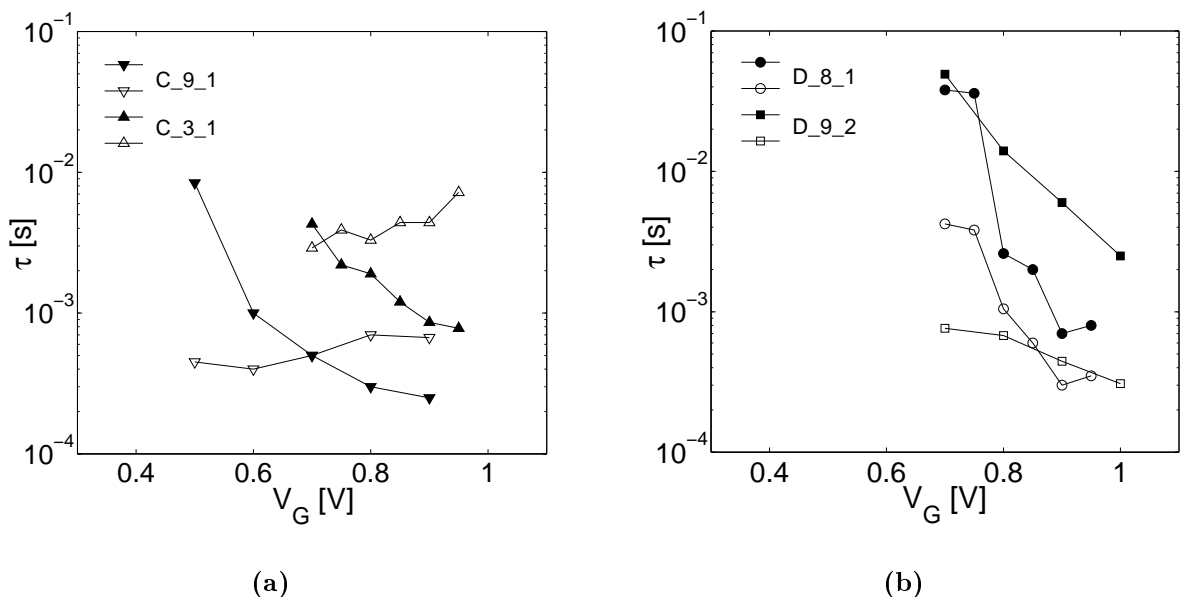


Bild 4.8: (a) und (b): Mittlere Zeitkonstanten „oben“ (gefüllte Symbole) und „unten“ (offene Symbole) der RT-Signale der Transistoren aus Bild 4.7.

stanten „oben“ abfallendes Verhalten, während die Zeitkonstanten „unten“ eher konstant bis leicht ansteigend auf erhöhte Gatespannung reagieren. Dieses Verhalten entspricht dem der klassischen Theorie im Drainstrom von Transistoren. Bild 4.8(b) zeigt jedoch im Unterschied dazu eine gleichzeitige Abnahme beider Zeitkonstanten mit der Gatespannung, die dieser Theorie widerspricht. Wie in Abschnitt 4.1.2 erwähnt, finden sich in der Literatur solche Abhängigkeiten der Zeitkonstanten für MOS-Tunneldioden. Auch in Bipolartransistoren [114,115] und MOS-Transistoren mit Tunneloxiden [116] wurden abnehmende Zeitkonstanten beobachtet.

4.3 Modellentwicklung

Die Situation bei durchgebrochenen Transistoren ist komplex, da in diesem System im Oxid ein neuer Leckstrompfad entstanden ist, durch den ein beträchtlicher Stromfluß zum Gate erfolgt. Die eingetretene Schädigung des Gitters läßt eine hohe Störstellendichte erwarten. Entlang des Durchbruchskanals gelegene Haftstellen können durch Be- und Entladung zu einer Veränderung der Leitfähigkeit des Durchbruchskanals analog zum klassischen, im Abschnitt 4.1.1 betrachteten Inversionskanal führen. Im Gegensatz zum ungeschädigten Transistor kann, wie in Bild 4.9 gezeigt, nach Durchbruch jedoch Einfang und Emission von Ladungsträgern durch die Haftstelle zum Inversions- und zusätzlich zum Durchbruch-

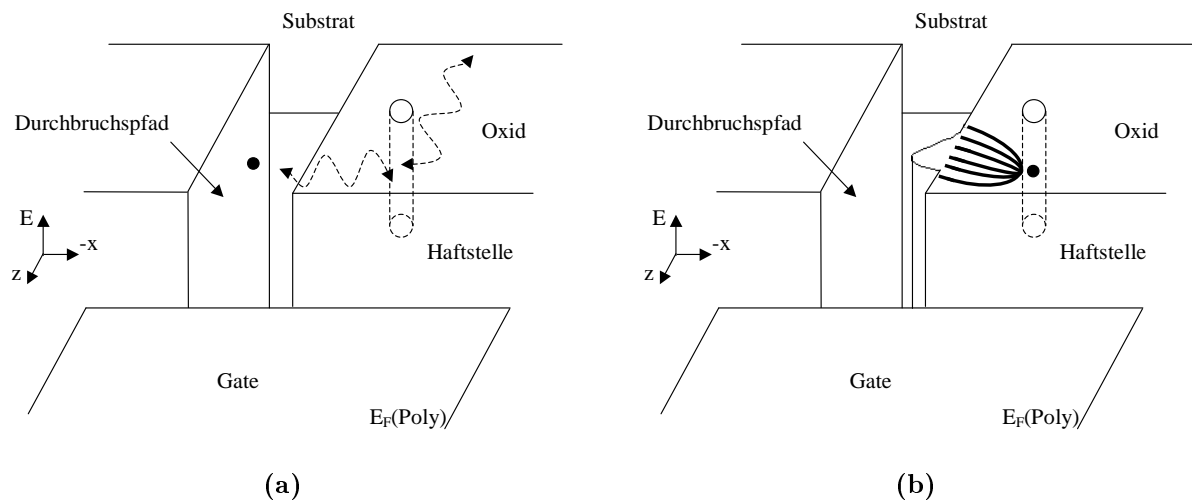


Bild 4.9: Dreidimensionales Energiediagramm des Durchbruchortes von der Gateseite. (a) Einfang und Emission kann je nach Lage der Haftstelle zum Durchbruchskanal oder zum Substrat erfolgen und führt (b) zu einer Störung des Stromflusses im Durchbruchskanal.

kanal erfolgen. Diese zwei Möglichkeiten führen zu den beobachteten zwei Gruppen von experimentellen Daten.

4.3.1 Herleitung aus der Multiphonontheorie

Die Shockley-Read-Hall-Theorie berücksichtigt zur Berechnung der Einfangrate nur die Ladungsträgerdichte und den Einfangwirkungsquerschnitt. In einem MOS-Transistor wird also nach Erreichen der Schwellspannung eine praktisch konstante Einfangzeitkonstante vorausgesagt, da die Ladungsträgerdichte im Inversionskanal bei hohen Spannungen nur noch sehr wenig ansteigt. Die durchgeführten Atlas-Simulationen zeigten, daß selbst im Fall des durchgebrochenen Oxids die Ladungsträgerdichte bei einer Gatespannung zwischen 0.4 und 1V fast konstant ist. Auch Bild 4.7(a) zeigt, daß der Gatestrom für $V_G > 0.5V$ innerhalb einer Größenordnung bleibt, was keinesfalls die beobachtete Abnahme der Einfangzeitkonstanten um mehr als eine Größenordnung erklärt.

Zur Ableitung der Einfang- und Emissionskinetik an der Haftstelle wird hier deswegen auf die Multiphonontheorie zurückgegriffen [117–119]. Diese Theorie beschreibt die Umwandlung der elektronischen Energie in Gitterenergie beim nicht strahlenden Einfang von Ladungsträgern. Da die umgewandelte Energie größer als die Energie eines einzelnen Phonons ist, sind an diesem Prozeß mehrere Phononen beteiligt, woraus die Bezeichnung Multiphononübergang entspringt. Das aus Elektron und den Gitterschwingungen zusammengesetzte System wechselt dabei zwischen freiem und gebundenen Zustand seine Energiekonfiguration. Zur Berechnung der Einfangrate ist, wie das Konfigurationskoordinaten-Diagramm in Bild 4.10 zeigt, die Energiedifferenz zwischen dem Ladungsträger im Leitungsband E_C und dem Energieniveau der Haftstelle E_T sowie die Elektron-Gitter-Kopplung von Bedeutung. Bei schwacher Kopplung, d. h. kleiner Gitterrelaxationsenergie E_R ergibt sich in Bild 4.10(a) eine monotone Zunahme der Einfangbarriere E_B und somit eine Abnahme der Einfangwahrscheinlichkeit, wenn die Energiedifferenz $E_C - E_T$ ansteigt. Bei starker Kopplung hingegen nimmt die Einfangbarriere in Bild 4.10(b) ab und steigt nur für sehr hohe Energiedifferenzen $E_C - E_T$ wieder an. Im Fall der starken Kopplung kann die Einfangwahrscheinlichkeit bei Raumtemperatur sehr sensibel auf eine Änderung von $E_C - E_T$ reagieren und über einige Größenordnungen variieren [118].

Die Multiphonontheorie sagt eine thermische Aktivierung bei hohen Temperaturen und ein temperaturunabhängiges Verhalten bei niedrigen Temperaturen voraus. Sie wird deswegen zur Erklärung der Temperaturabhängigkeit der Zeitkonstanten der RT-Signale verwendet [84, 85]. Bild 4.11 zeigt beispielhaft den großen Einfluß der Temperatur auf die

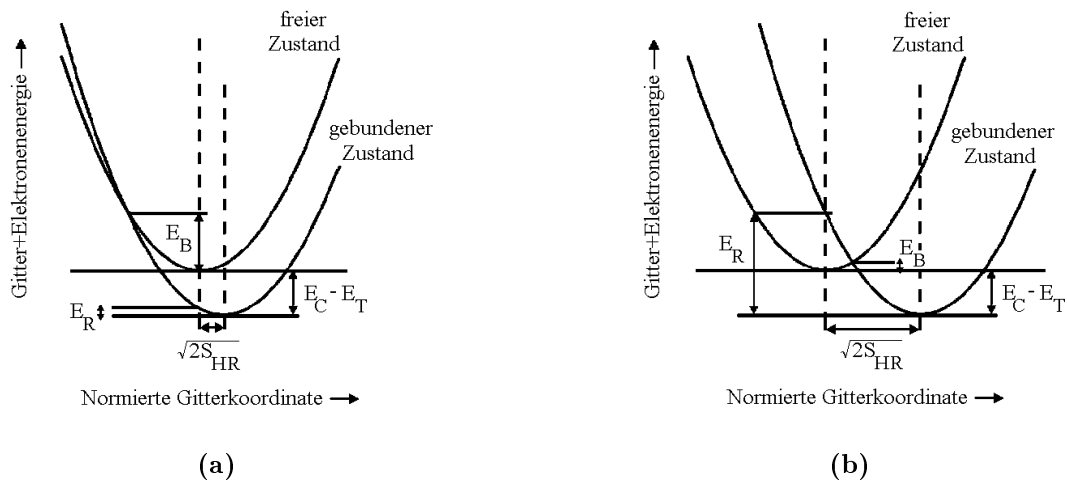


Bild 4.10: Energiekoordinatensystem bei (a) schwacher und (b) starker Elektron-Gitter-Kopplung [119].

Zeitkonstanten schon bei einer relativ geringen Erwärmung. Bei hohen Temperaturen verhält sich der Einfangwirkungsquerschnitt proportional zu $\exp(-E_B/kT)$ [117]. Experimentell ist eine Temperaturunabhängigkeit der Fluktuationen unterhalb von ca. 15K und die thermische Aktivierung oberhalb davon belegt worden [93], Raumtemperatur ist also eindeutig dem Hochtemperaturbereich zuzuordnen.

Die Multiphonontheorie wird nun verwendet, um zusätzlich die Spannungsabhängigkeit der RT-Signale zu modellieren. Im vorliegenden Fall kann eine starke Elektron-Gitter-Kopplung angenommen werden, da der Effekt des Einfangens und der Emission so deutlich im makroskopisch gemessenen Strom sichtbar ist. Der Einfang muß also zu einer starken

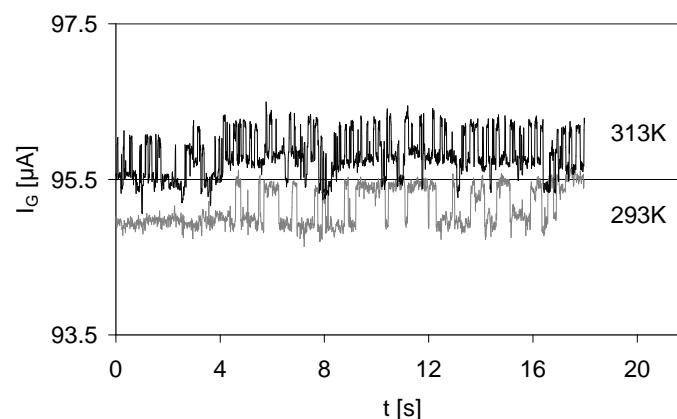


Bild 4.11: RT-Signale im Gatestrom von Probe 08_5_1 aus Bild 4.6 bei $V_G = 0.8V$ und $V_D = 50mV$ und unterschiedlichen Temperaturen.

Verschiebung der Gitterkoordinaten und des Coulombpotentials führen. Davon ist auch auszugehen, da durch den Oxiddurchbruch eine starke Schädigung der Gitterstruktur zu erwarten ist.

Die Einfangwahrscheinlichkeit P_{cap} kann für den Hochtemperaturfall genähert werden als [118]

$$P_{cap} \propto \exp \left[-\frac{(-S_{HR}\hbar\omega + E_C - E_T)^2}{4S_{HR}\hbar\omega kT} \right]. \quad (4.18)$$

Hier ist S_{HR} die Huang-Rhys Koppelkonstante und ω die Phononenfrequenz. In Gleichung 4.18 zeigt sich die bereits erwähnte thermische Aktivierung der Einfangrate. Bei starker Gitterkopplung ist $E_C - E_T < S_{HR}\hbar\omega$ und Gleichung 4.18 kann unter Vernachlässigung des quadratischen Terms vereinfacht werden zu

$$P_{cap} \propto \exp \left(-\frac{S_{HR}\hbar\omega}{4kT} + \frac{E_C - E_T}{2kT} \right). \quad (4.19)$$

Die Einfang- und Emissionsraten R_C and R_E werden deswegen abweichend von Abschnitt 4.1.1 um die Energieabhängigkeit erweitert und wie folgt aufgestellt [120]:

$$R_C = c_n \cdot \exp[(E_C - E_T)/2kT] \cdot n \cdot n_T^0 \quad (4.20)$$

$$R_E = e'_n \cdot n_T^- \quad (4.21)$$

Im Falle eines Volumenhalbleiters ist die Energiedifferenz $E_C - E_T$ konstant und Gleichung 4.20 vereinfacht sich zur bekannten Shockley-Read-Hall-Formel mit angepaßtem Koeffizienten c_n . In einem MOS-System ist diese Energiedifferenz zu einer Haftstelle, die sich innerhalb des Oxids befindet, abhängig von der angelegten Gatespannung und der Entfernung z von der Grenzfläche, wie Bild 4.12 zeigt. Diese kann zu

$$E_C - E_T = E_{CO} - E_{TO} + \frac{z}{t_{ox}} q V_{ox} \quad (4.22)$$

berechnet werden, wobei $E_{CO} - E_{TO}$ die Energiedifferenz bei Flachbandspannung und V_{ox} die Spannung über dem Oxid unter Berücksichtigung der Bandverbiegung ist (s. Gl. 3.16 auf S. 43). Es muß nun die Fallunterscheidung durchgeführt werden, ob der Ladungsträgeraustausch, wie in Bild 4.9(a) gezeigt, zum Substrat oder zum Durchbruchskanal stattfindet.

werden. Da das Fermi-Niveau aber von seinem Wert im Inversionskanal zum Niveau des Gates abfällt, wird die Emissionskinetik beeinflusst und die Gleichung für die Emissionszeitkonstante wird zu

$$\frac{1}{\tau_e} = c_n \cdot \exp\left(\frac{E_C - E_T}{2kT}\right) \cdot n \cdot \exp\left(-\frac{E_F(z) - E_T}{kT}\right), \quad (4.26)$$

wobei $E_F(z)$ das Fermi-Niveau am Ort z im Durchbruchkanal ist. Eine lineare Näherung für den Abfall von der Höhe im Silizium zum Wert im Gate wird wie folgt angesetzt:

$$E_F(z) - E_T = E_{CO} - E_{TO} + (E_F - E_C) \left(1 - \frac{z}{t_{ox}}\right) \quad (4.27)$$

Dies ist eine einfache Näherung, da der genaue Verlauf innerhalb des Durchbruchkanals unbekannt ist und Einfang und Emission im Prinzip zum Gate, zum Kanal und zum Substrat stattfinden kann. Gleichung 4.26 kann vereinfacht werden zu

$$\frac{1}{\tau_e} = c_n \cdot n \cdot \exp\left(\frac{E_C - E_F(z) + E_T - E_F(z)}{2kT}\right). \quad (4.28)$$

4.3.2 Modellimplementierung und Ergebnisse

Zur Implementierung des Modells wird die Näherung für das Oberflächenpotential im invertierten Kanal von van Langevelde und Klaassen verwendet [62]. Die Energiedifferenz $E_C - E_F$ wird in allen Gleichungen hieraus mit $-q\phi_s + E_g/2$ bestimmt (Gl. 3.2 auf S. 32). Das $\mathcal{F}_{1/2}$ Fermi-Dirac Integral wird nach einer Näherung von Blakemore berechnet, welche für alle Energiebereiche gültig ist [121]. Die Elektronendichte ergibt sich dann nach Nicollian und Brews [36, S. 67] zu

$$n = 3.02 \cdot 10^{15} \text{ cm}^{-3} \text{ K}^{-3/2} \cdot \left(T^{3/2} \cdot \frac{\exp[(q\phi_s - E_g/2)/kT]}{1 + 0.27 \exp[(q\phi_s - E_g/2)/kT]} \right). \quad (4.29)$$

Es ist eigentlich nicht notwendig, die Ladungsträgerdichte besonders genau zu modellieren, da sie, wie bereits erwähnt, bei hohen Spannungen praktisch konstant ist. Im relevanten Gatespannungsbereich zwischen 0.4 und 1V ergibt sich, wie Bild 4.13 zeigt, ein Wert von $n \sim 10^{19} \text{ cm}^{-3}$, der durch die Simulationen mit Atlas bestätigt wird. Der für jede Spannung berechnete Wert wurde für den gesamten Durchbruchkanal als konstant angenommen.

Bild 4.14 zeigt die modellierten Verläufe im Vergleich zu den Meßwerten. Beide Arten von experimentellen Ergebnissen werden gut wiedergegeben. Die unterschiedlichen Verläufe

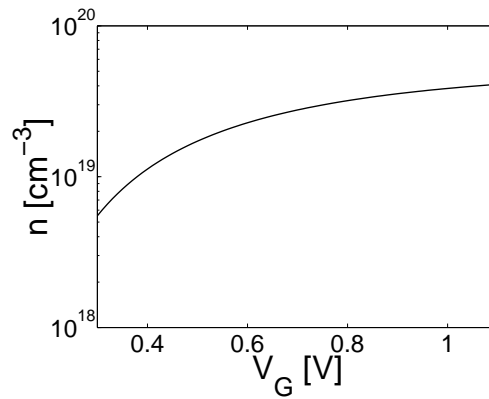
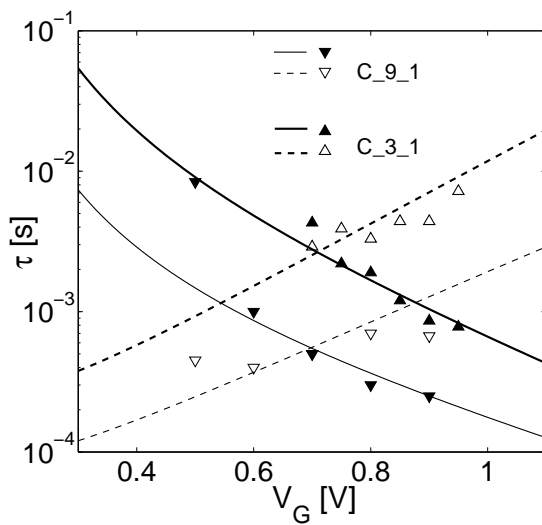
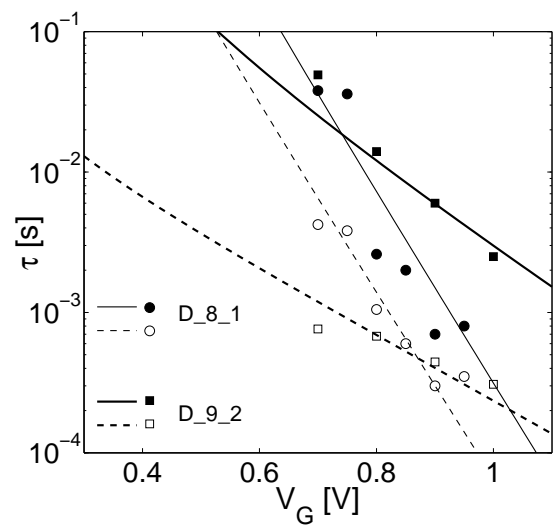


Bild 4.13: Nach Gleichung 4.29 und der Näherung für das Oberflächenpotential in Inversion [62] berechnete Elektronenkonzentration. $t_{ox} = 2.7\text{nm}$, $N_{Substrat} = 3 \cdot 10^{17}\text{cm}^{-3}$, $N_{Poly} = 7 \cdot 10^{19}\text{cm}^{-3}$.



(a)



(b)

Bild 4.14: Ergebnisse für (a) klassisches und (b) abnehmendes Verhalten. Die modellierte Einfang- (durchgezogene Linien) und Emissionszeitkonstanten (gestrichelt) wurden mit den Werten aus Tabelle 4.1 berechnet und mit den gemessenen Zeiten „oben“ (gefüllte Symbole) und „unten“ (offene Symbole) verglichen.

fe der Zeitkonstanten können in einer einheitlichen Modellvorstellung erklärt werden. Die bereits angewendete Interpretation der thermischen Aktivierung [84,85] der Zeitkonstanten aus der Multiphonontheorie wird somit hier um die Spannungsabhängigkeit ergänzt.

Die Modellparameter sind in Tabelle 4.1 aufgeführt. Es ergeben sich negative Energiedifferenzen für $E_{CO} - E_{TO}$, die Haftstelle liegt also bei Flachbandspannung über dem Leitungsband und wird somit erst aktiv, wenn die Gatespannung groß genug ist, um das Haftstellenniveau energetisch nahe an das Leitungsbandniveau zu ziehen. Auch dies wird durch die Meßwerte gestützt, bei denen unterhalb von 0.4V keine Fluktuationen gemessen wurden.

Tabelle 4.1: Verwendete Parameter des Modells für die Kurven aus Bild 4.14; $t_{ox} = 2.7\text{nm}$ für alle Berechnungen.

Transistor	Entfernung z [nm]	c_n [cm ³ /s]	$E_{CO} - E_{TO}$ meV	Emission zum...
C_3_1	0.6	$2.5 \cdot 10^{-17}$	-170	Inversionskanal
C_9_1	0.4	$1.2 \cdot 10^{-16}$	-130	Inversionskanal
D_8_1	2.2	$7.1 \cdot 10^{-23}$	-50	Durchbruchspfad
D_9_2	0.9	$1.3 \cdot 10^{-19}$	-100	Durchbruchspfad

4.3.3 Amplituden

Folgt man dem vorgestellten Modell, so können auch die relativen Amplituden aus Bild 4.7(b) wie folgt erklärt werden. Bei den Transistoren mit klassischem Verhalten „C“ erfolgt der Ladungsträgeraustausch mit dem Inversionskanal, während er bei den Proben mit abnehmendem Verhalten „D“ zum Durchbruchspfad erfolgt. In beiden Fällen wird der Stromfluß durch den Durchbruchkanal nach Besetzung der Haftstelle behindert. Es ergibt sich eine gleichzeitige Abnahme des Gatestroms und eine Zunahme des Drainstroms. Da ballistischer Transport durch den Durchbruchspfad angenommen wird, ist das durch die Besetzung der Haftstelle entstehende energetische Hindernis um so geringer, je niedriger das Energieniveau der Haftstelle ist.

Befinden sich die Haftstellen am Eingang des Durchbruchkanals, so hat eine Änderung der Gatespannung kaum Auswirkung auf die energetische Lage der Haftstelle relativ zum Leitungsband des Substrats (s. auch Bild 4.12). In diesem Fall ergeben sich konstante relative Amplituden. Je weiter entlang des Durchbruchkanals sich die Haftstelle befindet, desto energetisch niedriger liegt die Haftstelle gegenüber dem Leitungsband des Substrats

mit steigender Oxidspannung. Da die Auswirkung der Spannungsänderung auf die energetische Lage größer ist, ergeben sich mit höherer Gatespannung abnehmende relative Amplituden. Dies zeigt sich deutlich bei einem Vergleich der beiden Proben D_8_1 und D_9_2 in Bild 4.7(b). Die weiter entfernte Haftstelle unterliegt einer deutlich stärkeren Abhängigkeit der relativen Fluktuationsamplitude von der Gatespannung als die weiter vorn liegende. Die mit Hilfe der Analyse der Zeitkonstanten gewonnenen Erkenntnisse werden also durch die beobachteten Fluktuationsamplituden gestützt.

4.4 Zusammenfassung

Rauschen bleibt ein spannendes und schwieriges Forschungsfeld, welches bei skalierten Bauelementen mit Tunneloxiden möglicherweise eine größere Bedeutung erlangen wird. Im Rahmen dieser Arbeit wurden sogenannte Telegraphensignale (RTS) bzw. $1/f^2$ -Rauschen an Transistoren mit durchgebrochenen Gateoxiden untersucht. Korrelierte Rauschmessungen an mehreren Elektroden wurden als Hilfe zur Lokalisierung des Durchbruchs verwendet. Gerade bei Tunneloxiden wird das Erkennen eines Durchbruchs aus der I-U-Kennlinie erschwert, so daß sich Rauschmessungen als zusätzliche Charakterisierungsmethode anbieten. Hierfür ist es notwendig, Modelle für das zu erwartende Rauschen bei unterschiedlichen Konfigurationen zur Verfügung zu haben. Das unterschiedliche Verhalten der Zeitkonstanten in Abhängigkeit von der angelegten Gatespannung konnte mit einem Modell, was aus der Multiphonon- und Shockley-Read-Hall-Theorie abgeleitet wurde, sehr gut beschrieben werden. Dieses Modell klärt einige noch offene Fragen von RTS im Drainstrom und kann die zusätzlichen Beobachtungen bei durchgebrochenen Oxiden erklären.

5

Oxiddurchbrüche in Schaltungen

Mit Hilfe der in Kapitel 3 und 4 abgeleiteten Modelle kann die Auswirkung von Durchbrüchen auf ganze Schaltungen untersucht werden. Die dadurch gewonnenen Erkenntnisse können zur Einleitung geeigneter Gegenmaßnahmen zur Entwicklung fehlertoleranter Schaltungen verwendet werden.

Die Ersatzschaltbilder der durchgebrochenen Transistoren wurden in Gleichungsform in der Hardwarebeschreibungssprache Spectre HDL erzeugt. Spectre HDL eröffnet die Möglichkeit, Blöcke in eine SPICE-Simulationsumgebung einzufügen, die durch rein mathematische Gleichungen beschrieben werden [122, 123]. Es entsteht so ein Bauelement, dessen Strom- und Spannungszusammenhänge an den Klemmen durch frei definierbare mathematische Gleichungen bestimmt werden. Dies erleichtert einerseits den Umgang mit der Simulationsumgebung, da zur Simulation eines Durchbruchs nicht mehr mehrere Bauelemente in einer bestehenden Schaltung richtig verdrahtet werden müssen, um den Durchbruch zu simulieren. Andererseits ist die Modellierung auf diese Weise deutlich flexibler und kann auch auf kompliziertere Kennlinien erweitert werden, die durch Standard-SPICE-Bauelemente nur schwer zu realisieren sind.

Bild 5.1 zeigt das Blockschaltbild für einen durchgebrochenen Transistor. Es besteht einerseits aus den Transistorgleichungen mit degradierten Parametern für den Drain-Source-Feldeffektstrom und dem Modell für den Durchbruchspfad von Gate zu Drain oder Source. Bei einem harten Durchbruch zum Substrat werden nach Kaczer [68] zusätzlich zwei Bipolartransistoren und zwei Feldeffekttransistoren eingesetzt. Die recht hohe Anzahl von Bausteinen bei diesem Modell hat zahlreiche einstellbare Parameter zur Folge, die nur schwer zu bestimmen sind. Ebenso ergaben sich zum Teil Konvergenzprobleme. Einfachere Modelle mit weniger Parametern wären deshalb möglicherweise besser geeignet. Dieser

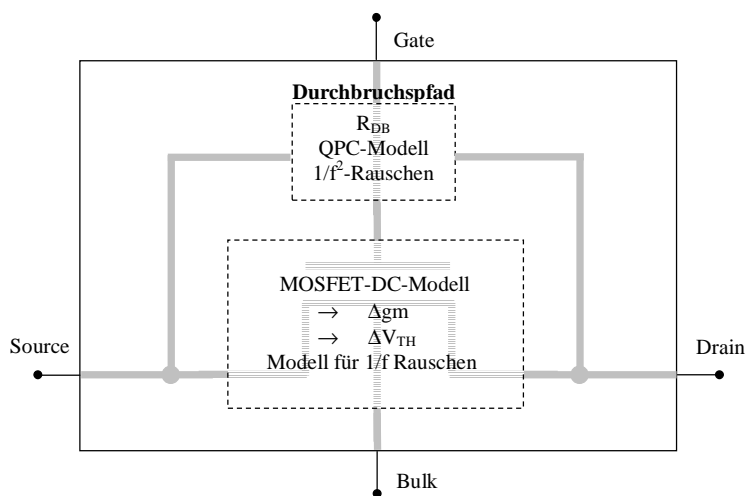


Bild 5.1: Blockschaltbild für einen durchgebrochenen Transistor.

Durchbruchstyp wurde deswegen nicht in den folgenden Untersuchungen berücksichtigt.

Für den Drainstrom wurde ein Modell nach Arora [124] definiert. Es ähnelt dem Level2-SPIICE-Modell, enthält also eine Beschreibung für den Subschwelligbereich. Die Änderung der MOS-Transistorkennlinien gegenüber dem ungeschädigten Fall wurde für harten und weichen Durchbruch durch eine Verdopplung der Schwellspannung und eine Abnahme der Steilheit um 20% erfasst. Ebenfalls wurde ein einfaches empirisches Modell für das $1/f$ Rauschen eingesetzt. Es erscheint nicht sinnvoll, ein komplizierteres Modell mit vielen Parametern für die durchgebrochenen Transistoren zu verwenden, da durch die statistische Verteilung der Durchbruchcharakteristik ohnehin eine starke Streuung zu erwarten ist.

Der Durchbruchpfad wurde beim harten Durchbruch durch einen Widerstand $R_{DB} = 10\text{k}\Omega$ modelliert. Für den weichen Durchbruch wurde das Quantenpunktmodell (QPC) verwendet. Das Auftreten von Random Telegraph Signals im Gatestrom wurde mit einem $1/f^2$ -förmigen Spektrum berücksichtigt. In Anhang B sind die verwendeten Modellgleichungen aufgeführt.

Gute Modelle sind somit zum Gateleckstrom sowie dessen Rauschen vorhanden, bei den weiteren Faktoren sind ungefähre Größenordnungen bekannt. Hiermit ist die Erstellung von „worst case“-Szenarien und eine Abschätzung der Auswirkung einzelner Faktoren auf die Schaltungen möglich. Die erzeugten Blöcke wurden nun testweise in analoge und digitale Schaltungen eingefügt und deren Funktion bei Durchbruch getestet.

5.1 Analoge Verstärker

Zwei Verstärkertypen wurden testweise ausgewählt und durchgebrochene Transistoren in deren Eingangsstufen eingesetzt. Simuliert wurde in einer $0.35\mu\text{m}$ AMS-Technologie mit $V_{DD}/V_{SS} = \pm 1.65\text{V}$. Obwohl bei dieser Technologie mit $t_{ox} = 7.5\text{nm}$ ein weicher Durchbruch in der Realität nicht auftritt, dienen die Simulationen als Anhaltspunkt für das Verhalten zukünftiger Generationen.

5.1.1 Differenzverstärker

In Bild 5.2 ist ein einfacher Differenzverstärker ohne Frequenzkompensation gezeigt. Beim weichen Durchbruch zeigt sich, daß der zusätzliche Gateleckstrom allein keine Änderung der Verstärkerkennlinien bewirkt. Der entscheidende Faktor für das Verstärkerverhalten ist die

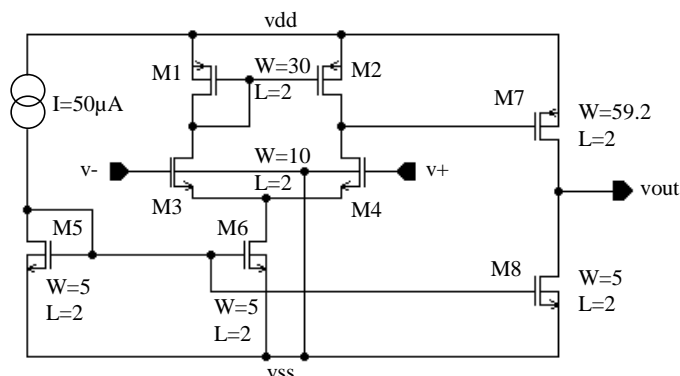


Bild 5.2: Differenzverstärker mit unkompensierter Ausgangsstufe, L und W in μm .

Schwelspannungsverschiebung, die zu Nichtübereinstimmung (Mismatch) der gleichartigen Transistorpaare führt. Matching ist eine Grundvoraussetzung für das Funktionieren der in analogen Schaltungen verwendeten Differenzstufen und Spiegelschaltungen [25, 125, 126]. Der prozeßbedingt entstehende Mismatch muß deswegen in einem robusten Design berücksichtigt werden. Zusätzlich kann noch in der Fertigung ein Feinabgleich durchgeführt werden, bevor das Bauelement ausgeliefert wird. Bei unsymmetrischer Belastung von Transistorpaaren kann allerdings im Betrieb Mismatch entstehen und zum Ausfall der Schaltung führen [127, 128]. Der Gateoxiddurchbruch eines Transistors stellt in dieser Hinsicht den schwersten Fall von Ungleichheit dar. Mismatch der Transistorpaare in der Differenzstufe beeinflusst in erster Linie den Offsetfehler [126].

In Bild 5.3 ist die Auswirkung von weichem und hartem Durchbruch im Vergleich dargestellt, wobei beim harten Durchbruch zur Veranschaulichung nur der zusätzliche Gatestrom

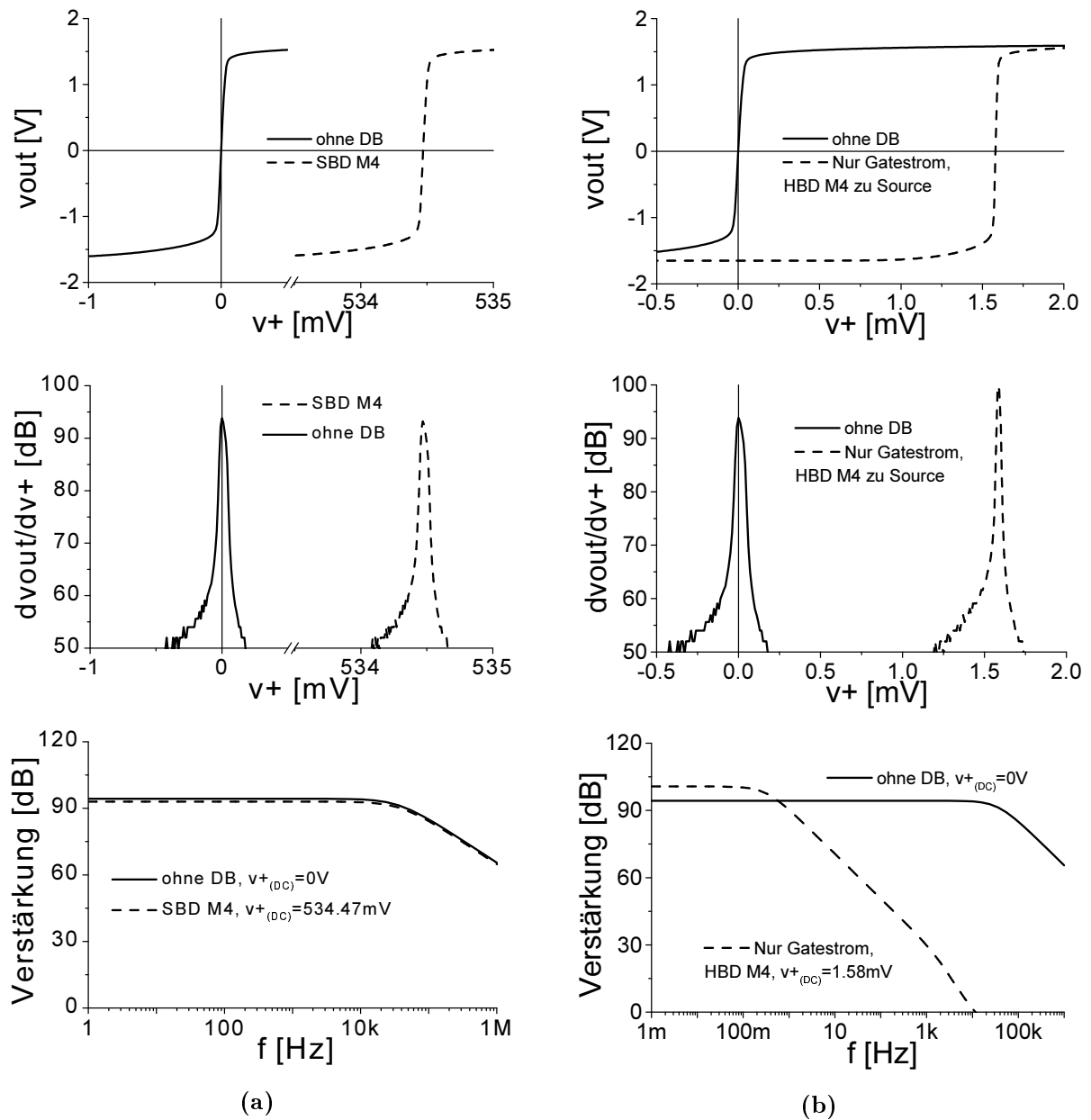


Bild 5.3: Simulierte DC-Kennlinie (oben), Ableitung der DC-Kennlinie (Mitte) und AC-Verstärkung nach Offsetkompensation am Eingang des Differenzverstärkers (unten) mit $v_-=0V$ vor und nach Durchbruch des Transistors M4, (a) weicher Durchbruch unter Berücksichtigung einer Schwellspannungsverdopplung und Steilheitsminderung um 20%, (b) harter Durchbruch zu Source unter Berücksichtigung nur des zusätzlichen Gateleckstroms.

berücksichtigt wurde. Am Augenscheinlichsten zeigt sich in beiden Fällen der entstehende Offset, der entweder durch die Schwellspannungsänderung (Bild 5.3(a)) oder den hohen Gateleckstrom (Bild 5.3(b)) verursacht wird. Schon eine geringe Schwellspannungsänderung im Millivoltbereich führt auch zu einem Offset im Millivoltbereich und somit zu einem Totalausfall der Schaltung. Die weitere Funktion des Verstärkers kann also nur über eine intelligente Offsetkorrektur während des Betriebs gewährleistet werden. Je nach Anwendung könnte dies über die bekannten Autozeroing und Chopper-Verfahren, über Floating-Gate Transistoren [129] oder durch eine einstellbare DC-Offsetspannung am Eingang erfolgen. Bild 5.3 zeigt, daß auch nach der Offsetkorrektur eine Veränderung der AC-Kennlinien festzustellen ist. Im Fall des weichen Durchbruchs ergibt sich durch die Steilheitsänderung ein vergleichsweise geringer Abfall der Verstärkung und eine weitgehend erhaltene Funktion. Je höher jedoch der parasitäre Gateleckstrom wird, desto stärker ist dessen Einfluß auf die AC-Kennlinien und desto mehr muß davon ausgegangen werden, daß der Operationsverstärker im praktischen Betrieb unbrauchbar wird. Der zusätzliche Strom führt zu einer Verschiebung des Arbeitspunktes der Schaltung, so daß in Bild 5.3(b) bei niedrigen Frequenzen sogar eine Zunahme der Verstärkung zu erkennen ist, das Verstärkungsbandbreiteprodukt nimmt jedoch stark ab.

Besonders ins Gewicht fällt der Durchbruch des Eingangstransistors bei hochohmigen Signalquellen, für die Verstärker mit FET-Eingang wegen ihres hohen Gatewiderstandes bevorzugt verwendet werden [130]. Hier zeigt Bild 5.4 einen starken Abfall der Verstärkung auch bei einem weichen Durchbruch, je höher der Innenwiderstand der Quelle ist. Bei einer komplexen Innenimpedanz wird zusätzlich die Phasenlage beeinflusst.

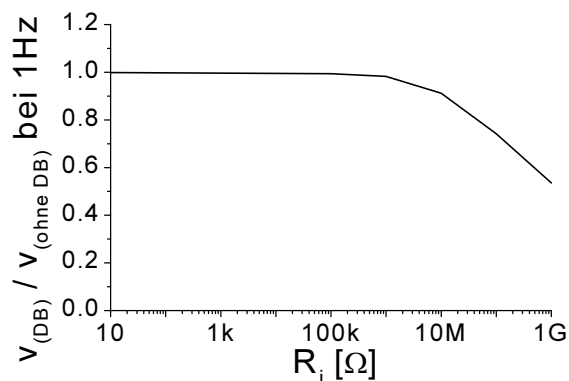


Bild 5.4: Auswirkung des Innenwiderstandes R_i der Signalquelle an v_+ auf die Verstärkung nach Offsetkompensation bei weichem Durchbruch von M4 und Berücksichtigung nur des Gateleckstroms.

5.1.2 Stromverstärker

Der in Abschnitt 2.3 vorgestellte Stromverstärker beruht in hohem Maße auf Stromspiegelschaltungen. Bild 5.5 zeigt den Verstärker ohne Ausgangsstufe. Wird bei der Simulation

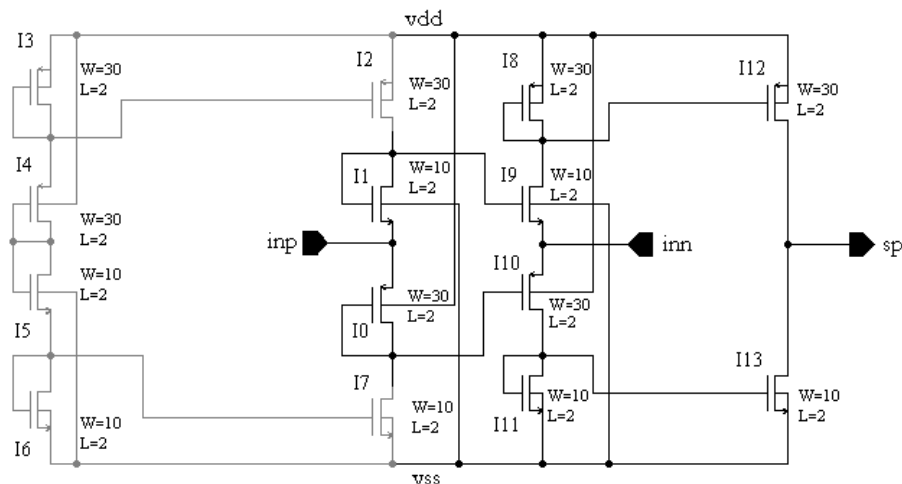


Bild 5.5: I-U-Verstärker, L und W in μm , in grau sind die zur Erzeugung eines konstanten Stroms verwendeten Transistoren eingezeichnet.

des Stromverstärkers mit weich durchgebrochenen Bauelementen ausschließlich der zusätzliche Gatestrom betrachtet, so zeigt sich auch hier ähnlich wie beim Differenzverstärker in den meisten Fällen kein Einfluß. Auch hier gilt jedoch, daß hauptsächlich der entstehende Offsetfehler, bedingt vor allem durch die Änderung der Schwellspannung, zum Ausfall der Schaltung führt. Ist einer der Transistoren in Spiegelschaltung I11-I13 oder I8-I12 betroffen, so zeigt sich schon bei einer Schwellspannungsänderung von einigen Millivolt ein hoher Offset. Werden die Transistoren I1-I9 bzw. I0-I10 betrachtet, so bewirkt ein Durchbruch eine Verschiebung des Arbeitspunktes der Schaltung und somit eine Änderung der Verstärkung. Bild 5.6 zeigt die Auswirkung eines weichen Durchbruchs von I1 bzw. I9, wo der Offset nur sehr wenig beeinflusst wird, aber die Verstärkung bei niedrigen Frequenzen durch die Arbeitspunktverschiebung sogar ansteigen kann. Hierdurch ändert sich auch das Rauschen. Solange die Höhe des Gateleckstroms unterhalb des aktiven Drainstroms liegt, ist dessen Rauschen im Vergleich zur Gesamtrauschleistung allerdings vernachlässigbar. Um das gesamte Rauschen im Transistor nach Durchbruch zu modellieren, müßte deshalb zusätzlich die Erhöhung des Rauschens im Drainstrom erfaßt werden.

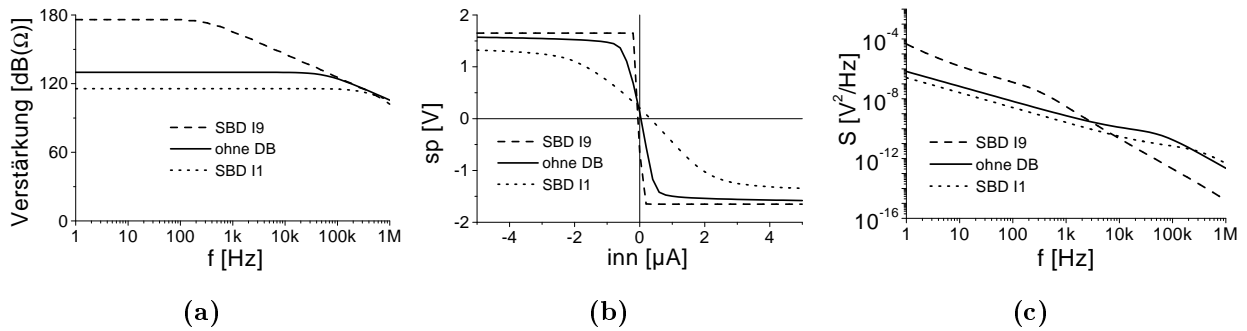


Bild 5.6: Simulierte Kennlinien des Verstärkers ohne und nach weichem Durchbruch von I1 bzw. I9 bei $inn=0V$, die AC- und Rauschkennlinien gelten für $inn=0A$.

5.2 Digitale Gatter

Zur Einschätzung der Auswirkung von Durchbrüchen auf digitale Schaltungen wurden die in Bild 5.7 gezeigten statischen NAND- und NOR-Gatter in CMOS-Technologie verwendet. Die Gatter wurden in $0.35\mu m$ AMS-Technologie in Mindestabmessungen mit $V_{DD} = 3.3V$ simuliert und testweise durchgebrochene Transistoren eingefügt. Kriterium für die weitere Funktion der Gatter war eine maximale Abweichung um 10% von V_{DD} . Für eine logische „1“ waren also minimal 2.97V und für eine logische „0“ maximal 330mV zulässig.

Bei den Simulationen mit harten Durchbrüchen ergibt sich ein gemischtes Bild. Tabelle 5.1 zeigt, daß die weitere Funktion des Gatters von Transistor und Durchbruchsort

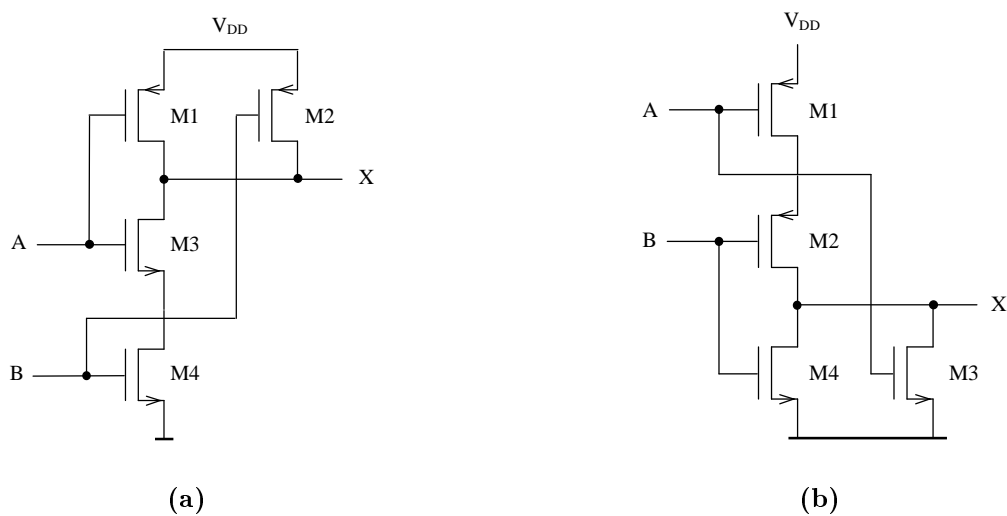
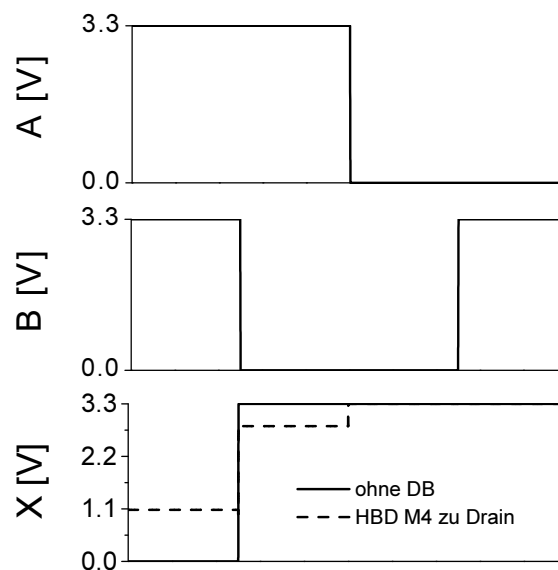


Bild 5.7: (a) NAND- und (b) NOR-Gatter in CMOS-Technologie

Tabelle 5.1: Funktion der Gatter aus Bild 5.7 nach hartem Durchbruch zu Drain und Source:

Transistor	NAND-Gatter mit Durchbruch des Transistors zu		NOR-Gatter mit Durchbruch des Transistors zu	
	Drain	Source	Drain	Source
M1	nein	ja	nein	ja
M2	nein	ja	nein	nein
M3	nein	nein	nein	ja
M4	nein	ja	nein	ja

abhängt. Die Ergebnisse lassen sich verallgemeinern durch die Feststellung, daß ein Durchbruch um so größere Auswirkungen hat, je weiter entfernt von der Versorgungsspannung er auftritt. Ein Durchbruch der Transistoren M3 beim NAND bzw. M2 beim NOR-Gatter führt deswegen immer zu einer Fehlfunktion. Bild 5.8 zeigt die Spannungen am NAND-Gatter nach einem Durchbruch von M4 zum Drain, in diesem Fall werden zwei logische Kombinationen falsch ausgegeben.

**Bild 5.8:** Simulation der Spannungspegel am NAND-Gatter ohne und mit hartem Durchbruch.

Um weniger ideale Bedingungen zu simulieren, wurde eine Last von $100\text{k}\Omega$ an den Ausgang hinzugefügt. Dies änderte jedoch ebenso wenig wie das Hinzufügen eines Innenwiderstandes bei den Spannungsquellen von 50Ω die Ergebnisse wesentlich. In günstigen Fällen werden die logischen Operationen also auch nach einem harten Durchbruch richtig ausgeführt.

Bei einem weichen Durchbruch sind die Gatter in jedem Fall weiterhin voll funktionsfähig. Im Gegensatz zu den analogen Schaltungen hat bei den digitalen Schaltungen weder der zusätzliche Gatestrom noch die Änderung der Schwellspannung entscheidende Auswirkung auf die prinzipielle Funktion, da stets mit den vollen Spannungspegeln operiert wird. Das Augenmerk muß hier also auf die Gatterlaufzeiten gelegt werden. Bild 5.9 zeigt, daß die Schaltgeschwindigkeit zum Erreichen der vollen Spannungspegel reduziert wird. Der

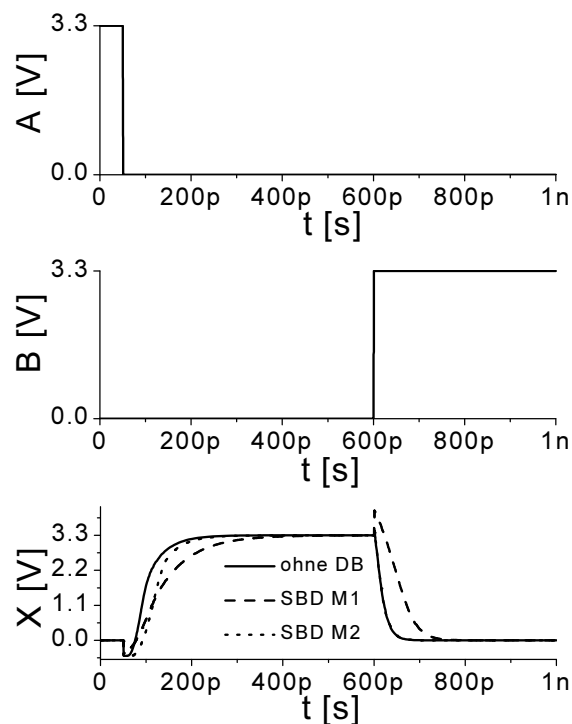


Bild 5.9: Simulierter Spannungsverlauf am NOR-Gatter ohne und mit weichem Durchbruch.

Durchbruch von M1 führt zu einer langsameren Be- und Entladung auf Maximal- bzw. Minimalpegel, während der Durchbruch von M2 sich nur auf die Aufladezeitkonstante auswirkt. Je nach Lage des Durchbruchs und den an den Eingängen anliegenden Spannungen ergibt sich damit der Einfluß auf die Zeitkonstanten. Bei digitalen Schaltungen muß deshalb außer auf die Signalpegel auch auf die noch erreichbaren Geschwindigkeiten und die Auswirkung von unterschiedlichen Gatterlaufzeiten [131] nach Oxiddurchbruch geachtet werden.

5.3 Fazit

Ähnlich wie die Bauelemente selber verlieren auch die aus Transistoren zusammengesetzten Schaltungen ihre Funktion bei einem Gateoxiddurchbruch nicht vollständig. Bei den Auswirkungen der Durchbrüche muß zwischen den unterschiedlichen Faktoren Gateleckstrom, Schwellspannungsverschiebung, Steilheitsminderung und Rauschen unterschieden werden.

Bei analogen Schaltungen tritt zunächst in den meisten Fällen ein Offsetfehler ein, der hauptsächlich durch die Schwellspannungsverschiebung des durchgebrochenen Transistors entsteht. Ohne Offsetkorrektur sind diese Schaltungen nach einem Durchbruch in den meisten Fällen nicht funktionsfähig. Hierin liegt also die Herausforderung, wenn eine fehlertolerante bzw. fehlerkorrigierende Schaltung entwickelt werden soll. Selbst nach der Offsetkompensation bleibt ein Einfluß auf das AC-Verhalten über die Steilheit und den Gateleckstrom vorhanden. Die Funktionsfähigkeit hängt in diesem Fall von der geforderten Bandbreite und Verstärkung der jeweiligen Anwendung ab. Bei einem Durchbruch der Eingangstransistoren spielt die Innenimpedanz der treibenden Signalquelle ebenfalls eine entscheidende Rolle. Je nach Einsatzgebiet muß hieraus eine Vorgabe definiert werden, die nach Durchbruch noch zu erfüllen ist.

Bei digitalen Schaltungen ergibt sich bei harten Durchbrüchen in vielen Fällen eine Fehlfunktion, die nicht korrigierbar ist. Abgesehen von der Funktion an sich, d. h. dem Erreichen eines bestimmten Spannungspegels, liegt hier das Problem in der reduzierten Schaltgeschwindigkeit. Dies gilt besonders bei den Schaltungen mit weichem Durchbruch, die zwar alle funktionieren, jedoch durch den Durchbruch langsamer werden. Auch hier muß die Vorgabe durch die jeweilige Anwendung erfolgen.

Innerhalb gewisser Grenzen scheint somit eine durchbruchstolerante oder zumindest -korrigierende Schaltung möglich zu sein. Ein großes Problem einer solchen Schaltung dürfte allerdings die erhöhte Leistungsaufnahme nach Durchbruch sein. Auch ist eine AC-Charakterisierung der durchgebrochenen Transistoren notwendig, um das Hochfrequenzverhalten zu untersuchen.

6

Zusammenfassung und Ausblick

Die CMOS-Technologie ist aufgrund der fortschreitenden Miniaturisierung bei Strukturgrößen im Nanometerbereich angelangt, was neue physikalische Effekte mit sich bringt. Die Bauelemente werden fragiler und ihre Belastung nimmt zu, so daß die Zuverlässigkeitsreserven abnehmen. Dabei wird die Dicke des für den Betrieb der MOS-Transistoren elementar wichtigen Gateoxids bald den Bereich weniger Atomlagen erreichen. Gateoxid-durchbrüche könnten deshalb in Zukunft schon während der geplanten Lebensdauer von 10 Jahren auftreten.

In dieser Arbeit wurde deshalb der Frage nachgegangen, wie und ob ein im herkömmlichen Sinne zerstörtes Bauelement in einer realen Schaltung zu betreiben ist. Kenntnis über das Verhalten eines MOS-Transistors mit Gateoxiddurchbruch und der Elemente, aus denen er besteht - Halbleiter, Isolator, Durchbruchskanal und deren Übergänge - sind hierfür notwendig. Zunächst wurden daher allgemein die Schwierigkeiten bei Anwendung von Charakterisierungsverfahren auf miniaturisierte Bauelemente erörtert, mit denen die Qualität des Isolators und der Isolator-Halbleiter-Grenzschicht ermittelt werden kann. Hierbei wurde ein Stromverstärkerkonzept vorgestellt, mit dem die zu messenden Signale bereits auf dem Chip verstärkt und somit die Hoffnung einer höheren Auflösung verbunden ist.

Bei Betrachtung der Durchbrüche konnten unterschiedliche Durchbrucharten identifiziert werden, die sich durch ihre Lage und Härte, also der Ausprägung des Durchbruchskanals unterscheiden. Anschließend erfolgte eine Klassifizierung, für die unterschiedliche Modelle und Ersatzschaltbilder entwickelt wurden. Dabei wurde festgestellt, daß die Kennlinien des Transistors sich zwar stark verändern, in gewissen Grenzen jedoch noch von einer Funktion des Bauelements gesprochen werden kann. Um so mehr stellt sich die Frage, inwieweit ganze Schaltungen von einem Durchbruch eines Transistors betroffen sind.

Bei skalierten Bauelementen muß zusätzlich berücksichtigt werden, daß der Einfluß einzelner Defekte immer stärker in den Vordergrund tritt. Einzelne Störstellen wirken sich auch makroskopisch sehr deutlich im Strom des Transistors aus. Bei durchgebrochenen Transistoren ist daher ein Charakteristikum die Zunahme des Rauschens und das Auftreten von Telegraphensignalen (Random Telegraph Signals) bzw. $1/f^2$ Rauschen im Gatestrom. Die Fluktuationen im Strom können dabei auf einen dominierenden Defekt zurückgeführt werden, der beim Durchbruch entstanden ist. Es wurde ein Modell vorgestellt, mit dem unterschiedliche Arten von Telegraphensignalen erklärt und Aufschluß über energetische und geometrische Lage des Defekts gewonnen werden kann. Die Abhängigkeit der Zeitkonstanten der Fluktuationen und somit des Rauschspektrums von der angelegten Gatespannung kann damit berechnet werden.

Die entwickelten Modelle für die Transistorcharakteristik nach Durchbruch wurden abschließend in ausgewählten Schaltungen eingesetzt, um die Auswirkung von Transistordurchbrüchen zu untersuchen. Dabei wurde festgestellt, daß sich bei digitalen Schaltungen ein Durchbruch in erster Linie auf die Schaltgeschwindigkeit auswirkt, während bei analogen Verstärkern hauptsächlich der Offset beeinflußt wird. Dabei könnten, abhängig von der Anwendung und bei geeigneten Gegenmaßnahmen, die Schaltungen weiterhin funktionsfähig bleiben. Mit dem gezeigten Ansatz kann den Schaltungsentwicklern ein Werkzeug an die Hand gegeben werden, mit dem sie eine solche Analyse bereits im Entwicklungsstadium durchführen können.

Für eine zuverlässige Vorhersage des Verhaltens von Schaltungen nach Durchbruch müßten die in dieser Arbeit ausschließlich an Einzeltransistoren durchgeführten Untersuchungen auf Testschaltungen erweitert werden. Ebenso müßten weitere Aspekte des Transistorverhaltens nach Oxiddurchbruch, wie zum Beispiel die Degradation der Parameter Steilheit und Schwellspannung oder auch das Hochfrequenzverhalten, systematisch analysiert werden. Damit könnte das Bild abgerundet und eine Vorhersage über das Verhalten zukünftiger Transistorgenerationen ermöglicht werden.

Näherung für die LambertW(x) Funktion

Näherungen für die *LambertW* Funktion wurden aus verschiedenen Quellen zusammengestellt und die Übereinstimmung mit der iterativen Lösung überprüft. Daraus ergaben sich drei Teilbereiche, in denen eine ausreichende Genauigkeit der jeweiligen Näherung festzustellen war. Verwendet wurden folgende Gleichungen:

Teilbereich 1, $x > 109.806996$

(Von <http://mathworld.wolfram.com/LambertsW-Function.html>)

$$k_1 = \ln(x), \quad k_2 = \ln(\ln(x))$$

$$LambertW = k_1 - k_2 + \frac{k_2}{k_1} + k_2 \frac{k_2 - 2}{2k_1^2} + k_2 \frac{6 - 9k_2 + 2k_2^2}{6k_1^3} + k_2 \frac{-12 + 36k_2 - 22k_2^2 + 3k_2^3}{12k_1^4} + k_2 \frac{60 - 300k_2 + 350k_2^2 - 125k_2^3 + 12k_2^4}{60k_1^5}$$

Teilbereich 2, $0.158730666 < x < 109.806996$

(Von <http://www.desy.de/~t00fri/qcdins/texhtml/lambertw/>)

$$LambertW = 0.665 \cdot [1 + 0.0195 \ln(x + 1)] \cdot \ln(x + 1) + 0.04$$

Teilbereich 3, $x < 0.158730666$

(Von <http://mathworld.wolfram.com/LambertsW-Function.html>)

$$LambertW = x - x^2 + 1.5x^3 - \frac{8}{3}x^4 + \frac{125}{24}x^5 - \frac{54}{5}x^6 + \frac{16807}{720}x^7$$

B

MOS-Transistormodell in Spectre HDL

In SpectreHDL wurden für die zu modellierenden Transistoren einfache Modelle eingesetzt, um die Übersichtlichkeit zu gewährleisten. Die Übereinstimmung mit den in Cadence implementierten genaueren BSIM Modellen war ausreichend. Da der Durchbruch eine statistische Unsicherheit beinhaltet, ist eine genauere Modellierung der Kennlinien nicht sinnvoll.

DC-Drainstrom

Für den DC-Teil wird eine Näherung nach Arora [124, S. 268] verwendet, die eine Beschreibung im Subschwelligbereich enthält. Hierbei wird $V_{on} = V_{TH} + \eta V_t$ festgelegt, wobei V_t die Temperaturspannung, V_{TH} die Schwellspannung und $\eta = 1.5$ ein Parameter ist, der die kapazitive Kopplung zwischen Gate und der Siliziumoberfläche angibt. Das Modell ist über den gesamten Spannungsbereich stetig.

Negative Spannung, $V_{GS} < 0$

$$I_D = 0$$

Linearer Subschwelligbereich, $V_{GS} < V_{on}$ & $V_{DS} < V_{GS} - V_{TH}$

$$I_D = K \cdot \left(V_{on} - V_{TH} - \frac{V_{DS}}{2} \right) \cdot V_{DS} \cdot \exp \left(\frac{V_{GS} - V_{on}}{\eta V_t} \right) \cdot (1 + \lambda V_{DS})$$

Subschwelligbereich in Sättigung, $V_{GS} < V_{on}$ & $V_{DS} > V_{GS} - V_{TH}$

$$I_D = \frac{K}{2} \cdot (V_{on} - V_{TH})^2 \cdot V_{DS} \cdot \exp \left(\frac{V_{GS} - V_{on}}{\eta V_t} \right) \cdot (1 + \lambda V_{DS})$$

Linearer Bereich, $V_{GS} > V_{on}$ & $V_{DS} < V_{GS} - V_{TH}$

$$I_D = K \cdot \left(V_{GS} - V_{TH} - \frac{V_{DS}}{2} \right) \cdot V_{DS} \cdot (1 + \lambda V_{DS})$$

Sättigungsbereich, $V_{GS} > V_{on}$ & $V_{DS} > V_{GS} - V_{TH}$

$$I_D = \frac{K}{2} \cdot (V_{GS} - V_{TH})^2 (1 + \lambda V_{DS})$$

Für die simulierte 0.35µm Technologie ist $t_{ox} = 7.5\text{nm}$, $V_{TH}(\text{nMOS}) = 0.5\text{V}$ und $V_{TH}(\text{pMOS}) = -0.65\text{V}$.

Weiterhin ist $\lambda = 0.01\text{V}^{-1}$, $K = \mu C'_{ox} \frac{W}{L}$, wobei für die Beweglichkeit $250\text{cm}^2/\text{Vs}$ beim nMOS und $100\text{cm}^2/\text{Vs}$ beim pMOS verwendet werden und C'_{ox} die spezifische Oxidkapazität ist.

AC-Simulation

Die Ergebnisse der AC-Simulation stimmen bis in den MHz-Bereich mit den BSIM Modellen überein, so daß keine genauere Modellierung notwendig ist.

1/f Rauschen

Das 1/f-Rauschen im Drainstrom des Transistors wird mit der empirischen Gleichung

$$S = M \cdot \frac{I_D^{1.25}}{C_{ox} f}$$

modelliert, wobei $M = 2 \cdot 10^{-4} \text{FA}^{0.75}$ verwendet wird.

Durchbruchmodell für den weichen Durchbruch

Die Änderung der Kennlinien wird durch eine Verdopplung der Schwellspannung und eine Abnahme der Beweglichkeit um 20% gegenüber dem ungeschädigten Fall berücksichtigt. Der weiche Durchbruchstrom teilt sich jeweils zur Hälfte auf Drain und Source auf, dies gilt für den DC-Strom und das Rauschen.

DC-Gatestrom

Der Gatestrom wird nach dem vereinfachten Quantenpunktmodell (ohne Einfluß der Bandverbiegung) berechnet:

$$I_{BD} = \frac{2q}{h} \frac{\pi q V_t \cdot \exp[-\alpha(\phi + \gamma T)]}{\sin(\pi \alpha q V_t)} \cdot \left\{ \exp \left[\frac{\alpha}{2} \cdot q \frac{V_{GS} + V_{GD}}{2} \right] - \exp \left[-\frac{\alpha}{2} \cdot q \frac{V_{GS} + V_{GD}}{2} \right] \right\}$$

Hierbei ist $\phi = 3.6\text{eV}$, $\alpha = 4\text{eV}^{-1}$, $\gamma = 3\text{meVK}^{-1}$.

1/f² Rauschen

Das aus den Random Telegraph Signals entstehende Spektrum im Gatestrom wird mit

$$S = \frac{4\Delta I_{BD}^2}{(\overline{\tau_{oben}} + \overline{\tau_{unten}}) f^2}$$

modelliert, wobei

$$\Delta I_{BD} = 1\% \cdot I_{BD}$$

$$\overline{\tau_{oben}} = 1\text{s} \cdot 10^{-(V_{GD} + V_{GS})/V}$$

$$\overline{\tau_{unten}} = 10^{-4}\text{s} \cdot 10^{(V_{GD} + V_{GS})/2V}$$

angesetzt wird.

Durchbruchmodell für die DC-Kennlinien nach hartem Durchbruch zum Drain-/Source-Überlappbereich

Der Durchbruch wird durch einen Widerstand $R_{DB} = 10\text{k}\Omega$ zwischen Gate und Drain oder Source modelliert. Auch hier wird die Änderung der Kennlinien durch eine Verdopplung der Schwellspannung und eine Abnahme der Beweglichkeit um 20% berücksichtigt.

Literaturverzeichnis

- [1] BRUGLER, J. S.; JESPERS, P. G. A.: Charge pumping in MOS devices. In: *IEEE Transactions on Electron Devices* **16** (1969), Nr. 3, S. 297–302
- [2] ELLIOT, A. B.: The use of charge pumping currents to measure surface state densities in MOS transistors. In: *Solid-State Electronics* **19** (1976), S. 241–7
- [3] GROESENEKEN, G.; MAES, H.; BELTRÁN, N.; DE KEERSMAECKER, R. F.: A reliable approach to charge-pumping measurements in MOS transistors. In: *IEEE Transactions on Electron Devices* **31** (1984), Nr. 1, S. 42–53
- [4] HADDARA, H.: Drain current charge pumping: A new technique for interface trap characterization in VLSI MOS transistors. In Tagungsband: *16th National Radio Science Conference, 1999*. – Cairo, Egypt, S. D5/1–7
- [5] FIKRY, W.; RAGHEB, M.; HADDARA, H.: Adaptation of "Drain current charge pumping technique" for interface trap characterization in short channel MOS transistors. In Tagungsband: *30th European Solid-State Device Research Conference, 2000*, S. 316–9
- [6] KADEN, G.; REIMER, H.: Ladungsträgereinfang in MOS-Strukturen mit externer Minoritätsträgerquelle. In: *physica status solidi (a)* **32** (1975), S. 183–94
- [7] PRZYREMBEL, G.: *Das Ladungspumpen als Meßmethode zur Charakterisierung von Kurzkanal-MOSFETs*, Technische Universität Berlin, Dissertation, 1989
- [8] Institute for Microelectronics, Technical University Vienna: *MINIMOS 6.1 User's guide*. 1999
- [9] BARBOTTIN, G. (Hrsg.); VAPAILLE, A. (Hrsg.): *Instabilities In Silicon Devices*. Bd. 2. Elsevier Science Publishers B.V., 1989
- [10] SHI, Y.; MA, T. P.; PRASAD, S.; DHANDA, S.: Polarity dependent gate tunneling currents in dual-gate CMOSFETs. In: *IEEE Transactions on Electron Devices* **45** (1998), Nr. 11, S. 2355–60

-
- [11] CAO, K. M.; LEE, W.-C.; LIU, W.; JIN, X.; SU, P.; FUNG, S. K. H.; AN, J. X.; YU, B.; HU, C.: BSIM4 Gate leakage model including Source-Drain partition. In Tagungsband: *International Electron Devices Meeting*, 2000, S. 815–8
- [12] VAN LANGEVELDE, R.; SCHOLTEN, A. J.; DUFFY, R.; CUBAYNES, F. N.; KNITEL, M. J.; KLAASSEN, D. B. M.: Gate current: Modeling, ΔL extraction and impact on RF performance. In Tagungsband: *International Electron Devices Meeting*, 2001, S. 13.2.1–4
- [13] SHIH, W.-K.; RIOS, R.; PACKAN, P.; MISTRY, K.; ABBOTT, T.: A general partition scheme for gate leakage current suitable for MOSFET compact models. In Tagungsband: *International Electron Devices Meeting*, 2001, S. 13.3.1–4
- [14] CAI, J.; SAH, C.-T.: Gate tunneling currents in ultrathin oxide metal-oxide-silicon transistors. In: *Journal of Applied Physics* **89** (2001), Nr. 4, S. 2272–85
- [15] PREIKSZAT, D.: *Charakterisierungsmethoden für MOS-Strukturen mit IBIOS-Schichten und deren Vergleich mit dünnsten thermischen Gate-Oxiden*, Technische Universität Hamburg-Harburg, Diplomarbeit, 1994
- [16] POMPL, T.: *Gateisolatoren für MOS-Feldeffekttransistoren*, Universität der Bundeswehr München, Dissertation, 2000
- [17] SELL, B.; AVELLÁN, A.; KRAUTSCHNEIDER, W. H.: Charge-Based Capacitance Measurements (CBCM) on MOS Devices. In: *IEEE Transactions on Device and Materials Reliability* **2** (2002), Nr. 1, S. 9–12
- [18] SELL, B.: *Interface characterization of metal-gate MOS-Structures and the application to DRAM-Capacitors*, Technische Universität Hamburg-Harburg, Dissertation, 2002
- [19] AVELLÁN, A.; KRAUTSCHNEIDER, W.: On-Chip-Stromverstärker zur Messung von Signalen im Zeitbereich. In Tagungsband: *Analog 2002*, VDE Verlag, 2002 (GMM-Fachbericht 38), S. 145–9
- [20] TOUMAZOU, C.; PAYNE, A.; LIDGEY, J.: Current-feedback versus voltage feedback amplifiers: History, insight and relationships. In Tagungsband: *International Symposium on Circuits and Systems*, 1993, S. 1046–9
- [21] BOWERS, D. F.: The so-called current-feedback operational amplifier - Technological breakthrough or engineering curiosity? In Tagungsband: *International Symposium on Circuits and Systems*, 1993, S. 1054–7

-
- [22] ZELE, R. H.; LEE, S.-S.; ALLSTOT, D. J.: A high gain current-mode operational amplifier. In Tagungsband: *International Symposium on Circuits and Systems*, 1992, S. 2852–5
- [23] MUCHA, I.: Current operational amplifiers: Basic architecture, properties, exploitation and future. In: *Analog Integrated Circuits and Signal Processing* **7** (1995), Nr. 3, S. 243–55
- [24] KARKI, J.: *Voltage Feedback vs Current Feedback Op Amps*. Texas Instruments, November 1998. – Application Report SLVA051
- [25] TUINHOUT, H. P.: Impact of parametric mismatch and fluctuations on performance and yield of deep-submicron CMOS technologies. In Tagungsband: *32nd European Solid-State Device Research Conference*, 2002, S. 95–101
- [26] SUÑÉ, J.; PLACENCIA, I.; BARNIOL, N.; FARRÉS, E.; AYMERICH, X.: Degradation and breakdown of gate oxides in VLSI devices. In: *physica status solidi (a)* **111** (1989), S. 675–85
- [27] STATHIS, J. H.: Physical and predictive models of ultrathin oxide reliability in CMOS devices and circuits. In: *IEEE Transactions on Device and Materials Reliability* **1** (2001), Nr. 1, S. 43–59
- [28] MIZUBAYASHI, W.; YOSHIDA, Y.; MIYAZAKI, S.; HIROSE, M.: Statistical analysis of soft breakdown in ultrathin gate oxides. In Tagungsband: *Symposium on VLSI Technology*, 2001, S. 95–6
- [29] WU, E. Y.; VOLLERTSEN, R.-P.: On the Weibull shape factor of intrinsic breakdown of dielectric films and its accurate experimental determination-Part I: Theory, methodology, experimental techniques. In: *IEEE Transactions on Electron Devices* **49** (2002), Nr. 12, S. 2131–40
- [30] WU, E. Y.; SUÑÉ, J.; LAI, W.: On the Weibull shape factor of intrinsic breakdown of dielectric films an its accurate experimental determination-Part II: Experimental results and the effects of stress conditions. In: *IEEE Transactions on Electron Devices* **49** (2002), Nr. 12, S. 2141–50
- [31] MONSIEUR, F.; VINCENT, E.; ROY, D.; BRUYÈRE, S.; VILDEUIL, J. C.; PANANAKAKIS, G.; GHIBAUDO, G.: A thorough investigation of progressive breakdown in ultra-thin oxides. Physical understanding and application for industrial reliability assessment. In Tagungsband: *40th International Reliability Physics Symposium*, 2002, S. 45–54

-
- [32] LINDER, B. P.; STATHIS, J. H.; FRANK, D. J.; LOMBARDO, S.; VAYSHENKER, A.: Growth and scaling of oxide conduction after breakdown. In Tagungsband: *41st International Reliability Physics Symposium*, 2003, S. 402–5
- [33] AVELLÁN, A.; MIRANDA, E.; SELL, B.; SCHROEDER, D.; KRAUTSCHNEIDER, W.: Temperature dependence of the hard breakdown current of MOS capacitors. In Tagungsband: *32nd European Solid-State Device Research Conference*, 2002, S. 463–6
- [34] SILVACO International: *ATLAS User's Manual*. 1998
- [35] SCHROEDER, D.: *Modelling of interface carrier transport for device simulation*. Springer-Verlag, 1994
- [36] NICOLLIAN, E. H.; BREWS, J. R.: *MOS (Metal Oxide Semiconductor) Physics and Technology*. John Wiley & Sons, 1982
- [37] BEARDA, T.; WOERLEE, P. H.; WALLINGA, H.; HEYNS, M. M.: Charge transport after hard breakdown in gate oxides. In: *Japanese Journal of Applied Physics, Part 1* **41** (2002), Nr. 4B, S. 2431–6
- [38] SZE, S. M.: *Physics of Semiconductors Devices*. 2. Aufl. John Wiley & Sons, 1981
- [39] ARORA, N. D.; HAUSER, J. R.; ROULSTON, D. J.: Electron and hole mobilities in silicon as a function of concentration and temperature. In: *IEEE Transactions on Electron Devices* **29** (1982), Nr. 2, S. 292–5
- [40] NEUDECK, G. W.: *The PN junction diode*. Addison-Wesley Publishing Company, 1983 (Modular series on solid state devices 2)
- [41] WERNER, J. H.: Schottky barrier and pn-junction I/V plots - small signal evaluation. In: *Applied Physics A* **47** (1988), S. 291–300
- [42] GUHA, S.; GUSEV, E. P.; OKORN-SCHMIDT, H.; COPEL, M.; RAGNARSSON, L.-Å.; BOJARCZUK, N. A.; RONSHEIM, P.: High temperature stability of Al₂O₃ dielectrics on Si: Interfacial metal diffusion and mobility degradation. In: *Applied Physics Letters* **81** (2002), Nr. 16, S. 2956–8
- [43] SIN, J. K. O.; SALAMA, C. A. T.: Hybrid Schottky injection MOS-gated power transistor. In: *Electronics Letters* **22** (1986), Nr. 19, S. 1003–5
- [44] KIMURA, M.; MATSUDATE, T.: A new type of Schottky tunnel transistor. In: *IEEE Electron Device Letters* **15** (1994), Nr. 10, S. 412–4

- [45] MEHROTRA, M.; BALIGA, B. J.: Trench MOS barrier Schottky (TMBS) rectifier: A Schottky rectifier with higher than parallel plane breakdown voltage. In: *Solid-State Electronics* **38** (1995), Nr. 4, S. 801–6
- [46] GROVE, A. S.: *Physics and technology of semiconductor devices*. John Wiley & Sons, 1967
- [47] BRIÈRE, O.; CHROBOCZEK, J. A.; GHIBAUDO, G.: Random telegraph signal in the quasi-breakdown current of MOS capacitors. In Tagungsband: *26th European Solid-State Device Research Conference*, 1996, S. 759–62
- [48] CESTER, A.; CIMINO, S.; PACCAGNELLA, A.; GHIDINI, G.; GUEGAN, G.: Collapse of MOSFET drain current after soft breakdown and its dependence on the transistor aspect ratio W/L. In Tagungsband: *41st International Reliability Physics Symposium*, 2003, S. 189–95
- [49] OKADA, K.; TANIGUCHI, K.: Electrical stress-induced variable range hopping conduction in ultrathin silicon dioxides. In: *Applied Physics Letters* **70** (1997), Nr. 3, S. 351–3
- [50] HOUSSA, M.; NIGAM, T.; MERTENS, P. W.; HEYNS, M. M.: Soft breakdown in ultrathin gate oxides: Correlation with the percolation theory of nonlinear conductors. In: *Applied Physics Letters* **73** (1998), Nr. 4, S. 514–6
- [51] ALAM, M. A.; WEIR, B. E.; SILVERMAN, P. J.; MA, Y.; HWANG, D.: The statistical distribution of percolation resistance as a probe into the mechanics of ultra-thin oxide breakdown. In Tagungsband: *International Electron Devices Meeting*, 2000, S. 529–32
- [52] MIRANDA, E.; SUÑÉ, J.: Analytic modeling of leakage current through multiple breakdown paths in SiO₂ films. In Tagungsband: *39th International Reliability Physics Symposium*, 2001, S. 367–79
- [53] MILLER, W. H.: Semiclassical treatment of multiple turning-point problems - phase shifts and Eigenvalues. In: *Journal of Chemical Physics* **48** (1968), Nr. 4, S. 1651–8
- [54] BÜTTIKER, M.: Quantized transmission of a saddle-point constriction. In: *Physical Review B* **41** (1990), Nr. 11, S. 7906–9
- [55] SUÑÉ, J.; MIRANDA, E.: Post soft breakdown conduction in SiO₂ gate oxides. In Tagungsband: *International Electron Devices Meeting*, 2000, S. 533–6
- [56] BRONSTEIN, I. N.; SEMENDJAJEW, K. A.: *Taschenbuch der Mathematik*. 23. Aufl. BSB B. G. Teubner Verlagsgesellschaft, 1987

-
- [57] MIRANDA, E.; SUÑÉ, J.; RODRÍGUEZ, R.; NAFRÍA, M.; AYMERICH, X.; FONSECA, L.; CAMPABADAL, F.: Soft breakdown conduction in ultrathin (3-5nm) gate dielectrics. In: *IEEE Transactions on Electron Devices* **47** (2000), Nr. 1, S. 82-9
- [58] MIRANDA, E.; SUÑÉ, J.; RODRÍGUEZ, R.; NAFRÍA, M.; MARTÍN, F.; AYMERICH, X.: Soft breakdown in ultrathin SiO₂ layers: the conduction problem from a new point of view. In: *Japanese Journal of Applied Physics, Part 1* **38** (1999), Nr. 4B, S. 2223-6
- [59] AVELLÁN, A.; MIRANDA, E.; SCHROEDER, D.; KRAUTSCHNEIDER, W.: Unified model for the voltage and temperature dependence of the soft breakdown conduction mechanism in ultrathin gate oxides. In Tagungsband: *Insulating films on semiconductors*, 2003
- [60] SCHROEDER, D.; AVELLÁN, A.: Physical explanation of the barrier height temperature dependence in metal-oxide-semiconductor leakage current models. In: *Applied Physics Letters* **82** (2003), Nr. 25, S. 4510-2
- [61] AVELLÁN, A.; MIRANDA, E.; SELL, B.; KRAUTSCHNEIDER, W.: Experimental study and modeling of the temperature dependence of soft breakdown conduction in ultrathin gate oxides. In Tagungsband: *41st International Reliability Physics Symposium*, 2003, S. 580-1
- [62] VAN LANGEVELDE, R.; KLAASSEN, F. M.: An explicit surface-potential-based MOSFET model for circuit simulation. In: *Solid-State Electronics* **44** (2000), S. 409-418
- [63] PAVAN, S.; TSIVIDIS, Y.; NAGARAJ, K.: Modeling of accumulation MOS capacitors for analog design in digital VLSI processes. In Tagungsband: *International Symposium on Circuits and Systems*, 1999, S. 202-5
- [64] YEOH, T.-S.; KAMAT, N. R.; NAIR, R. S.; HU, S.-J.: Gate oxide breakdown model in MOS transistors. In Tagungsband: *33rd International Reliability Physics Symposium*, 1995, S. 149-55
- [65] YEOH, T.-S.; HU, S.-J.: Influence of MOS transistor gate oxide breakdown on circuit performance. In Tagungsband: *International Conference on Semiconductor Electronics*, 1998, S. 59-63
- [66] KACZER, B.; DEGRAEVE, R.; GROESENEKEN, G.; RASRAS, M.; KUBICEK, S.; VANDAMME, E.; BADENES, G.: Impact of MOSFET oxide breakdown on digital circuit operation and reliability. In Tagungsband: *International Electron Devices Meeting*, 2000, S. 553-6

- [67] KACZER, B.; DEGRAEVE, R.; DE KEERSGIETER, A.; VAN DE MIEROOP, K.; BEARDA, T.; GROESENEKEN, G.: Consistent model for short-channel nMOSFET post-hard-breakdown characteristics. In Tagungsband: *Symposium on VLSI Technology*, 2001, S. 121–2
- [68] KACZER, B.; DEGRAEVE, R.; DE KEERSGIETER, A.; VAN DE MIEROOP, K.; SIMONS, V.; GROESENEKEN, G.: Consistent model for short-channel nMOSFET after hard gate oxide breakdown. In: *IEEE Transactions on Electron Devices* **49** (2002), Nr. 3, S. 507–13
- [69] POMPL, T.; WURZER, H.; KERBER, M.; WILKINS, R. C. W.; EISELE, I.: Influence of soft breakdown on NMOSFET device characteristics. In Tagungsband: *37th International Reliability Physics Symposium*, 1999, S. 82–7
- [70] TSAI, M.-Y.; LIN, H.-C.; LEE, D.-Y.; HUANG, T.-Y.: Post-soft-breakdown characteristics of deep submicron NMOSFETs with ultrathin gate oxides. In: *IEEE Electron Device Letters* **22** (2001), Nr. 7, S. 348–50
- [71] WU, E.; NOWAK, E.; AITKEN, J.; ABADDER, W.; HAN, L. K.; LO, S.: Structural dependence of dielectric breakdown in ultra-thin gate oxides and its relationship to soft breakdown modes and device failure. In Tagungsband: *International Electron Devices Meeting*, 1998, S. 187–90
- [72] ALERS, G. B.; WEIR, B. E.; KRISCH, K. S.; MONROE, D.: Tunneling current noise and reliability of thin MOS oxides. In Tagungsband: *14th International Conference on Noise in Physical Systems and 1/f Fluctuations*, 1997, S. 247–52
- [73] WEIR, B. E.; SILVERMAN, P. J.; MONROE, D.; KRISCH, K. S.; ALAM, M. A.; ALERS, G. B.; SORSCH, T. W.; TIMP, G. L.; BAUMANN, F.; LIU, C. T.; MA, Y.; HWANG, D.: Ultra-thin gate dielectrics: They break down, but do they fail? In Tagungsband: *International Electron Devices Meeting*, 1997, S. 73–6
- [74] BITTEL, H.; STORM, L.: *Rauschen*. Springer-Verlag, 1971
- [75] MÜLLER, R.: *Rauschen*. Springer-Verlag, 1979
- [76] DUTTA, P.; HORN, P. M.: Low-frequency fluctuations in solids: 1/f noise. In: *Reviews of Modern Physics* **53** (1981), Nr. 3, S. 497–516
- [77] WEISSMAN, M. B.: 1/f noise and other slow, nonexponential kinetics in condensed matter. In: *Reviews of Modern Physics* **60** (1988), Nr. 2, S. 537–71
- [78] VANDAMME, L. K. J.; TREFÁN, G.: 1/f Noise in MOSFETs. In Tagungsband: *Semiconductor Advances for Future Electronics*, 2001, S. 195–208

- [79] MCWORTHER, A. L.: 1/f noise and germanium surface properties. In Tagungsband: *Conference on the physics of semiconductor surfaces*, 1957, S. 207–28
- [80] HOOGE, F. N.; VANDAMME, L. K. J.: Lattice scattering causes 1/f noise. In: *Physics Letters A* **66** (1978), Nr. 4, S. 315–6
- [81] HANDEL, P. H.: Quantum approach to 1/f noise. In: *Physical Review A* **22** (1980), Nr. 2, S. 745–57
- [82] MACHLUP, S.: Noise in semiconductors: Spectrum of a two-parameter random signal. In: *Journal of Applied Physics* **25** (1954), Nr. 3, S. 341–3
- [83] UREN, M. J.; DAY, D. J.; KIRTON, M. J.: 1/f and random telegraph noise in silicon metal-oxide-semiconductor field-effect transistors. In: *Applied Physics Letters* **47** (1985), Nr. 11, S. 1195–7
- [84] KIRTON, M. J.; UREN, M. J.: Noise in solid-state microstructures: A new perspective on individual defects, interface states and low-frequency (1/f) noise. In: *Advances in physics* **38** (1989), Nr. 4, S. 367–468
- [85] UREN, M. J.; COBDEN, D. H.: Generation of random telegraph noise by single Si/SiO₂ interfacial defects. In: HELMS, C. R. (Hrsg.); DEAL, B. E. (Hrsg.): *The Physics and Chemistry of SiO₂ and the Si-SiO₂ interface 2*, Plenum Press, New York, 1993, S. 373–82
- [86] SCHULZ, M.: Coulomb energy of traps in semiconductor space-charge regions. In: *Journal of Applied Physics* **74** (1993), Nr. 4, S. 2649–57
- [87] MUELLER, H. H.; SCHULZ, M.: Review: Individual interface traps at the Si-SiO₂ interface. In: *Journal of materials science: Materials in electronics* **6** (1995), S. 65–74
- [88] MUELLER, H. H.; SCHULZ, M.: Conductance modulation of submicrometer metal-oxide-semiconductor field-effect transistors by single-electron trapping. In: *Journal of Applied Physics* **79** (1996), Nr. 8, S. 4178–86
- [89] MUELLER, H. H.; SCHULZ, M.: Random telegraph signal: An atomic probe of the local current in field-effect transistors. In: *Journal of Applied Physics* **83** (1998), Nr. 3, S. 1734–41
- [90] KANDIAH, K.; DEIGHTON, M. O.; WHITING, F. B.: A physical model for random telegraph signal currents in semiconductor devices. In: *Journal of Applied Physics* **66** (1989), Nr. 2, S. 937–48

- [91] SIMOEN, E.; DIERICKX, B.; CLAEYS, C. L.; DECLERCK, G.: Explaining the amplitude of RTS noise in submicrometer MOSFETs. In: *IEEE Transactions on Electron Devices* **39** (1992), Nr. 2, S. 422–9
- [92] SHI, Z.; MIÉVILLE, J.-P.; DUTOIT, M.: Random telegraph signals in deep submicron n-MOSFETs. In: *IEEE Transactions on Electron Devices* **41** (1994), Nr. 7, S. 1161–8
- [93] SCOFIELD, J. H.; BORLAND, N.; FLEETWOOD, D. M.: Temperature-independent switching rates for a random telegraph signal in a silicon metal-oxide-semiconductor field-effect transistor at low temperatures. In: *Applied Physics Letters* **76** (2000), Nr. 22, S. 3248–50
- [94] GODOY, A.; GÁMIZ, F.; PALMA, A.; JIMÉNEZ-TEJADA, J. A.; BANQUERI, J.; LÓPEZ-VILLANUEVA, J. A.: Influence of mobility fluctuations on random telegraph signal amplitude in n-channel metal-oxide-semiconductor field-effect transistors. In: *Journal of Applied Physics* **82** (1997), Nr. 9, S. 4621–8
- [95] ASENOV, A.; BALASUBRAMANIAM, R.; BROWN, A. R.; DAVIES, J. H.; SAINI, S.: Random Telegraph Signal amplitudes in sub 100nm (decanano) MOSFETs: A 3D 'atomistic' simulation study. In Tagungsband: *International Electron Devices Meeting*, 2000, S. 279–82
- [96] LÜKEN, R.: *Oberflächeneinfluß auf das 1/f-Rauschen von Germanium*, Technische Universität Carolo-Wilhelmina zu Braunschweig, Dissertation, 1969
- [97] BREDERLOW, R.: *Niederfrequentes Rauschen in analogen CMOS-Schaltungen*, Technische Universität Berlin, Dissertation, 1999
- [98] HUNG, K. K.; KO, P. K.; HU, C.; CHENG, Y. C.: A unified model for the flicker noise in metal-oxide-semiconductor field-effect transistors. In: *IEEE Transactions on Electron Devices* **37** (1990), Nr. 3, S. 654–65
- [99] ROUX-DIT-BUISSON, O.; GHIBAUDO, G.; BRINI, J.: Model for drain current RTS amplitude in small-area MOS transistors. In: *Solid-State Electronics* **35** (1992), Nr. 9, S. 1273–6
- [100] GHIBAUDO, G.; ROUX-DIT-BUISSON, O.: Low frequency fluctuations in scaled down silicon CMOS devices status and trends. In Tagungsband: *24th European Solid-State Device Research Conference*, 1994, S. 693–700
- [101] SCHAUMBURG, H.: *Halbleiter*. B. G. Teubner Stuttgart, 1991 (Werkstoffe und Bauelemente der Elektrotechnik 2)

-
- [102] PIERRET, R. F.: *Advanced Semiconductor Fundamentals*. Addison-Wesley Publishing Company, 1987 (Modular series on solid state devices 6)
- [103] SHOCKLEY, W.; READ JR., W. T.: Statistics of the recombination of holes and electrons. In: *Physical Review* **87** (1952), Nr. 5, S. 835–42
- [104] ROGERS, C. T.; BUHRMAN, R. A.: Composition of 1/f noise in metal-insulator-metal tunnel junctions. In: *Physical Review Letters* **53** (1984), Nr. 13, S. 1272–5
- [105] ROGERS, C. T.; BUHRMAN, R. A.: Nature of single-localized-electron states derived from tunneling measurements. In: *Physical Review Letters* **55** (1985), Nr. 8, S. 859–62
- [106] FARMER, K. R.; SALETTI, R.; BUHRMAN, R. A.: Current fluctuations and silicon oxide wear-out in metal-oxide-semiconductor tunnel diodes. In: *Applied Physics Letters* **52** (1988), Nr. 20, S. 1749–51
- [107] FARMER, K. R.; BUHRMAN, R. A.: Defect dynamics and wear-out in thin silicon oxides. In: *Semiconductor Science and Technology* **4** (1989), S. 1084–105
- [108] FARMER, K. R.; ROGERS, C. T.; BUHRMAN, R. A.: Localized-state interactions in metal-oxide-semiconductor tunnel diodes. In: *Physical Review Letters* **58** (1987), Nr. 21, S. 2255–8
- [109] ANDERSSON, M. O.; XIAO, Z.; NORRMAN, S.; ENGSTRÖM, O.: Model based on trap-assisted tunneling for two-level current fluctuations in submicrometer metal-silicon-dioxide-silicon diodes. In: *Physical Review B* **41** (1990), Nr. 14, S. 9836–42
- [110] HUNG, K. K.; KO, P. K.; HU, C.; CHENG, Y. C.: An automated system for measurement of random telegraph signals in metal-oxide-semiconductor field-effect transistors. In: *IEEE Transactions on Electron Devices* **36** (1989), Nr. 6, S. 1217–9
- [111] MISITI, M.; MISITI, Y.; OPPENHEIM, G.; POGGI, J.-M.: *Wavelet Toolbox User's Guide*. The MathWorks Inc.
- [112] NEWLAND, D. E.: *An introduction to random vibrations, spectral and wavelet analysis*. 3. Aufl. Longman Scientific & Technical, 1993
- [113] AVELLÁN, A.; KRAUTSCHNEIDER, W.; SELL, B.: Softbreakdown and RTS in gate and drain currents of MOS transistors with ultrathin oxides. In Tagungsband: *16th International Conference on Noise in Physical Systems and 1/f Fluctuations*, 2001, S. 375–8

- [114] LÜ, J.-Q.; KOCH, F.: Random Telegraph Noise in advanced self-aligned bipolar transistors. In: *Japanese Journal of Applied Physics, Part 1* **35** (1996), Nr. 2B, S. 826–32
- [115] VON HAARTMAN, M.; SANDÉN, M.; ÖSTLING, M.: Random telegraph signal noise in SiGe heterojunction bipolar transistors. In: *Journal of Applied Physics* **92** (2002), Nr. 8, S. 4414–21
- [116] CHEN, M.-J.; LU, M.-P.: On-off switching of edge direct tunneling currents in metal-oxide-semiconductor field-effect transistors. In: *Applied Physics Letters* **81** (2002), Nr. 18, S. 3488–90
- [117] HENRY, C. H.; LANG, D. V.: Nonradiative capture and recombination by multiphonon emission in GaAs and GaP. In: *Physical Review B* **15** (1977), Nr. 2, S. 989–1016
- [118] ISLER, M.; LIEBIG, D.: Enhanced multiphonon capture of hot electrons by deep centers with strong lattice coupling: A Monte Carlo study of InP:Fe. In: *Physical Review B* **61** (2000), Nr. 11, S. 7483–8
- [119] ISLER, M.: *Hochfeld-Transporteffekte in Heterostruktur-Feldeffekttransistoren und ihr Einfluss auf das Hochfrequenzverhalten*, Technische Universität Hamburg-Harburg, Dissertation, 2003
- [120] AVELLÁN, A.; SCHROEDER, D.; KRAUTSCHNEIDER, W.: Modeling random telegraph signals in the gate current of metal oxide semiconductor field effect transistors after oxide breakdown. In: *Journal of Applied Physics* **94** (2003), Nr. 1, S. 703–8
- [121] BLAKEMORE, J. S.: Approximations for Fermi-Dirac integrals, especially the function $F_{1/2}(\eta)$ used to describe electron density in a semiconductor. In: *Solid-State Electronics* **25** (1982), Nr. 11, S. 1067–76
- [122] Cadence Design Systems, Inc.: *Spectre HDL Reference, Product Version 4.4.3*. 1998
- [123] NOUMAINE, R.: *Entwicklung eines SPICE-Modells zur Beschreibung des MOS-Transistorverhaltens nach Oxiddurchbruch*, Technische Universität Hamburg-Harburg, Diplomarbeit, 2003
- [124] ARORA, N.: *MOSFET models for VLSI circuit simulation*. Springer-Verlag, 1993
- [125] PELGROM, M. J. M.; DUINMAIJER, A. C. J.; WELBERS, A. P. G.: Matching properties of MOS transistors. In: *IEEE Journal of Solid-State Circuits* **24** (1989), Nr. 5, S. 1433–40

-
- [126] SU, H.; MICHAEL, C.; ISMAIL, M.: Yield optimization of analog MOS integrated circuits including transistor mismatch. In Tagungsband: *International Symposium on Circuits and Systems*, 1993, S. 1801–4
- [127] MICHAEL, C.; WANG, H.; TENG, C. S.; SHIBLEY, J.; LEWICKI, L.; SHYU, C.-M.; LAHRI, R.: Mismatch drift: A reliability issue for analog MOS circuits. In Tagungsband: *30th International Reliability Physics Symposium*, 1992, S. 81–4
- [128] CHEN, Y.; ZHOU, J.; TEDJA, S.; HUI, F.; OATES, A. S.: Stress-induced MOS-FET mismatch for analog circuits. In Tagungsband: *IEEE International Integrated Reliability Workshop*, 2001, S. 41–3
- [129] LANZONI, M.; TONDI, G.; GALBIATI, P.; RICCÓ, B.: Automatic and continuous offset compensation of MOS operational amplifiers using floating gate transistors. In: *IEEE Journal of Solid-State Circuits* **33** (1998), Nr. 2, S. 287–90
- [130] TIETZE, U.; SCHENK, C.: *Halbleiter-Schaltungstechnik*. 10. Aufl. Springer-Verlag, 1993
- [131] PELGROM, M. J. M.; TUINHOUT, H. P.; VERTREGT, M.: Transistor matching in analog CMOS applications. In Tagungsband: *International Electron Devices Meeting*, 1998, S. 915–8