

**Fakultät für Elektrotechnik und Informationstechnik
der Technischen Universität München
Lehrstuhl für Technische Elektrophysik**

Physikalische Modellierung und numerische Simulation von Nanodraht- Tunnelfeldeffekttransistoren

Alexander Maximilian Heigl

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines
Doktors-Ingenieur (Dr.-Ing.)
genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr.-Ing. H.-G. Herzog

Prüfer der Dissertation: 1. Univ.-Prof. Dr. rer. nat. G. Wachutka
2. Univ.-Prof. Dr.-Ing. W. Hansch
Universität der Bundeswehr München

Die Dissertation wurde am 29.09.2011 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 05.02.2015 angenommen.

Zusammenfassung

Die fortschreitende Miniaturisierung in der Halbleitertechnologie führt zu Problemen bei der Skalierung von MOSFETs (*metal oxide semiconductor field-effect transistor*) durch sogenannte Kurzkanaleffekte. So nehmen beispielsweise die unerwünschten Leckströme durch das Gateoxid zu und die Schalteigenschaften werden schlechter. Um eine solche Abnahme der Leistungsfähigkeit zu kompensieren, ist ein erheblicher Designaufwand zu betreiben. Hieraus folgen Transistoren mit meist signifikant komplexeren Dotierprofilen und Bauelementgeometrien. Eine andere Möglichkeit stellt die Einführung alternativer Bauelementkonzepte dar. An diese sind allerdings einige Anforderungen zu stellen: Die gegenwärtig eingesetzten Prozessierungstechnologien sollten alleine aus Kostengründen auch weiterhin anwendbar sein. Weitere Anforderungen sind gute Schalteigenschaften und die Fähigkeit komplementäre Logikschaltungen zu realisieren. Eine solche Option ist der Tunnelfeldeffekttransistor (TFET).

Im Rahmen dieser Arbeit werden die Eigenschaften von Nanodraht-Tunneltransistoren in unterschiedlichen Bauformen durch numerische Simulationen analysiert. Im Prinzip handelt es sich bei diesen Bauelementen um die Anwendung des Multigatekonzepts auf Tunnelfeldeffekttransistoren. Hierbei zeigt sich, dass mit den hier eingesetzten Modellen zur Beschreibung des Band-zu-Band Tunnelns erstmals die vollständigen Kennlinien von Tunneln und Tunneln auf realistische Weise simuliert werden können. Den Kern dieser Arbeit bildet die Ausarbeitung von Wegen zur Optimierung der Schalteigenschaften von Nanodraht-Tunneltransistoren. Zu diesem Zweck werden die Auswirkungen von Dotierprofilen, Geometrie und Materialeigenschaften untersucht und daran anschließend die zugrunde liegenden Mechanismen erklärt. Auf diese Weise folgt ein Leitfaden zur Optimierung von n- und p-Kanal Tunnelfeldeffekttransistoren. Hierbei wird auch auf die Besonderheiten von rechteckigen Kanalquerschnitten eingegangen.

Den Abschluss bildet die eingehende Analyse des eingesetzten Modells zur Beschreibung von Band-zu-Band Tunnelprozessen: Es zeigt sich, dass bei der Ableitung der Tunnelraten ein folgenreicher Fehler aufgetreten ist. In dieser Arbeit wird der Fehler behoben und das daraus folgende korrigierte Tunnelmodell wird anschließend in den Bauelementsimulator Socrates Device implementiert und ausgiebig getestet. Zudem werden einige alternative Modellierungskonzepte zur Beschleunigung der Simulationen vorgestellt.

Abstract

Due to the progressive shrinking of the physical feature sizes of modern nanometer-scale MOS-devices (metal oxide semiconductor) the influence of undesired shortchannel effects like punch-through, drain-induced barrier lowering and roll-off is getting more and more important. Partially these can be reduced and in some cases even avoided by introducing elaborate doping profiles or scaling the vertical dimensions of the devices. Nevertheless several challenges are still remaining, in particular the leakage currents caused by parasitic tunneling such as the gate-induced drain leakage or direct tunneling of electrons through the gate oxide. An alternative approach to cope with shortchannel effects relies on new device concepts. One feasible option is the tunneling field-effect transistor (TFET). The primary working principle of the TFET is gate-controlled band-to-band tunneling through a heavily doped p^+n -junction.

Using physical device simulation, the functionality and performance of various nanowire tunneling field-effect transistors is investigated. Basically these devices are the application of multigate concepts on the TFET. Including quantum confinement effects in terms of the density gradient approach and the band-to-band tunneling model of Schenk, it is possible to obtain for the first time the complete characteristics of tunneling diodes and transistors within the framework of a drift-diffusion based device simulator. The focal point of this work is directed on the optimisation of the switching behaviour of nanowire tunneling transistors by the variation of the doping profiles and the geometrical dimensions. Special attention is focused on the possibility to improve the device performance by considering alternative materials for the gate-stack and the source region. Furthermore, some specifics of rectangular channel diameters are discussed. In conclusion a detailed guideline to the optimisation of n- and p-type tunneling transistors is presented.

The final chapter deals with a critical evaluation of the band-to-band tunneling model of Schenk: An approximation error in the original model is leading to an erroneous formula of the band-to-band tunneling rate. In this work the error is identified as wrong series expansion of the Airy-integral and a better suited approximation is introduced. Since the complete model is depending on a sum of different types of Airy-functions and integrals an alternative way to obtain a more precise approximation is presented. Afterwards, the resulting model is implemented in Sentaurus Device and elaborate investigations on the application of the corrected band-to-band tunneling model to the simulation of cylindrical nanowire tunneling transistors are performed.

Inhaltsverzeichnis

1. Einleitung	1
1.1 Nanostrukturierte Halbleiterbauelemente	1
1.2 Klassifizierung von Halbleiterbauelementen	2
1.3 Hierarchien der Bauelementmodellierung	3
1.4 Konzepte zur Modellierung von Transportprozessen	4
2. Modellierung von Band-zu-Band Tunnelbauelementen	5
2.1 Grundlagen der Halbleiterphysik	5
2.1.1 Bandstruktur und effektive Massennäherung	5
2.1.2 Fermi-Dirac-Statistik und Ladungsträgerdichte	7
2.1.3 Kopplung der Ladungsträger: Poisson-Gleichung	8
2.2 Ladungsträgertransport und Boltzmann-Gleichung	8
2.2.1 Ableitung des Drift-Diffusions-Modells	9
2.2.2 Randbedingungen bei der Halbleitermodellierung	10
2.2.3 Scharfetter-Gummel-Diskretisierung der Transportgleichungen	12
2.3 Generationsmodelle	13
2.3.1 Shockley-Read-Hall-Rekombination	13
2.3.2 Defektunterstütztes Tunneln	14
2.3.3 Stoßionisation	15
2.4 Phononenunterstütztes Tunneln	16
2.4.1 Direktes Band-zu-Band Tunneln	17
2.4.2 Indirektes Band-zu-Band Tunneln	18
2.5 Auswirkungen von Quanteneffekten auf die Ladungsdichten	23

3. Numerische Simulation von Tunnelbauelementen	27
3.1 Untersuchung elementarer Tunnelbauelemente	27
3.1.1 Esaki-Tunnelodiode	27
3.1.2 MOSFET	29
3.1.3 Tunneltransistor	33
3.1.4 Zener-Feldeffekttransistor	37
3.2 Modellkalibrierung und Parametervariation	38
3.2.1 Parameter des Band-zu-Band Tunnelmodells	39
3.2.2 Kritische Betrachtung der Stoßionisation in Tunneltransistoren	41
4. Optimierung von Tunnelbauelementen	45
4.1 Einfluss von Quanteneffekten auf die TFET-Charakteristiken	45
4.2 Auswirkungen der Dotierprofile auf das Bauelementverhalten	49
4.2.1 Drain- und Sourcekonzentration	50
4.2.2 Kanalkonzentration	53
4.2.3 Variation der Spacerweite	55
4.2.4 Gradient der Sourcedotierung	58
4.3 Skalierungseigenschaften bei zylindrischen Nanodraht-TFETs	61
4.3.1 Gatelänge	61
4.3.2 Durchmesser	65
4.4 Auswirkungen unterschiedlicher Materialien	69
4.4.1 Variation des Gate-Stapels (Material und Dicke)	70
4.4.2 Einfluss verschiedener Materialien im Kanalgebiet	76
4.5 Optimierungsstrategien bei zylindrischen Nanodraht-TFETs	79
4.6 Analyse von 3D-strukturierten Tunneltransistoren	81
4.6.1 Doppelgate als Vorstufe zu Nanodraht-TFETs	82
4.6.2 Eckeneffekte bei Quadgate-Nanodraht-TFETs	83
4.6.3 Vergleich unterschiedlicher Querschnittstypen	87
5. Verbesserte Modellierung von Band-zu-Band Tunnelprozessen	89
5.1 Korrektur der Band-zu-Band Tunnelmodelle	89

5.1.1	Fehleranalyse beim bislang verwendeten Tunnelmodell	90
5.1.2	Näherungsmethoden für Airy-Integrale	92
5.1.3	Ableitung eines korrigierten Band-zu-Band Tunnelmodells	94
5.2	Alternative Modellkonzepte für das Band-zu-Band Tunneln	99
5.2.1	Vereinfachte Beschreibung des Band-zu-Band Tunnelns	99
5.2.2	Simulation von Tunneltransistoren auf Kompaktmodellebene	103
6.	Zusammenfassung und Ausblick	107
	Symbolverzeichnis	111
	Naturkonstanten und Modellparameter	117
	Literaturverzeichnis	121
	Danksagung	131

1. Einleitung

Halbleiterbauelemente haben heute Einzug in nahezu alle Bereiche des täglichen Lebens gehalten. Einerseits sind es die bekannten Anwendungen als Transistoren in Prozessoren oder Speicherelementen in Digitalkameras und Computern und andererseits gibt es auch eine Vielfalt an alternativen Einsatzgebieten: von der Sensorik im Automobilbau über die Beleuchtungsindustrie bis hin zur Energieversorgung. Diese Halbleiterbauelemente haben alle gemeinsam, dass einerseits grundlegende Materialeigenschaften identisch sind, und andererseits der Bedarf besteht diese exakt und schnell virtuell nachzubilden. Dazu ist es erforderlich geeignete Modelle zur Beschreibung zu entwickeln, die möglichst nah an der physikalischen Wirklichkeit sind und gleichzeitig die Problemstellungen signifikant vereinfachen. Der Grund für dieses Ziel hat darin seine Ursache, dass die Evaluierung und Optimierung der Bauelementeigenschaften in geringer Zeit und damit auch zu geringen Kosten durchführbar sein soll.

Im Rahmen dieser Arbeit werden verschiedene zukünftige Transistorkonzepte analysiert und hinsichtlich ihres Einsatzes als Schaltelemente optimiert. Zu diesem Zweck werden diese am Computer nachgebildet und mit Hilfe von geeigneten Beschreibungsmodellen simuliert. Von zentraler Bedeutung ist hierbei auch die Entwicklung, Erweiterung und Verifikation von Modellen für die Simulation dieser nanostrukturierten Bauelemente. Im ersten Abschnitt werden die untersuchten Bauelemente hinsichtlich ihrer Funktionsweise und Anwendungsgebiete eingeordnet. Anschließend werden verschiedene Konzepte und Hierarchien zur Modellierung von Halbleiterbauelementen vorgestellt. Von besonderer Bedeutung ist dabei die Beschreibung von Ladungsträgertransportprozessen.

1.1 Nanostrukturierte Halbleiterbauelemente

Mit dem Begriff nanostrukturierte Halbleiterbauelemente werden meist die Bauelemente bezeichnet, deren physikalisch relevante Abmessungen im Nanometerbereich liegen. Da in jedem MOSFET (*metal oxide semiconductor field-effect transistor*) das Gateoxid nur wenige Nanometer dünn ist, fällt nach dieser Definition diese gesamte Transistorklasse unter die nanostrukturierten Bauelemente. In der nachfolgenden Untersuchung werden Feldeffekttransistoren allerdings nur dann als Nanometerbauelemente bezeichnet, wenn auch die Gatelänge deutlich unter einem Mikrometer liegt. Als Folge der fortschreitenden Miniaturisierung von Transistoren werden inzwischen auch die Kanalbreiten im unteren Nanometerbereich realisiert. Da in diesen Fällen die Gatemetallisierungen zusätzlich an den Seiten des Transportkanals angebracht werden, entstehen sogenannte Multigate-FETs (MuGFETs). Ein solches Bauelement ist in Abbildung 1.1 (links) dargestellt. Im oberen

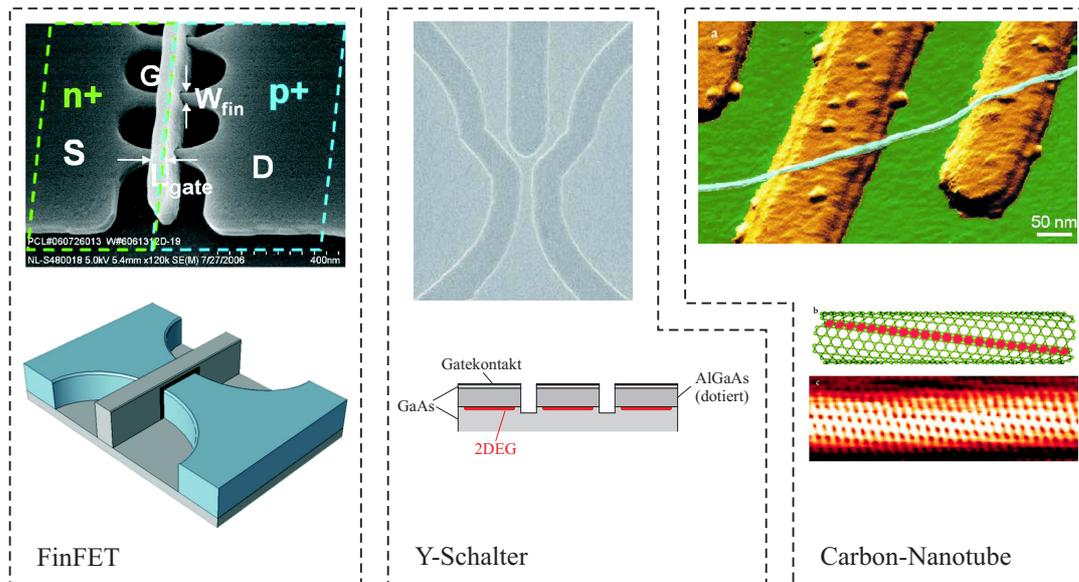


Abb. 1.1: Auswahl an unterschiedlichen nanostrukturierten Halbleiterbauelementen: FinFET [1], Y-Schalter [2] und Carbon-Nanotube [3].

Teil ist eine SEM-Aufnahme (*scanning electron microscope*) gezeigt. Auf diesem können mehrere parallele Kanäle identifiziert werden, die von einer dünnen Oxidschicht bedeckt und auf drei Seiten von einem durchlaufenden Gate ummantelt sind. Das untere Teilbild gibt zusätzlich den schematischen Aufbau des Bauelements wieder. Eine solche komplexe Strukturierung ist erforderlich, da auf diese Weise auch bei sehr kleinen Kanallängen eine gute Steuerbarkeit durch den Gatekontakt gewährleistet werden kann. In der folgenden Arbeit werden solche Multigatekonzepte auf den Tunneltransistor [4, 5] angewandt und dann mit Hilfe eines geeigneten Beschreibungsmodells [6] simuliert. Eine zweite Gruppe nanostrukturierter Transistoren sind quantenballistische Bauelemente. Bei diesen ist die mittlere freie Weglänge der Elektronen zwischen zwei Stoßprozessen deutlich größer als die Bauelementabmessungen. Erreicht werden kann diese Situation, indem geschichtete Halbleiterheterostrukturen erzeugt werden. In diesen bildet sich bei geeignetem Design ein zweidimensionales Elektronengas unterhalb der hochdotierten Schicht, wodurch eine hohe Beweglichkeit der Elektronen erzielt wird. Ein Beispiel für ein solches Bauelement ist der in Abbildung 1.1 (Mitte) dargestellte Y-Schalter. Der Vorteil dieser Strukturen liegt darin, dass nicht nur Transistoren, sondern Zähler und andere Logikbauelemente direkt realisiert werden können. Auf der rechten Seite dieser Abbildung ist ein drittes Beispiel nanostrukturierter Transistoren abgebildet. Bei diesem ist der leitende Kanal durch ein Kohlenstoff-Nanoröhrchen (*carbon nanotubes*) realisiert.

1.2 Klassifizierung von Halbleiterbauelementen

Für die Einteilung von Halbleiterbauelementen steht eine Vielzahl an Möglichkeiten zur Verfügung. So kann eine Klassifizierung beispielsweise nach dem Anwendungsgebiet

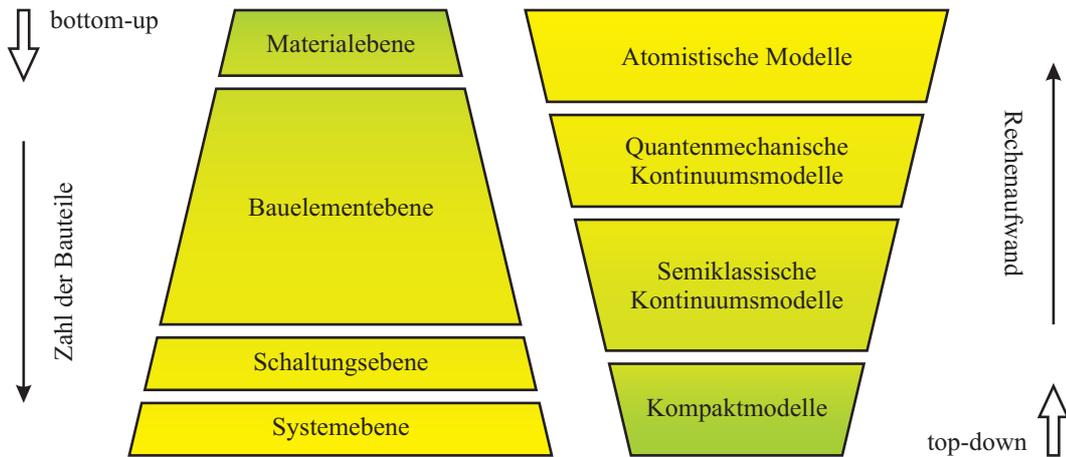


Abb. 1.2: Hierarchien der Modellierung in der Halbleiterphysik.

vorgenommen werden. In der Sensorik und Aktorik werden Halbleiterbauelemente als Beschleunigungssensoren und Mikropumpen eingesetzt. Alternativ können diese auch Transistoren in Logikschaltungen oder IGBTs (*insulated gate bipolar transistors*) in der Automobilelektronik sein. Weitere Kriterien sind die Abmessungen der Bauelemente oder die beteiligten Energiedomänen: Dabei können elektrische, optische, mechanische und fluidische Effekte gekoppelt auftreten. In der vorliegenden Arbeit werden elektronische nanostrukturierte Bauelemente untersucht. Diese wiederum können nach dem beteiligten Transportmechanismus eingeteilt werden. Bei gegebenen Bauelementabmessungen L_{dev} und der freien Weglänge der Elektronen λ_{free} gibt es:

- dissipativen Transport ($L_{\text{dev}} \gg \lambda_{\text{free}}$)
- ballistischen Transport ($L_{\text{dev}} \approx \lambda_{\text{free}}$)
- quantenballistischen Transport ($L_{\text{dev}} \ll \lambda_{\text{free}}$)

Einen entscheidenden Einfluss hat dabei die Dotierung des Halbleiters, da die Störstellen die freie Weglänge signifikant beeinflussen. So haben beispielsweise in einem mit einer Konzentration von 10^{15} cm^{-3} dotierten Halbleiter die Störstellen einen mittleren Abstand von 100 nm je Raumrichtung.

1.3 Hierarchien der Bauelementemodellierung

Zumeist werden nicht einzelne Halbleiterbauelemente eingesetzt, sondern eine Vielzahl solcher. In diesem Fall stellen sich auch unterschiedliche Anforderungen an das virtuelle Nachbilden dieser Bauelemente und Systeme. Eine Zusammenstellung der Hierarchien in der Modellierung von Problemen der Halbleiterphysik ist in Abbildung 1.2 gezeigt. So ist in der linken Hälfte des Bildes nach der Anzahl der beteiligten Bauelemente und auf der rechten Seite nach dem Modelltyp unterschieden. Da prinzipiell mehr Bauelemente auch

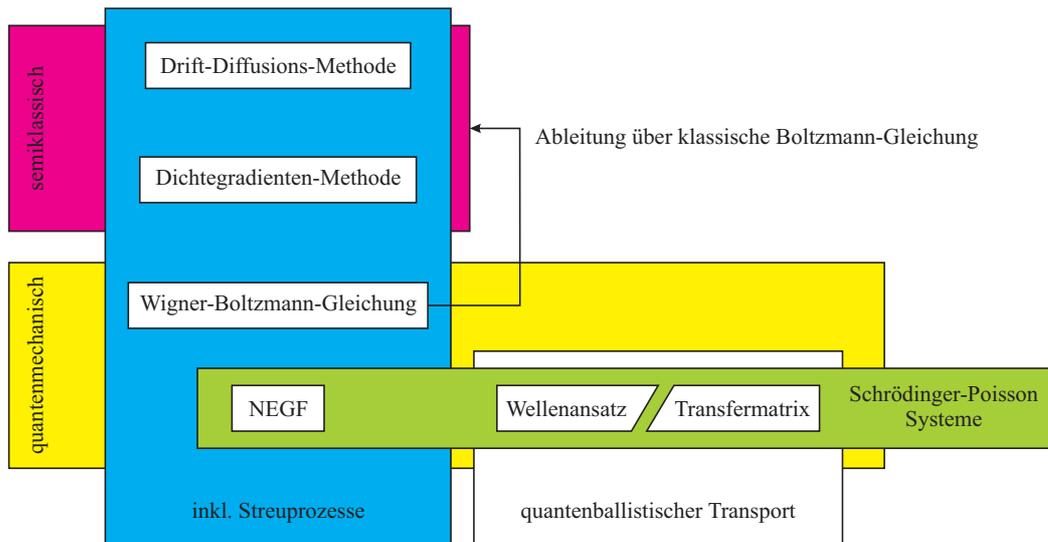


Abb. 1.3: Konzepte zur semiklassischen und quantenmechanischen Modellierung von Halbleiterbauelementen.

einen erhöhten Rechenaufwand bedeuten, werden in Systemen die einzelnen Bausteine zumeist mit Hilfe von Kompaktmodellen beschrieben. Je weniger Bauelemente und damit Teilchen beteiligt sind, desto genauere Modelle sind rechentechnisch möglich.

1.4 Konzepte zur Modellierung von Transportprozessen

In der vorliegenden Arbeit werden vornehmlich Kontinuumsmodelle zur Simulation von Transportprozessen herangezogen. Die semiklassischen Modelle basieren hierbei auf der Boltzmann-Gleichung und führen im einfachsten Fall zur Drift-Diffusions-Methode, die bisweilen durch quantenmechanische Korrekturen erweitert wird [7]. Diesen gegenüber stehen die quantenmechanischen Modelle, die entweder auf der Wigner-Gleichung [8, 9] oder auf Schrödinger-Poisson-Systemen beruhen [10]. Werden Aufgabenstellungen ohne kohärenzbrechende Streuprozesse (quantenballistischer Transport) untersucht, so können direkte Verfahren zur Lösung des Schrödinger-Poisson-Systems, wie Wellenansätze oder die Transfermatrixmethode, eingesetzt werden [11, 12, 13]. Sind zusätzlich Streuprozesse für die korrekte Beschreibung des Bauelementverhaltens von Bedeutung, dann kann die NEGF-Methode (Nichtgleichgewichts Green-Funktionen) eingesetzt werden [14, 15]. In Abbildung 1.3 sind die einzelnen Verfahren noch einmal schematisch zusammengestellt und zueinander in Relation gesetzt.

2. Modellierung von Band-zu-Band Tunnelbauelementen

In diesem Kapitel werden die grundlegenden Modelle zur Simulation von Band-zu-Band Tunnelbauelementen vorgestellt. Hierbei wird im ersten Abschnitt, ausgehend von der Bandstruktur, die effektive Massennäherung eingeführt und die elektrostatische Kopplung der Ladungsträger durch die Poisson-Gleichung besprochen. Daran anschließend wird das Ladungstransportproblem ausgehend von der Boltzmann-Gleichung betrachtet und somit das Drift-Diffusions-Modell abgeleitet. Nach einer kurzen Präsentation der physikalisch relevanten Randbedingungen, wird dann auf die numerische Lösung der resultierenden partiellen Differentialgleichungen eingegangen. Im dritten und vierten Abschnitt werden die für die Simulation erforderlichen Generationsmodelle zur paarweisen Erzeugung von Elektronen und Löchern eingeführt. Dabei wird ein Schwerpunkt auf die Beschreibung von Band-zu-Band Tunnelprozessen gelegt und das vorgestellte Tunnelmodell wird in eine für die Bauelementsimulation besser geeignete Form überführt. Auf diese Weise können letztlich Band-zu-Band Tunnelprozesse durch eine Generationsrate beschrieben werden. Den Abschluss dieses Kapitels bildet dann die Beschreibung der Auswirkungen von Quanteneffekten auf die Ladungsdichten und die mögliche Behandlung in Form einer Erweiterung des Drift-Diffusions-Modells durch die Dichtegradientenmethode.

2.1 Grundlagen der Halbleiterphysik

Die physikalisch korrekte Beschreibung von Transportprozessen in Halbleitermaterialien bei komplexen zwei- und dreidimensionalen Strukturen erfordert einige vereinfachende Annahmen, da eine große Klasse an Aufgabenstellungen wegen des damit verbundenen rechentechnischen Aufwands nur selten vollständig quantenmechanisch gelöst werden kann. Daher werden dann sogenannte semiklassische Modelle eingesetzt bei denen der atomare Aufbau der Materialien durch ein Kontinuumsmodell beschrieben wird [16].

2.1.1 Bandstruktur und effektive Massennäherung

Von entscheidender Bedeutung für die Eigenschaften von Halbleiterbauelementen ist die zugrundeliegende Struktur des eingesetzten Halbleitermaterials. An dieser Stelle wird nur auf Silizium eingegangen, da dieses zum gegenwärtigen Zeitpunkt das Standard-CMOS

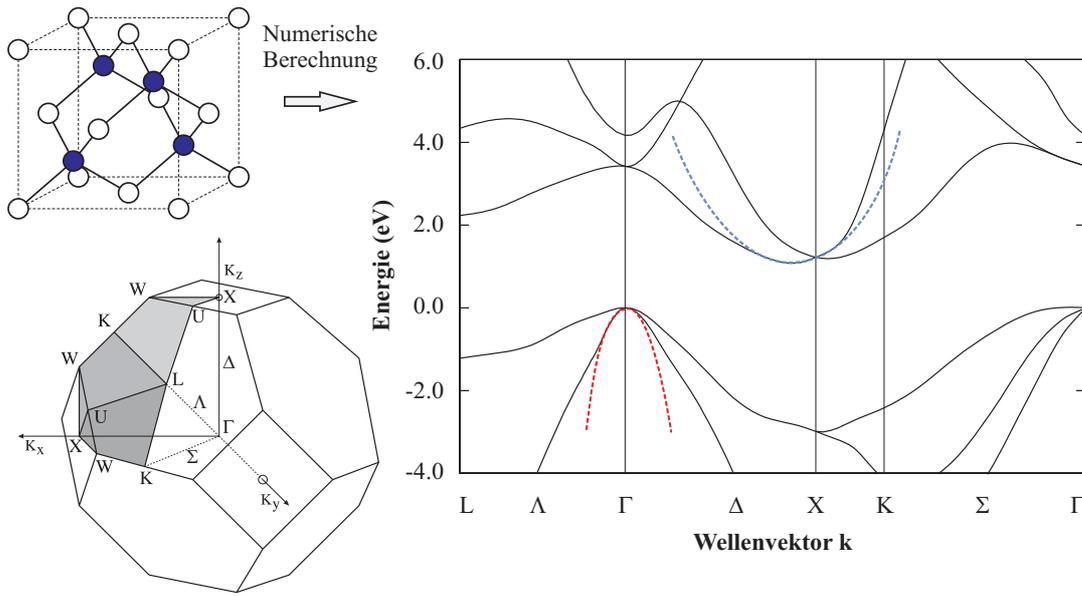


Abb. 2.1: Kristallstruktur, erste Brillouin-Zone des kubisch flächenzentrierten Gitters und Energie-Impuls-Banddiagramm von Silizium [17]. In das Banddiagramm ist zusätzlich die parabolische Effektivmassennäherung eingezeichnet.

Material ist und die hier durchgeführten Untersuchungen von Tunnelbauelementen sich überwiegend mit solchen siliziumbasierten Bauelementen beschäftigen. Die einzelnen Atome sind im Siliziumkristall auf zwei, zueinander um ein Viertel der Diagonalen der Einheitszelle verschobenen, kubisch-flächenzentrierten Gitter angeordnet. In Abbildung 2.1 oben links ist die Einheitszelle von Silizium mit den entsprechenden vier kovalenten Bindungen der einzelnen Atome dargestellt.

Wegen des periodischen Gitters von Halbleitern können die Einteilchenzustände von Elektronen durch einen (Quasi-)Impuls charakterisiert werden. Aus dem in Abbildung 2.1 dargestellten numerisch berechneten Verlauf des Energie-Impuls-Banddiagramms geht hervor, dass Silizium ein indirekter Halbleiter ist. So ist das Leitungsbandminimum (nahe dem X-Punkt) bei einem anderen Wellenvektorwert als das Maximum des Valenzbandes (im Γ -Punkt). Aufgrund der Symmetrie hat kristallines Silizium in der Energie-Impuls-Darstellung sechs äquivalente Leitungsbandminima zwischen dem Γ - und X-Punkt. In der Nähe der Bandkanten können die realen Energiebänder $\varepsilon_\nu(\mathbf{k})$ des ν -ten Bandes im Extremum \mathbf{k}_e durch eine parabolische Funktion approximiert werden

$$\tilde{\varepsilon}_\nu(\mathbf{k}) \approx \varepsilon_\nu(\mathbf{k}_e) + \frac{\hbar^2}{2} \cdot (\mathbf{k} - \mathbf{k}_e) \cdot \left[\frac{1}{m_{\nu, \mathbf{k}_e}^*} \right] \cdot (\mathbf{k} - \mathbf{k}_e) \quad (2.1)$$

wobei der effektive Massentensor definiert ist durch

$$\left[\frac{1}{m_{\nu, \mathbf{k}_e}^*} \right]_{ij} = \frac{1}{\hbar^2} \cdot \left. \frac{\partial^2 \varepsilon_\nu(\mathbf{k})}{\partial k_i \partial k_j} \right|_{\mathbf{k}=\mathbf{k}_e} \quad (2.2)$$

Wegen der unterschiedlichen Krümmungen der Bandkanten im Leitungsband berechnet sich die effektive Masse der Elektronen im Silizium aus einem transversalen und einem

longitudinalen Anteil

$$m_e = 6^{2/3} \cdot (m_t^2 \cdot m_l)^{1/3} \quad (2.3)$$

Auf diese Weise kann ebenso für die Löcher eine effektive Masse bestimmt werden. Im Fall der Löcher ist allerdings zu berücksichtigen, dass sich diese aus den Anteilen drei verschiedener Subbänder zusammensetzt. Mit den effektiven Massen können letztlich die dazugehörigen effektiven Zustandsdichten für das Leitungsband

$$N_c = 2 \cdot \left(\frac{2\pi k m_e T}{h^2} \right)^{3/2} \quad (2.4)$$

und das Valenzband

$$N_v = 2 \cdot \left(\frac{2\pi k m_h T}{h^2} \right)^{3/2} \quad (2.5)$$

berechnet werden. Hierbei ist k die Boltzmann-Konstante, T die Temperatur und h das Plancksche Wirkungsquantum. Die realistische Bandstruktur der Halbleiter wird also in den durchgeführten Bauelementsimulationen nicht berücksichtigt. Durch die vorgestellte Näherung werden parabolische Bänder für Elektronen und Löcher betrachtet. Auf diese Weise erfolgt eine Parametrisierung der Bänder in Abhängigkeit der effektiven Massen der Elektronen m_e und Löcher m_h , der Elektronenaffinität χ und der Bandlücke E_g . Die Beschreibung der Impulsabhängigkeit wird vollständig aufgegeben und darüberhinaus ist bei dieser Vorgehensweise keine Unterscheidung von direkten und indirekten Halbleitern möglich. Bei der Herleitung von Modellen zur Beschreibung von Interband-Übergängen (phononenunterstütztes Tunneln vom Valenz- in das Leitungsband) wird eine erneute Analyse der Bandstruktur erforderlich werden.

2.1.2 Fermi-Dirac-Statistik und Ladungsträgerdichte

Die Ladungsträgerdichte der Elektronen und Löcher in einem Halbleiter wird durch die Integration über die Zustandsdichte und der Besetzung der Zustände berechnet. Sind die Halbleiter stark dotiert, muss die Fermi-Dirac-Statistik für die Besetzung der Zustände eingesetzt werden und es folgt dann für die Ladungsträgerdichten

$$n = N_c \cdot \mathcal{F}_{1/2} \left(\frac{F_n - E_c}{kT} \right) \quad (2.6)$$

$$p = N_v \cdot \mathcal{F}_{1/2} \left(\frac{E_v - F_p}{kT} \right) \quad (2.7)$$

mit dem Fermi-Dirac-Integral der Ordnung $1/2$, den Quasi-Fermi-Niveaus der Elektronen F_n und Löcher F_p und der Leitungs- und Valenzbandenergie

$$E_c = -\chi - q \cdot (\Phi - \Phi_{\text{ref}}) \quad (2.8)$$

$$E_v = -\chi - q \cdot (\Phi - \Phi_{\text{ref}}) - E_g \quad (2.9)$$

wobei das Referenzpotential Φ_{ref} auf das Fermi-Potential des intrinsischen Halbleiters gesetzt wird und q die Elementarladung ist. Da die Quasi-Fermi-Niveaus nur im Gleichgewichtsfall einfach zu berechnen sind, ist für die Berechnung der Ladungsträgerdichten jeweils eine zusätzliche Bestimmungsgleichung erforderlich.

2.1.3 Kopplung der Ladungsträger: Poisson-Gleichung

Bereits aus der Elektrodynamik ist bekannt, dass elektrische Ladungsträger miteinander wechselwirken. Die Kopplung zwischen dem elektrischen Potential Φ und den einzelnen Ladungsträgerdichten wird durch die nichtlineare Poisson-Gleichung

$$\nabla \cdot (\varepsilon \nabla \Phi) = q (n - p + N_D^+ - N_A^-) \quad (2.10)$$

beschrieben. In dem Quellterm auf der rechten Seite der Gleichung ist N_D^+ die aktivierte Donatoren- und N_A^- die aktivierte Akzeptorenkonzentration im Halbleiter, welche wie die Elektronen- und Löcherdichte vom elektrischen Potential abhängen.

2.2 Ladungsträgertransport und Boltzmann-Gleichung

Den Ausgangspunkt für die Ableitung von Ladungsträgertransportgleichungen bildet bei den hier untersuchten Problemstellungen die Boltzmann-Gleichung. Aus dieser können dann vereinfachte Transportgleichungen durch eine Momentenentwicklung abgeleitet werden. Allerdings bleibt an dieser Stelle festzuhalten, dass eine solche Vereinfachung im Einzelfall die Qualität der Simulationsergebnisse wesentlich beeinflussen kann. Die Boltzmann-Gleichung lautet dabei

$$\left[\frac{\partial}{\partial t} + \mathbf{v} \frac{\partial}{\partial \mathbf{x}} + \frac{q\mathbf{F}}{\hbar} \frac{\partial}{\partial \mathbf{k}} \right] f(\mathbf{x}, \mathbf{k}, t) = S(f(\mathbf{x}, \mathbf{k}, t)) \quad (2.11)$$

mit der Gruppengeschwindigkeit $\mathbf{v} = \hbar^{-1} \partial_{\mathbf{k}} E(\mathbf{k})$, dem elektrischen Feld \mathbf{F} und dem Stoßintegral für klassische Teilchen

$$S(f(\mathbf{x}, \mathbf{k}, t)) = \int \left[W(\mathbf{k}', \mathbf{k}, \mathbf{x}) \cdot f(\mathbf{x}, \mathbf{k}', t) - W(\mathbf{k}, \mathbf{k}', \mathbf{x}) \cdot f(\mathbf{x}, \mathbf{k}, t) \right] d^3 \mathbf{k}' \quad (2.12)$$

und der Streurrate $W(\mathbf{k}', \mathbf{k}, \mathbf{x})$, die die Wahrscheinlichkeit angibt, mit der ein Teilchen aus dem Volumen \mathbf{k} in das Volumen \mathbf{k}' gestreut wird. Für quantenmechanische Teilchen gilt zudem das Pauli-Verbot. So muss bei dem Übergang beachtet werden, ob der Zielzustand bereits besetzt ist. Dann folgt für das Stoßintegral

$$S(f(\mathbf{x}, \mathbf{k}, t)) = \int \left[W(\mathbf{k}', \mathbf{k}, \mathbf{x}) \cdot f(\mathbf{x}, \mathbf{k}', t) \cdot (1 - f(\mathbf{x}, \mathbf{k}, t)) \right. \\ \left. - W(\mathbf{k}, \mathbf{k}', \mathbf{x}) \cdot f(\mathbf{x}, \mathbf{k}, t) \cdot (1 - f(\mathbf{x}, \mathbf{k}', t)) \right] d^3 \mathbf{k}' \quad (2.13)$$

Die Multiplikation von Gleichung (2.11) mit einer impulsabhängigen Funktion $\psi = \psi(\mathbf{k})$ und Integration über den Impulsraum liefert dann eine Beschreibungsgleichung für den lokalen Mittelwert dieser Funktion, multipliziert mit der lokalen Teilchendichte. Damit folgt dann

$$\int \psi(\mathbf{k}) \frac{\partial f}{\partial t} d^3 \mathbf{k} + \int \psi(\mathbf{k}) \mathbf{v} \frac{\partial f}{\partial \mathbf{x}} d^3 \mathbf{k} + \int \psi(\mathbf{k}) \frac{q\mathbf{F}}{\hbar} \frac{\partial f}{\partial \mathbf{k}} d^3 \mathbf{k} = \int \psi(\mathbf{k}) S(f) d^3 \mathbf{k} \quad (2.14)$$

Werden für $\psi(\mathbf{k})$ Potenzen von \mathbf{k} eingesetzt, so ist Gleichung (2.14) das sogenannte n-te Moment der Boltzmann-Gleichung. Es bleibt festzuhalten, dass der zweite Term auf der linken Seite der Momentengleichung immer das nächsthöhere Moment enthält. Damit ist eine geeignete Abbruchbedingung für das Entwicklungsverfahren nötig. Das n-te Moment der Verteilungsfunktion ist das Integral über die n-te Potenz des Impulses, multipliziert mit der Verteilungsfunktion. Diese Mittelwerte sind damit Sonderfälle der allgemeinen Form

$$n(\mathbf{x}, t) \cdot \langle \psi \rangle = \int \psi(\mathbf{k}) f(\mathbf{x}, \mathbf{k}, t) d^3 \mathbf{k} \quad (2.15)$$

Wie bereits im ersten Kapitel beschrieben, steht zur Modellierung von Quantentransport eine große Zahl an Wegen zur Verfügung. So ist beispielsweise die direkte Lösung von Schrödinger-Poisson-Systemen [18, 19] wie auch der Einsatz von Nichtgleichgewichts Green-Funktionen (NEGF) möglich [20, 21, 22, 23]. Darüberhinaus kann auch eine mit der oben vorgestellten Boltzmann-Gleichung verwandte Methode benutzt werden. Dabei wird eine Bilanzgleichung für die Wigner-Funktion betrachtet [9, 24]. Diese kann als Fourier-Transformierte der Dichtematrix interpretiert werden.

2.2.1 Ableitung des Drift-Diffusions-Modells

Ein Beschreibungsmodell für das Transportverhalten von Ladungsträgern in Halbleitern kann aus den Bilanzgleichungen (2.14) abgeleitet werden. Das einfachste Modell besteht aus jeweils einer Erhaltungsgleichung für die Ladungsträgerdichten von Elektronen und Löchern. Eine Abbruchbedingung wird aus der Annahme, dass sich die Ladungsträger mit dem Halbleiterkristallgitter im thermischen Gleichgewicht befinden, aus dem zweiten und dritten Moment der Boltzmann-Gleichung bestimmt. Damit resultiert das sogenannte Drift-Diffusions-Modell bestehend aus den beiden Kontinuitätsgleichungen

$$\frac{\partial}{\partial t} n(\mathbf{x}, t) - \frac{1}{q} \nabla \cdot \mathbf{j}_n = G - R \quad (2.16)$$

$$\frac{\partial}{\partial t} p(\mathbf{x}, t) + \frac{1}{q} \nabla \cdot \mathbf{j}_p = G - R \quad (2.17)$$

Durch die Poisson-Gleichung (2.10) wird die Kopplung der Ladungsträger beschrieben und das Gleichungssystem abgeschlossen. Mit der Einstein-Relation zur Verknüpfung der Diffusionskoeffizienten D_i mit den Beweglichkeiten μ_i

$$D_i = \frac{kT\mu_i}{q} \quad (2.18)$$

und den Abbruchbedingungen der Kontinuitätsgleichung für die Elektronenstromdichte

$$\mathbf{j}_n = -q\mu_n n \nabla \Phi + qD_n \nabla n \quad (2.19)$$

und die Löcherstromdichte

$$\mathbf{j}_p = -q\mu_p p \nabla \Phi - qD_p \nabla p \quad (2.20)$$

In den nachfolgenden Untersuchungen von Tunnelbauelementen wird, soweit nicht anders angegeben, das Drift-Diffusions-Modell eingesetzt. Damit verbleiben als Unbekannte des Gleichungssystems die Ladungsträgerkonzentrationen der Elektronen n und Löcher p und das elektrostatische Potential Φ . Bei der Ableitung des Drift-Diffusions-Modells aus der Boltzmann-Gleichung wurden die Beweglichkeiten für die Elektronen μ_n und Löcher μ_p eingeführt. Definiert ist die Beweglichkeit durch den Erwartungswert im Impulsraum

$$\mu = \left\langle \frac{q}{\hbar} \nabla_{\mathbf{k}} [\tau \mathbf{v}^T] \right\rangle_{\mathbf{k}} = q \left\langle \tau m^{-1} + \frac{d\tau}{d\varepsilon} \mathbf{v} \mathbf{v}^T \right\rangle_{\mathbf{k}} \quad (2.21)$$

mit der mikroskopischen Relaxationszeit τ , dem inversen Massentensor m^{-1} und der Gruppengeschwindigkeit \mathbf{v} . Die zweite Beziehung gilt, wenn die Relaxationszeit nur über die Energie vom Wellenvektor abhängt. Die mikroskopische Relaxationszeit kann hierbei durch die Streurrate $W(\mathbf{k}, \mathbf{k}', \mathbf{x})$ berechnet werden. Zur Berechnung dieser kann Fermis goldene Regel herangezogen werden

$$W(\mathbf{k}, \mathbf{k}', \mathbf{x}) = \frac{2\pi}{\hbar} \cdot |M_{\mathbf{k}'\mathbf{k}}|^2 \cdot \delta(\varepsilon(\mathbf{k}') - \varepsilon(\mathbf{k}) - \varepsilon^{\text{trans}}(\mathbf{k}, \mathbf{k}')) \quad (2.22)$$

mit dem Matrixelement des Streupotentials $M_{\mathbf{k}'\mathbf{k}}$ und der Transferenergie $\varepsilon^{\text{trans}}$. Insgesamt gibt es in Halbleitern eine große Zahl an unterschiedlichen Streuprozessen:

- Optische Zwischentalstreuung
- Akustische Deformationspotentialstreuung
- Streuung an ionisierten Störstellen
- Elektron-Loch-Streuung
- Oberflächenstreuung an Materialgrenzflächen

Unter der Annahme, dass die einzelnen Streuprozesse unabhängig voneinander sind, kann die Trägerbeweglichkeit nach der Regel von Mathiessen aus den Beweglichkeiten der Teilprozesse näherungsweise bestimmt werden

$$\frac{1}{\mu} = \frac{1}{\mu_1} + \frac{1}{\mu_2} + \frac{1}{\mu_3} + \dots \quad (2.23)$$

Da allerdings vielfach nur empirische Modelle [25, 26, 27, 28] zur Beschreibung der einzelnen Beweglichkeitsterme vorliegen, ist dieses Vorgehen der einzige Weg um die Gesamtträgerbeweglichkeiten zu berechnen.

2.2.2 Randbedingungen bei der Halbleitermodellierung

Bei Bauelementsimulationen kommt den Randwerten eine erhebliche Bedeutung zu. So werden angelegte Spannungen und Materialwechsel im Wesentlichen durch Rand- und

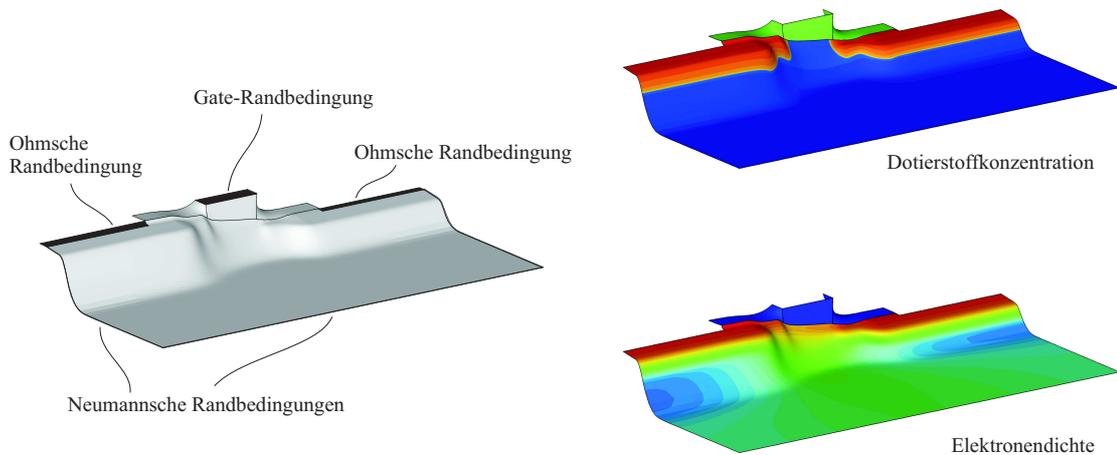


Abb. 2.2: Randbedingungen und gegebene Dotierstoffkonzentration für die Simulation eines zweidimensionalen MOSFETs und daraus resultierende Ergebnisse für das elektrische Potential und die Elektronendichte.

Grenzflächenbedingungen beschrieben. Im Rahmen des Drift-Diffusions-Modells sind hierbei Angaben für das Verhalten der Ladungsträgerdichten und des elektrostatischen Potentials erforderlich. Aus mathematischer Sicht werden dazu Dirichlet-, Neumann- und gemischte Randbedingungen eingesetzt. Allerdings müssen diese mit der physikalischen Wirklichkeit der Problemstellung im Einklang stehen. Von zentraler Bedeutung sind die stromführenden Kontakte der Halbleiterbauelemente. Meist werden (leitende) Ohmsche Kontakte und Gatekontakte benötigt. Zusätzlich ist eine Aussage darüber notwendig, was an den verbleibenden sonstigen Berandungen zu gelten hat. In Abbildung 2.2 (links) sind beispielhaft die erforderlichen Bedingungen für die Bauelementsimulation eines lateralen MOSFETs eingezeichnet.

Im Fall eines Ohmschen Kontakts befindet sich der Halbleiter im thermodynamischen Gleichgewicht und damit gilt $np = n_i^2$ und Ladungsträgerneutralität. Ausgehend von den Beziehungen (2.6) und (2.7) können die Randbedingungen für das elektrische Potential und die Ladungsträgerdichten bestimmt werden. Mit Hilfe des Fermi-Potentials Φ_F , das der angelegten Spannung entspricht, folgen schließlich Dirichlet-Randbedingungen für das elektrische Potential und die Ladungsträgerdichten. Im Hinblick auf Transistoren sind Gatekontakte von erheblicher Bedeutung. So wird das Bauelementverhalten wesentlich von der angelegten Gatespannung gesteuert. Für das elektrostatische Potential an einem Gatekontakt gilt

$$\Phi_g = \Phi_F - \Delta\Phi_{WF} \quad (2.24)$$

mit dem Fermi-Potential Φ_F , das erneut der angelegten Spannung entspricht, und der Differenz der Arbeitsfunktion $\Delta\Phi_{WF}$ zwischen dem Metall und einem intrinsischen Referenz-Halbleiter. An allen anderen Berandungen gelten homogene Neumannsche Randbedingungen für das elektrische Feld und die Ladungsträgerstromdichten.

Insgesamt folgt aus den Gleichungen (2.10) und (2.16) bis (2.20) im Zusammenspiel mit den vorgestellten Randbedingungen ein vollständiges Differentialgleichungssystem zur

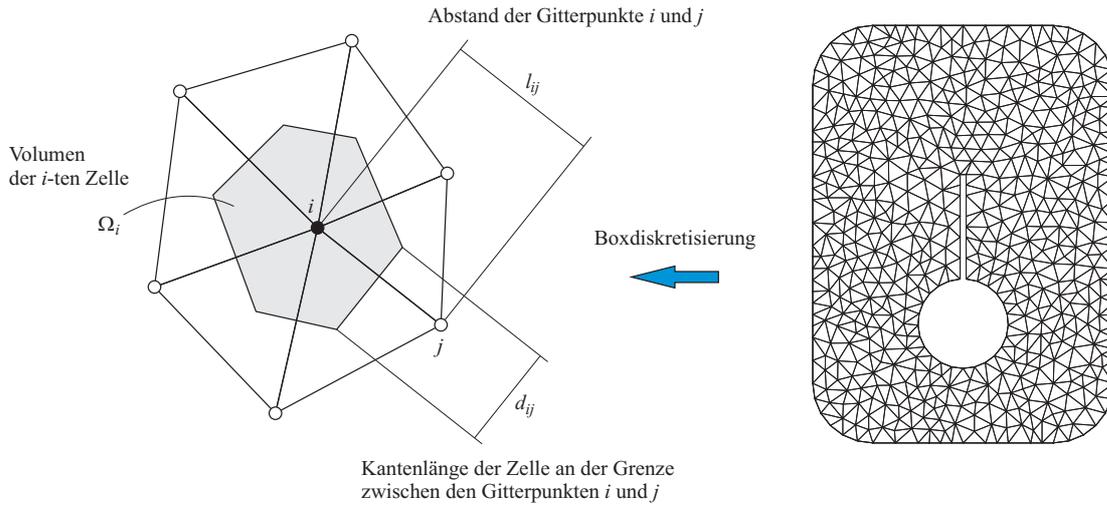


Abb. 2.3: Diskretisierung mittels der Boxmethode: Definition der Zellengrößen (links) und Beispiel eines zweidimensionalen Gitters (rechts).

numerischen Beschreibung von klassischen Halbleitertransportproblemen. Angewandt auf das Beispiel in Abbildung 2.2 folgen dann, ausgehend von der oben rechts gezeigten Dotierstoffkonzentration, die dargestellten Resultate für das elektrische Potential (siehe Höhenprofil der Teilbilder) und die Elektronendichte (rechts unten).

2.2.3 Scharfetter-Gummel-Diskretisierung der Transportgleichungen

Um das gegebene partielle Differentialgleichungssystem numerisch zu lösen, müssen die einzelnen Differentialgleichungen jeweils in geeigneter Weise diskretisiert werden. Die einfachste Vorgehensweise ist die Finite-Differenzen-Methode, die allerdings im Fall der beiden Kontinuitätsgleichungen erhebliche Probleme hinsichtlich der Konvergenz und Stabilität aufweist. Aus diesem Grund wird die besser geeignete Scharfetter-Gummel-Boxmethode [29] in der Halbleitertransportsimulation eingesetzt. In Abbildung 2.3 sind die wesentlichen geometrischen Kenngrößen der Boxmethode und ein zweidimensionales Beispielgitter dargestellt. Mit der Bernoulli-Funktion

$$B(x) = \frac{x}{\exp(x) - 1} \quad (2.25)$$

folgt die diskrete Form der beiden Kontinuitätsgleichungen (2.16) und (2.17)

$$\sum_j \mu_n^{(i)} \frac{d_{ij}}{l_{ij}} \left[B(\Phi_j - \Phi_i) n_j - B(\Phi_i - \Phi_j) n_i \right] = q \Omega_i \left(R_i - G_i + \frac{dn_i}{dt} \right) \quad (2.26)$$

$$\sum_j \mu_p^{(i)} \frac{d_{ij}}{l_{ij}} \left[B(\Phi_i - \Phi_j) p_i - B(\Phi_j - \Phi_i) p_j \right] = q \Omega_i \left(R_i - G_i + \frac{dp_i}{dt} \right) \quad (2.27)$$

mit den über die Zelle gemittelten Beweglichkeiten $\mu_n^{(i)}$ und $\mu_p^{(i)}$ und der Summation über alle Nachbarzellen j der Zelle i . Für die Poisson-Gleichung (2.10) ergibt sich

$$\sum_j \varepsilon^{(i)} \frac{d_{ij}}{l_{ij}} (\Phi_i - \Phi_j) = \Omega_i \varrho_i \quad (2.28)$$

Aus praktischen Erwägungen wird diese diskrete Form der Differentialgleichungen meist elementweise zugewiesen, da somit die materialabhängigen Größen auch elementweise berechnet werden können. Gerade an Materialgrenzflächen, an denen Materialparameter unstetig sind, ist ein solches Vorgehen von Vorteil.

2.3 Generationsmodelle

2.3.1 Shockley-Read-Hall-Rekombination

Ein entscheidender Mechanismus in Halbleitern ist der Übergang eines Elektrons aus dem Leitungsband auf ein Störstellenniveau in der Bandlücke und die daran folgende Rekombination mit einem Loch, das aus dem Valenzband auf dieses Störstellenniveau springt. Weil dieser nichtstrahlende, mehrstufige Prozess deutlich wahrscheinlicher ist als das direkte Rekombinieren von Elektron und Loch über die gesamte Bandlücke, wird die Shockley-Read-Hall-Rekombination über tiefe Störstellenniveaus dominiert [30]. Solche tiefe Störstellen entstehen durch Kristalldefekte und Dotieratome, die im Wesentlichen Punktdefekte darstellen. Die Rekombinationsrate hat dann die Form

$$R^{\text{SRH}} = \frac{np - \gamma_n \gamma_p n_i^2}{\tau_p (n + \gamma_n n_i) + \tau_n (p + \gamma_p n_i)} \quad (2.29)$$

mit den dichteabhängigen Termen $\gamma_n(n)$ und $\gamma_p(p)$. Für die Lebensdauer der Elektronen und Löcher gilt folgender Zusammenhang

$$\tau_c = \tau_d \cdot \frac{f(T)}{1 + g_c(F)} \quad (2.30)$$

wobei $c \in \{n, p\}$. Die Lebensdauer ist hierbei von der Dotierung, der Temperatur und der elektrischen Feldstärke abhängig. Aus experimentellen Befunden folgt die Abhängigkeit der Lebensdauern von der Dotierstoffkonzentration [31]. Dieser Zusammenhang ist schon deshalb nachvollziehbar, da Dotieratome als Störstellen interpretiert werden können. Der Einfluss der Dotierung auf die Lebensdauer wird letztlich durch die Scharfetter-Relation beschrieben

$$\tau_d = \tau_{\min} + \frac{\tau_{\max} - \tau_{\min}}{1 + \left(\frac{N_D + N_A}{N_0} \right)^{\gamma_d}} \quad (2.31)$$

Mittels Messungen kann ein Anstieg der Minoritätsträgerlebensdauer mit der Temperatur beobachtet werden [32]. Die Abhängigkeit der Lebensdauer von der Temperatur wird

durch das Potenzgesetz

$$f(T) = \left(\frac{T}{T_0} \right)^{-3/2} \quad (2.32)$$

mit $T_0 = 300$ K modelliert. Aus der Theorie folgt nach einigen Näherungen der gleiche Zusammenhang [33]. Eine Behandlung der Shockley-Read-Hall-Rekombination im Fall von Grenzflächenzuständen erfolgt analog zu der im Bulkmaterial. Auf den Einfluss der elektrischen Feldstärke wird im nächsten Abschnitt über defektunterstütztes Tunneln noch genauer eingegangen.

2.3.2 Defektunterstütztes Tunneln

Bei hohen elektrischen Feldern folgt eine nicht-verschwindende Wahrscheinlichkeit, dass Teilchen über Störstellenniveaus durch die Bandlücke tunneln. Grundsätzlich lassen sich defektunterstützte Tunnelprozesse mit der SRH-Rekombination beschreiben. Unter der Annahme, dass nur ein Störstellenniveau in der Bandlücke aktiv ist, erhält man mit Hilfe von Gleichung (2.30) eine Rekombinationsrate mit einer feldabhängigen Korrektur der Lebensdauern. Nach Schenk kann ein Modell zur Beschreibung von defektunterstütztem Tunneln aus der Analyse feldunterstützter Multiphonon-Übergänge (hervorgerufen durch Elektron-Phonon-Kopplung) abgeleitet werden [6]. Hierbei wird angenommen, dass das elektrische Feld konstant über der Tunnelwegstrecke ist und das Phononenspektrum in Silizium wird durch einen effektiven Phononenmodus der Energie $\hbar\omega_0$ ersetzt. Mit dem Huang-Rhys-Faktor S als Maß für die Diagonalelemente der Elektron-Phonon-Kopplung folgt die Gitterrelaxationsenergie $\epsilon_R = S\hbar\omega_0$ und daraus das Energieniveau des optimalen horizontalen Übergangspfades

$$E_0 = 2\sqrt{\epsilon_F} \cdot \left(\sqrt{\epsilon_F + E_{tc} + \epsilon_R} - \sqrt{\epsilon_F} \right) - \epsilon_R \quad (2.33)$$

mit der Abkürzung

$$\epsilon_F = \frac{(2\epsilon_R kT)^2}{(\hbar\Theta_c)^3} \quad (2.34)$$

und der elektro-optischen Frequenz

$$\Theta_c = \left(\frac{e^2 F^2}{2\hbar m_c} \right)^{1/3} \quad \text{mit } c \in \{n, p\} \quad (2.35)$$

Hierbei ist m_c die Ladungsträgertunnelmasse in Richtung der elektrischen Feldstärke, und für die modifizierte Störstellenenergie gilt

$$E_{tc} = \begin{cases} \frac{1}{2}E_g + \frac{3}{4}kT \ln \left(\frac{m_n}{m_p} \right) - E_t & \text{für Elektronen} \\ \frac{1}{2}E_g - \frac{3}{4}kT \ln \left(\frac{m_n}{m_p} \right) + E_t & \text{für Löcher} \end{cases} \quad (2.36)$$

Auf diese Weise resultiert für den feldabhängigen Teil der Minoritätsträgerlebensdauer in der Niedertemperaturnäherung:

$$\begin{aligned}
g_c(F) &= \left(1 + \frac{(\hbar\Theta_c)^{3/2}\sqrt{E_{tc} - E_0}}{E_0\hbar\omega_0}\right)^{-1/2} \frac{(\hbar\Theta_c)^{9/4}(E_{tc} - E_0)^{1/4}}{2(kT)^{3/2}\sqrt{E_{tc}E_0}} \times \\
&\times \exp\left[-\frac{E_{tc} - E_0}{\hbar\omega_0} + \frac{E_{tc} - E_0}{kT} + \frac{\hbar\omega_0 - kT}{2\hbar\omega_0} - \frac{4}{3}\left(\frac{E_{tc} - E_0}{\hbar\Theta_c}\right)^{3/2}\right] \times \\
&\times \exp\left(\frac{E_{tc} + kT/2}{\hbar\omega_0} \ln\left(\frac{E_{tc}}{\epsilon_R}\right) - \frac{E_0}{\hbar\omega_0} \ln\left(\frac{E_0}{\epsilon_R}\right)\right) \quad (2.37)
\end{aligned}$$

Es ist dabei zu beachten, dass die Energie des Übergangspfades E_0 für Elektronen und Löcher unterschiedlich ist, da diese sowohl von der elektro-optischen Frequenz Θ_c , als auch von der modifizierten Störstellenenergie E_{tc} abhängt. An dieser Stelle sei für eine detaillierte Erklärung der einzelnen Komponenten auf [17] verwiesen. Im Fall niedriger elektrischer Felder gilt $g_c \ll 1$, weshalb die Minoritätslebensdauern der Elektronen und Löcher dann in der Shockley-Read-Hall-Rekombinationsrate feldunabhängig sind. Das vorgestellte Modell ist lokal und berücksichtigt deshalb die Änderung der Fermi-Niveaus über der Tunnelwegstrecke nicht. Zudem wird angenommen, dass immer Anfangs- und Endzustände für den Tunnelvorgang existieren. Durch ein nicht-lokales Modell kann die auftretende unphysikalische Sättigung der störstellenunterstützten Generation bei steilen pn-Übergängen verhindert werden. Alternativ besteht die Möglichkeit durch eine lokale Dichtekorrektur diese Sättigung ebenso zu beheben. Dazu werden in Gleichung (2.29) die Elektronen- und Löcherdichte durch folgende Terme ersetzt:

$$n \rightarrow n \left(\frac{n_i}{N_c}\right)^{\frac{|\nabla F_n|}{F}} \quad \text{und} \quad p \rightarrow p \left(\frac{n_i}{N_v}\right)^{\frac{|\nabla F_p|}{F}} \quad (2.38)$$

Ein alternatives Beschreibungsmodell basiert auf dem Artikel von Hurkx [34] und geht ebenso von einem linearen Potential über der Tunnelwegstrecke aus. In den folgenden Untersuchungen wird allerdings das vorgestellte, physikalisch besser fundierte Modell von Schenk bevorzugt.

2.3.3 Stoßionisation

Ein weiterer wichtiger Generationsterm in Halbleitern ist die Stoßionisation. Durch diese können Elektron-Loch-Paare erzeugt werden, wenn einerseits die elektrische Feldstärke über einem kritischen Grenzwert liegt und andererseits die Raumladungszone ausreichend ausgedehnt ist. Die Rate hat dabei die Form

$$G^{ii} = \alpha_n n v_n + \alpha_p p v_p \quad (2.39)$$

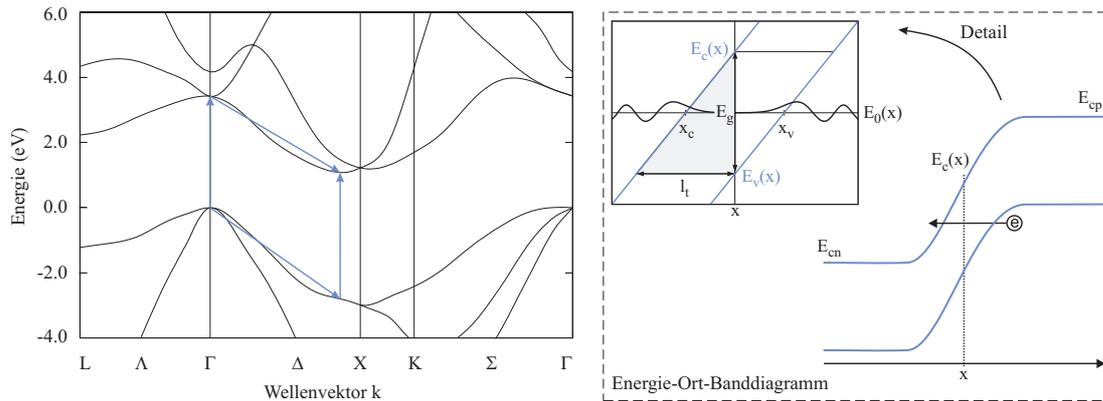


Abb. 2.4: Band-zu-Band Tunneln im indirekten Halbleiter Silizium [17]: Übergangspfade im Energie-Impuls (links) und im Energie-Ort Banddiagramm (rechts).

mit den Driftgeschwindigkeiten der Elektronen v_n und Löcher v_p und den entsprechenden Ionisationskoeffizienten nach Chynoweth [35] und van Overstraeten und de Man [36]

$$\alpha_c = \gamma^{ii} a_c \exp\left(-\frac{\gamma^{ii} b_c}{F^{ii}}\right) \quad (2.40)$$

wobei der Vorfaktor γ^{ii} temperaturabhängig und die treibende Kraft F^{ii} das elektrische Feld in Stromrichtung ist. Inwieweit letztlich die Stoßionisation bei Tunnelbauelementen von Bedeutung sein kann, wird in den nachfolgenden Kapiteln genauer untersucht.

2.4 Phononenunterstütztes Tunneln

Bei Bauelementen wie dem Tunnelfeldeffekttransistor beruht das Funktionsprinzip im wesentlichen auf dem Tunneln zwischen den Energiebändern. Hierbei dominiert zumeist das phononenunterstützte Tunneln. Dieses kann bei Esakioden anhand des negativen differentiellen Widerstandes beobachtet werden. Beruhend auf demselben Prinzip tritt bei vielen modernen nanostrukturierten Bauelementen ein meist unerwünschter Tunnelstrom auf (*gate-induced drain-leakage*). Das Ziel dieses Abschnitts ist es einen Einblick in das phononenunterstützte Band-zu-Band Tunnelmodell von Schenk [6] zu geben. In diesem Zusammenhang wird jedoch nicht auf die konkrete Herleitung eingegangen, da diese in der Literatur nachgelesen werden kann. Wesentlich wichtiger ist vielmehr die Erklärung der zugrundeliegenden Annahmen und die Untersuchung des Gültigkeitsbereiches des Modells.

Für Silizium ist das Energie-Impuls Banddiagramm in Abbildung 2.4 (links) gezeigt. Es gibt zwei Übergänge vom Valenzbandmaximum im Γ -Punkt zum Leitungsbandminimum am Δ_1 -Punkt: Zum einen über das Leitungsband im Gammapunkt und zum anderen über das Valenzband im Δ_1 -Punkt. Der Weg von Ladungsträgern durch die Bandlücke kann als Transport auf einem imaginären k-Band interpretiert werden. Im Fall direkter Halbleiter

findet der Übergang ohne die Unterstützung von Phononen statt, ist aber ebenso über ein k -Band im imaginären Raum zu beschreiben.

Das Modell für phononenunterstütztes Interbandtunneln wird über den Formalismus der Green-Funktionen abgeleitet. Auf den folgenden Seiten werden die wesentlichen Schritte aus [6] wiederholt. Eine Beschreibung von Band-zu-Band Tunneleffekten ist durch eine Tunnelstromdichte möglich. Diese kann durch Integration der Tunnelleitfähigkeit über die Feldstärke berechnet werden

$$j_t = \int_0^F dF' \sigma(F') \quad (2.41)$$

Unter Vernachlässigung von inelastischen Kollisionen kann die Leitfähigkeit in einen Inter- und einen Intradband-Anteil zerlegt werden. Für den Leitfähigkeitstensor folgt unter Anwendung der Kubo-Formel [37]

$$\sigma(\omega, \mathbf{F}) = i \frac{q^2 N}{m_0(\omega + i\epsilon)} \mathbf{I} + \frac{1}{\hbar(\omega + i\epsilon)} \int_0^\infty e^{i\omega t} \langle [\mathbf{j}(0), \mathbf{j}(-t)]_- \rangle_\infty dt \quad (2.42)$$

mit der Kurzform für die Korrelationsfunktion des Stromdichteoperators

$$\langle [\mathbf{j}(0), \mathbf{j}(-t)]_- \rangle_\infty = \lim_{t' \rightarrow \infty} \mathbf{Sp} \{ \rho(t') [\mathbf{j}(0), \mathbf{j}(-t)]_- \} \quad (2.43)$$

der Frequenz ω , dem Vektor der elektrischen Feldstärke \mathbf{F} , der Zahl der Elektronen N und dem Dichtetensor des großkanonischen Ensembles $\rho(t)$. Die im Dichtetensor enthaltene Hamiltonfunktion beschreibt die Elektron-Phonon-Kollisionen, wie auch die Kopplung mit dem elektrischen Feld und dem Phononenensemble. Da die Bandlücke zwischen dem Γ - und dem X-Punkt monoton steigt, müssen Elektron- und Loch-Phonon-Kollisionen berücksichtigt werden. Der Stromdichtetensor \mathbf{j} wird in Form der zweiten Quantisierung beschrieben.

2.4.1 Direktes Band-zu-Band Tunneln

Bei der Herleitung der Übergangswahrscheinlichkeiten wurde angenommen, dass das elektrische Feld auf der gesamten Tunnelwegstrecke konstant ist (siehe Abbildung 2.4 rechts). Damit folgt für die Tunnelweglänge

$$l_t = |x_c - x_v| = \frac{E_g^d}{q|F|} \quad (2.44)$$

mit der direkten Bandlücke E_g^d . Unter der Annahme von parabolischen und isotropen Bändern nahe des Gammapunktes resultiert als Stromdichte für direkte Übergänge

$$j_t^{\text{dir}} = 3\gamma_{cv} \cdot \frac{q\mu_\Gamma^2 \Theta_\Gamma^2}{\hbar m_0} \cdot \left[\text{Ai}'^2 \left(\frac{E_g^d}{\hbar\Theta_\Gamma} \right) - \frac{E_g^d}{\hbar\Theta_\Gamma} \text{Ai}^2 \left(\frac{E_g^d}{\hbar\Theta_\Gamma} \right) \right] \cdot (f_v - f_c) \quad (2.45)$$

mit der Airy-Funktion $\text{Ai}(x)$, der Ableitung dieser $\text{Ai}'(x)$, der in dem nächsten Abschnitt noch genauer besprochenen Differenz der Fermi-Verteilungsfunktionen ($f_v - f_c$) und der elektro-optischen Frequenz des Tunnelübergangs

$$\Theta_\Gamma = \left(\frac{q^2 F^2}{2\mu_\Gamma \hbar} \right)^{1/3} \quad (2.46)$$

der Oszillatorstärke

$$\gamma_{cv} = 2 \frac{|p^{cv}(0)|^2}{m_0 E_g^d} \quad (2.47)$$

und der reduzierten effektiven Masse μ_Γ im Gammapunkt ($\mathbf{k} = 0$). Für das verbliebene Impulsmatrixelement p^{cv} wird im nächsten Abschnitt eine Näherung angegeben.

2.4.2 Indirektes Band-zu-Band Tunneln

Wird ein indirektes Halbleitermaterial wie Silizium betrachtet, ist die Situation ein wenig komplizierter. In diesem gilt für die direkte Bandlücke $E_g^d(\Gamma) = 3.42 \text{ eV}$ und für die minimale indirekte Bandlücke $E_g^i = E_c(\Delta_1) - E_v(\Gamma) = 1.12 \text{ eV}$. Wie aus Abbildung 2.4 zu entnehmen ist, sind nun zwei Übergangspfade bei der Herleitung zu betrachten.

Stromdichte

Nach einer etwas längeren Rechnung folgt mit dem Wellenvektor im Δ_1 -Punkt \mathbf{k}_0 und der Differenz der Fermi-Verteilungsfunktionen $D^* = f_v(0) - f_c(\mathbf{k}_0)$ ein Ausdruck für die Tunnelstromdichte

$$j_t^{\text{ind}} = A^* \cdot D^* \cdot F \cdot \sum_{\alpha=x,y,z} \sqrt{m_\perp^\alpha m_\parallel^\alpha} \cdot \left\{ (E_g^i \mp \hbar\omega_0) f_B H(x_\alpha^\mp) + (E_g^i \pm \hbar\omega_0) (f_B + 1) H(x_\alpha^\pm) \right\} \quad (2.48)$$

mit der Phononenbesetzungszahl f_B (Bose-Einstein-Verteilung) und dem Vorfaktor

$$A^* = \frac{12\pi q^2}{\hbar^5} \cdot M^{\text{ind}} \cdot \frac{\sqrt{m_t m_v^{3/2}}}{m_0^2} \quad (2.49)$$

Auf den Term M^{ind} wird im übernächsten Teilabschnitt etwas genauer eingegangen, da für die Implementierung in einen Bauelementsimulator einige zusätzliche Annahmen zu treffen sind. Die Funktion

$$H(x) = \frac{\text{Ai}(x)}{x^2} + \frac{\text{Ai}'(x)}{x} + \text{Ai}_1(x) \quad (2.50)$$

enthält neben der bereits erwähnten Airy-Funktion auch das Integral

$$\text{Ai}_1(x) = \int_x^\infty \text{Ai}(y) dy \quad (2.51)$$

und wird ausgewertet an den Stellen

$$x_{\alpha}^{\pm} = 2^{2/3} \cdot \frac{E_g^i \pm \hbar\omega_0}{\hbar\Theta_{r\parallel}^{\alpha}} \quad (2.52)$$

Für die elektro-optische Frequenz gilt im Fall des indirekten Band-zu-Band Tunnelns

$$\hbar\Theta_{r\parallel}^{\alpha} = \left(\frac{q^2 \hbar^2 F^2}{2\mu_{\parallel}^{\alpha}} \right)^{1/3} \quad (2.53)$$

Ausgehend von der effektiven Massennäherung aus Abschnitt 2.1.1 genügt, im Fall von Silizium, zur Beschreibung der Anisotropie der Bänder eine laterale und eine transversale Elektronenmasse. Die longitudinale Masse m_l ist durch die [100]-Richtung definiert und die transversale m_t ist der Anteil in der Ebene orthogonal zur dieser Richtung. Aus diesen effektiven Massen kann die effektive Elektronenmasse in Feldrichtung und orthogonal zu dieser bestimmt werden durch

$$m_{\parallel}^{\alpha} = \frac{m_l m_t}{m_l - (m_l - m_t) \cdot F_{\alpha}^2 / F^2} \quad (2.54)$$

$$m_{\perp}^{\alpha} = \frac{m_l m_t}{m_t - (m_t - m_l) \cdot F_{\alpha}^2 / F^2} \quad (2.55)$$

Damit folgt für die relative effektive Tunnelmasse parallel zur Feldrichtung:

$$\frac{1}{\mu_{\parallel}^{\alpha}} = \frac{1}{m_v} + \frac{1}{m_{\parallel}^{\alpha}} \quad (2.56)$$

Generationsrate

Für die Bauelementsimulation ist eine Rekombinationsrate besser zur Beschreibung von Tunnelströmen geeignet. Unter der Annahme, dass das elektrische Feld über der gesamten Tunnelwegstrecke konstant ist, gilt für die Tunnelstrecke

$$l_t = \frac{E_g^i \pm \hbar\omega_0}{q \cdot F} \quad (2.57)$$

Aus der stationären Kontinuitätsgleichung (2.16) folgt der Zusammenhang zwischen der Tunnelstromdichte und der gesuchten Rekombinationsrate. Da die Tunnelstromdichte der Elektronen entgegengesetzt zu der Richtung des elektrischen Feldes fließt, resultiert aus der Weglänge und der Annahme eines konstanten elektrischen Feldes die Rate

$$R_t^{\text{ind}} = \frac{j_t^{\text{ind}}}{q \cdot l_t} \quad (2.58)$$

Einsetzen von Gleichung (2.48) und (2.57) in Gleichung (2.58) liefert dann

$$R_t^{\text{ind}} = A^* \cdot D^* \cdot F^2 \cdot \sum_{\alpha} \sqrt{m_{\perp}^{\alpha} m_{\parallel}^{\alpha}} \cdot \left[f_B H(x_{\alpha}^{\mp}) + (f_B + 1) H(x_{\alpha}^{\pm}) \right] \quad (2.59)$$

mit der Bose-Einstein-Verteilungsfunktion für die mittlere Phononenbesetzungszahl

$$f_B = \left(\exp \left[\frac{\hbar\omega_0}{kT} \right] - 1 \right)^{-1} \quad (2.60)$$

Werden nach Schenk [17] die Airy-Funktion, das Integral (2.51) und die Ableitung der Airy-Funktion durch asymptotische Näherungen für große Werte ersetzt, folgt

$$H(x) \simeq \frac{x^{-9/4}}{2\sqrt{\pi}} \cdot \exp \left[-\frac{2}{3}x^{3/2} \right] \quad (2.61)$$

und damit auch eine Näherungslösung für die Rekombinationsrate

$$R_t^{\text{ind}} \simeq A \cdot D^* \cdot F^{7/2} \cdot \sum_{\alpha} \sqrt{m_{\perp}^{\alpha} m_{\parallel}^{\alpha}} \cdot \left\{ f_B (F_c^{\alpha\mp})^{-3/2} \exp \left[-\frac{F_c^{\alpha\mp}}{F} \right] + (f_B + 1) (F_c^{\alpha\pm})^{-3/2} \exp \left[-\frac{F_c^{\alpha\pm}}{F} \right] \right\} \quad (2.62)$$

Hierbei bezieht sich das obere Vorzeichen von $F_c^{\alpha\pm}$ auf den Fall der Tunnelgeneration ($f_v > f_c$) und das untere auf Rekombination ($f_v < f_c$). In der Rekombinationsrate ist zudem die kritische elektrische Feldstärke

$$F_c^{\alpha\pm} = \frac{4}{3} \cdot \frac{\sqrt{2\mu_{\parallel}^{\alpha}}}{q\hbar} \cdot (E_g^i \pm \hbar\omega_0)^{3/2} \quad (2.63)$$

mit der effektiven Phononenenergie bei Multiphonon-Übergängen $\hbar\omega_0$ eingeführt. Die kritische Feldstärke ist die obere Schranke für die Gültigkeit des Tunnelmodells. Mit der Rekombinationsrate aus Gleichung (2.62) steht schließlich ein Modell zur Beschreibung von phononenunterstützten Tunnelprozessen in indirekten Halbleitern zur Verfügung. Es zeigt sich, dass dieses Modell noch einige Ungenauigkeiten in den Approximationen der Airy-Funktionen aufweist. Auf diesen Sachverhalt wird in Kapitel 5 genauer eingegangen und letztlich wird eine verbesserte Variante des Tunnelmodells präsentiert.

Berechnung der Vorfaktoren

In dem Vorfaktor M^{ind} aus Gleichung (2.49) sind unterschiedliche Terme enthalten, die bisher nicht spezifiziert wurden. Unter der Annahme, dass die Elektron-Phonon-Kopplung durch akustische Intraband-Streuung beschrieben wird, gilt

$$M^{\text{ind}} \simeq \frac{\hbar\omega_{\text{TA}} \cdot D_{ac,n}^2}{\rho \cdot c_s^2} \cdot \left(\frac{m_0^2}{4\mu_{\Gamma}} \right)^2 \quad (2.64)$$

mit der Massendichte ρ , dem Deformationspotential der akustischen Elektronenstreuung am Kristallgitter $D_{ac,n}$ und der Schallgeschwindigkeit c_s . Zudem ist $\hbar\omega_{\text{TA}}$ die Energie des akustischen Phonons und $\mu_{\Gamma} \approx 0.1m_0$ die reduzierte effektive Masse im Γ -Punkt.

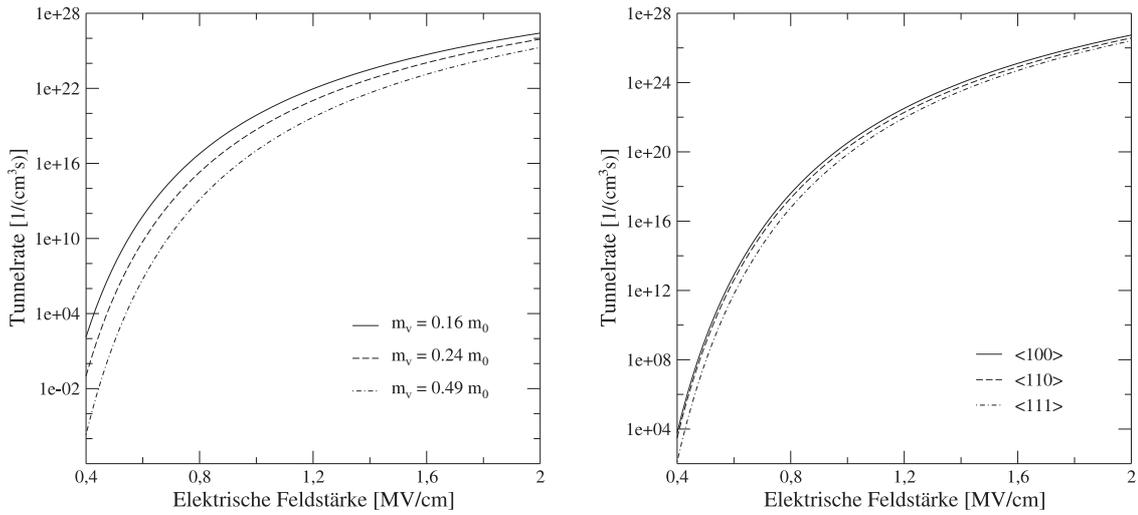


Abb. 2.5: Band-zu-Band Tunnelgenerationsrate in Silizium für unterschiedliche effektive Löchermassen (links) und Kristallorientierungen (rechts).

Die Zahlenwerte der Parameter sind in den Tabellen im Anhang zusammengefasst. Der Zusammenhang zwischen den Vorfaktoren A^* aus Gleichung (2.59) und A aus Gleichung (2.62) wird beschrieben durch

$$A = A^* \cdot \left(\frac{2}{3}\right)^{3/2} \cdot \frac{1}{2\sqrt{\pi}} \quad (2.65)$$

An dieser Stelle bleibt darauf hinzuweisen, dass dieser Vorfaktor von einigen Parametern abhängig ist. So ist der exakte Wert für die effektive Löchermasse m_v ebenso unsicher wie der des Deformationspotentials der akustischen Elektronenstreuung am Gitter. Aus diesem Grund ist der Vorfaktor keine allgemein gültige Konstante. Das Tunnelmodell kann deshalb nur dann sinnvoll eingesetzt werden, wenn bekannt ist, in welcher Weise diese Schwankungen die Simulationsergebnisse beeinflussen. Die Auswirkungen der Wahl unterschiedlicher effektiver Löchermassen ist in Abbildung 2.5 dargestellt. Dabei ist zu erkennen, dass dieser Parameter einen erheblichen Einfluss auf die Rekombinationsrate hat. Die einzelnen Werte sind die effektive Masse des lh-Bandes $m_v = 0.16 m_0$, die des hh-Bandes $m_v = 0.49 m_0$ und die doppelte reduzierte effektive Löchermasse. Zwischen der durchgezogenen und der gestrichelten Linie befindet sich letztendlich der unsichere Bereich, da das hh-Band nur einen geringen Einfluß auf den Tunnelprozess hat.

Es verbleibt damit zu klären, wie die Differenz der Fermi-Verteilungsfunktionen D^* in einem Bauelementsimulator berechnet werden kann. Diese beschreibt die Verfügbarkeit von Elektronen- und Lochzuständen an den Start- und Endpunkten der Tunnelwegstrecke und kann vereinfacht werden zu

$$D^* = f_v(0) - f_c(\mathbf{k}_0) = \frac{n_i}{n_i + \gamma \cdot p} - \frac{n}{n + \gamma \cdot n_i} \quad (2.66)$$

mit dem zusätzlichen Faktor γ . Wie auch im Fall des defektunterstützten Tunnelns kann die unphysikalische Sättigung der Band-zu-Band Tunnelraten durch die Dichtekorrektur aus Gleichung (2.38) vermieden werden.

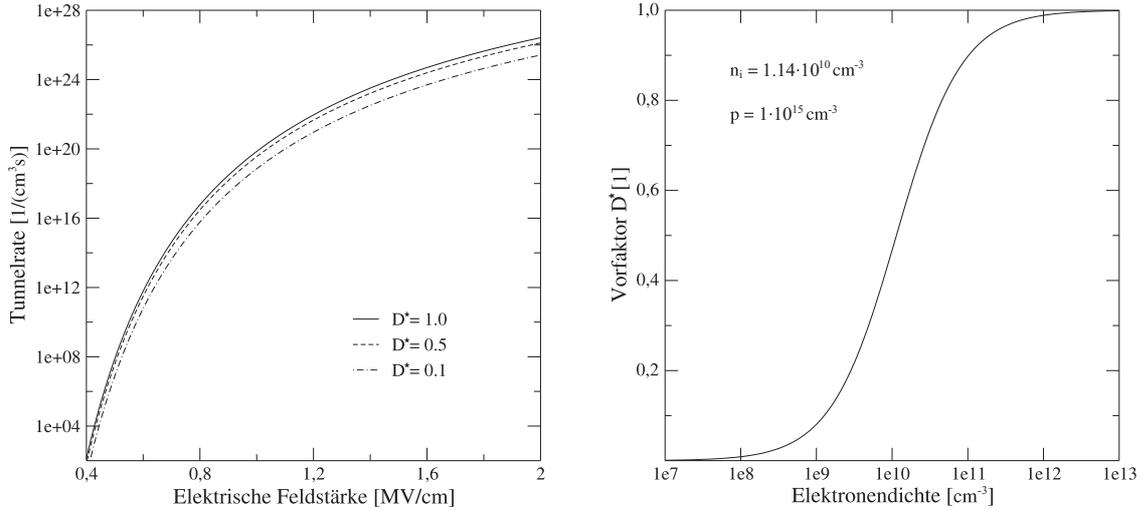


Abb. 2.6: Band-zu-Band Tunnelrate in Silizium in Abhängigkeit von der elektrischen Feldstärke für unterschiedliche Vorfaktoren (links) und in Abhängigkeit von der Elektronendichte (rechts).

Richtungsunabhängiges Tunnelmodell

Da viele kommerzielle Bauelementsimulatoren, wie auch das im folgenden eingesetzte Sentaurus Device [38], keine Unterscheidung der Kristallrichtungen erlauben, wird eine richtungsunabhängige Variante des Tunnelmodells vorgestellt. Unter der Annahme, dass das elektrische Feld in [111]-Richtung weist, gilt für die einzelnen Feldkomponenten

$$F_{\alpha} = \frac{F}{\sqrt{3}} \quad (2.67)$$

Wird dieser Wert in die Rekombinationsrate aus Gleichung (2.62) eingesetzt, resultiert die richtungsunabhängige Tunnelrekombinationsrate

$$R_t^{\text{ind}} \simeq B \cdot D^* \cdot F^{7/2} \cdot \left\{ f_B (F_c^{\mp})^{-3/2} \exp \left[-\frac{F_c^{\mp}}{F} \right] + (f_B + 1) (F_c^{\pm})^{-3/2} \exp \left[-\frac{F_c^{\pm}}{F} \right] \right\} \quad (2.68)$$

in der nun auch F_c^{\pm} aus Gleichung (2.63) nicht mehr richtungsabhängig ist. Für den neuen Vorfaktor gilt hierbei

$$B = A \cdot \frac{9m_l m_t}{m_0 \sqrt{(2m_l + m_t)(2m_t + m_l)}} \quad (2.69)$$

In Abbildung 2.5 (rechts) ist der Einfluss von unterschiedlichen Feldrichtungen auf die Rekombinationsrate dargestellt. Der Einfluss der lokalen Ladungsträgerkonzentrationen ist dadurch eliminiert, dass $D^* = 1$ gesetzt wird. Hierbei zeigt sich deutlich, dass die

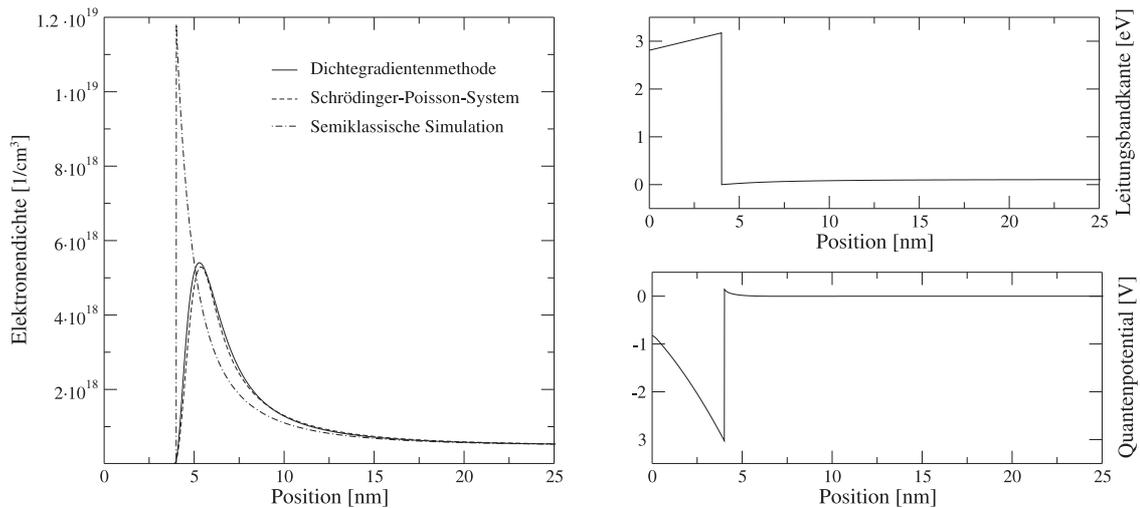


Abb. 2.7: Vergleich der Elektronendichten unter einem MOS-Gate bei semiklassischer (mit und ohne Quantenkorrektur) und quantenmechanischer Simulation (links) und Leitungsbandkante und Quantenpotential im Fall der DGM (rechts).

Unterschiede nur gering sind. Aus diesem Grund ist die vermeintliche Einschränkung auf eine richtungsunabhängige Rate durchaus gerechtfertigt. Es steht zu erwarten, dass die Teilchendichten von Elektronen und Löchern über den Faktor D^* aus Gleichung (2.66) signifikante Auswirkungen auf die Rekombinationsrate haben. In Abbildung 2.6 (links) ist die Tunnelrate in Abhängigkeit von der elektrischen Feldstärke für unterschiedliche Werte von D^* dargestellt. Der direkte Einfluss der Elektronendichte auf die Tunnelrate ist für typische Werte der Ladungsträgerdichten und der intrinsischen Dichte im Fall hoch dotierter und abrupter pn-Übergänge in der rechten Teilabbildung gezeigt. Als Ergebnis bleibt festzuhalten, dass die Auswirkungen der Ladungsträgerdichten bei der Berechnung der Rekombinationsraten nicht einfach vernachlässigt werden dürfen, ohne die Resultate entscheidend zu beeinflussen.

2.5 Auswirkungen von Quanteneffekten auf die Ladungsdichten

Im Kanal von MOS-Bauelementen treten Quantisierungseffekte auf, die sich sowohl in der Elektronenverteilung wie auch der Löcherverteilung widerspiegeln. Diese zeigen sich vornehmlich in einer Reduktion der Teilchendichte an der Si-SiO₂-Grenzfläche unter dem Gateoxid. In Abbildung 2.7 ist eine solche Situation unter einem 4 nm Oxid gezeigt. Dabei ist zu erkennen, dass die klassische Vorgehensweise einen unphysikalischen Maximalwert direkt an der Grenzfläche ergibt. Wegen des extrem großen rechentechnischen Aufwands verbietet sich in den meisten Fällen eine streng quantenmechanische Behandlung (wie die selbstkonsistente Lösung des Schrödinger-Poisson-Systems). Die klassische Behandlung von nanostrukturierten Bauelementen kann jedoch durch zusätzliche Quantenpotentiale für die Elektronen und Löcher gerechtfertigt werden. Diese Potentiale können hierbei als

quantenmechanische Korrekturen der (Quasi-)Fermi-Niveaus interpretiert werden. Den Ausgangspunkt dieser Betrachtungsweise bilden dabei die Arbeiten von Bohm [39, 40] und Ancona [41]. Die Dichte-Gradienten-Methode basiert auf der Annahme, dass das Quantenpotential der Elektronen geschrieben werden kann als

$$\Lambda_n = -\frac{\gamma\hbar^2}{6m\sqrt{n}}\nabla^2\sqrt{n} \quad (2.70)$$

mit der effektiven Masse der Elektronen m und einem empirischen Parameter γ , der die relative Besetzung der unterschiedlichen Subbänder beschreibt. Für die Elektronendichte gilt dann mit Gleichung (2.6) die korrigierte Form

$$n = N_c \cdot \mathcal{F}_{1/2} \left(\frac{F_n - E_c - q\Lambda_n}{kT} \right) \quad (2.71)$$

Die Kombination der Beziehungen (2.70) und (2.71) führt auf eine Gleichung für \sqrt{n} , die selbstkonsistent mit der Poisson- und den jeweiligen Transportgleichungen gelöst werden muss. Mit diesem Vorgehen kann eine äquivalente Gleichung für die quantenmechanische Korrektur der Löcherdichten aufgestellt werden. Aus praktischen Gesichtspunkten ist es vielfach besser eine Differentialgleichung für das Korrekturpotential aufzustellen. Es folgt dann die Differentialgleichung

$$\Lambda = \frac{\gamma\hbar^2\beta}{12m} \left[\nabla^2(\Phi + \Lambda) - \frac{\beta}{2}(\nabla\Phi + \nabla\Lambda) \right] \quad (2.72)$$

mit $\beta = q/kT$. Auch hier folgt eine äquivalente Gleichung für die Löcherdichten. Mit der in Abschnitt 2.2.3 vorgestellten Scharfetter-Gummel-Methode [29] kann diese Gleichung folgendermaßen diskretisiert werden

$$\Omega_i\Lambda_i = \frac{\gamma\hbar^2}{6m_i} \sum_j \frac{d_{ij}}{l_{ij}} \left(1 - \exp \left[\frac{\Phi_i + \Lambda_i}{2kT} - \frac{\Phi_j + \Lambda_j}{2kT} \right] \right) \quad (2.73)$$

Der große Vorteil dieser Methode liegt einerseits in dem relativ geringen Aufwand diese in ein klassisches Bauelements simulationsprogramm einzubinden und andererseits in der nicht übermäßigen Erhöhung der benötigten Rechenzeit. Allerdings ist der Vergleich mit der quantenmechanischen Rechnung das entscheidende Kriterium dieses Verfahren als geeignet zu klassifizieren. In Abbildung 2.7 (links) ist die Elektronendichte unter dem MOS-Gate im Vergleich zu der klassischen und der quantenmechanischen Berechnung aufgetragen. So ist zu erkennen, dass die Dichtegradientenmethode eine hervorragende Übereinstimmung mit der Schrödinger-Poisson Methode aufweist. Auf der rechten Seite der Abbildung 2.7 ist das dazugehörige Leitungsband und das Quantenkorrekturpotential der Elektronen eingezeichnet.

Gerade bei dreidimensional-strukturierten Bauelementen kommt quantenmechanischen Effekten eine erhöhte Bedeutung zu. So wird beispielsweise in einem FinFET der leitende Kanal durch Gräben auch lateral begrenzt. Durch die *silicon-on-insulator* Technologie

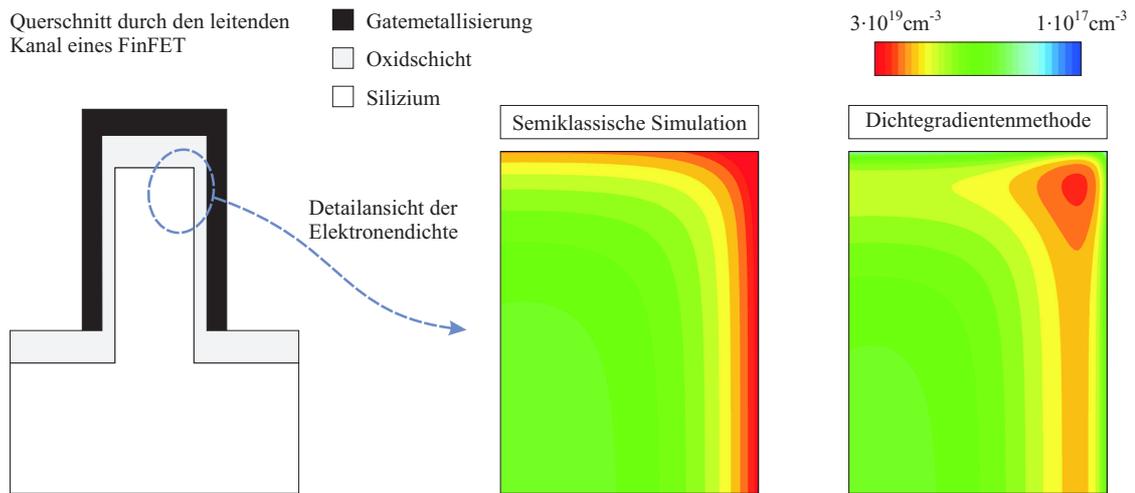


Abb. 2.8: Querschnitt durch den leitenden Kanal eines FinFETs: Schematischer Aufbau und Elektronendichteprofil unter dem Gate mit und ohne quantenmechanische Korrektur durch das Dichtegradientenpotential.

ergeben sich daher mehr und mehr eindimensionale Quantendrähte als Kanalregionen. Es folgen somit Verhältnisse wie im Fall des in Abbildung 2.8 dargestellten FinFETs. Die aus den Simulationen folgenden Elektronenverteilungen im Kanal weisen einen erheblichen Unterschied auf. So ist auch zu erwarten, dass sich die Transporteigenschaften durch die quantenmechanischen Korrekturen signifikant ändern werden. Aus dem vorangehenden Abschnitt ist bekannt, dass auch die Tunnelrate von den Ladungsträgerdichten abhängt und damit gerade bei Tunneltransistoren ein zusätzlicher Einfluss auf die Charakteristiken zu erwarten ist. Auf diesen Sachverhalt wird im vierten Kapitel genauer eingegangen.

3. Numerische Simulation von Tunnelbauelementen

Der grundsätzliche Aufbau und das zugrundeliegende Funktionsprinzip von elementaren Tunnelbauelementen wird in diesem Kapitel vorgestellt. Zu diesem Zweck werden auch erste Simulationsergebnisse präsentiert. Der zweite Teil des Kapitels behandelt dann die Auswirkungen der Modellparameter auf die Simulationsergebnisse. Durch die Kenntnis dieser Einflussfaktoren kann somit prinzipiell, das heißt bei Vorliegen von verlässlichen Prozessierungsdaten, die Bauelementsimulation an die jeweilige Herstellungstechnologie kalibriert werden.

3.1 Untersuchung elementarer Tunnelbauelemente

Den Ausgangspunkt bei der Untersuchung von nahezu allen Tunnelbauelementen stellt die Esaki-Tunneldiode dar. Bei diesem eindimensionalen Bauelement tritt wegen der sehr hohen und abrupten Dotierkonzentrationen am pn-Übergang phononenunterstütztes Band-zu-Band Tunneln auf. Da auch in MOSFETs Tunnelströme eine nicht unerhebliche Rolle spielen, werden auch solche Bauelemente untersucht. Allerdings stellt bei diesen der Band-zu-Band Tunnelstrom nur einen parasitären Effekt und nicht das grundsätzliche Funktionsprinzip dar. Ein besonderer Schwerpunkt liegt hierbei auf der Untersuchung von Kurzkanaleffekten und den Möglichkeiten diese zu beheben. Den zentralen Abschnitt des Kapitels bildet die Untersuchung von Tunnelfeldeffekttransistoren (TFETs) [42]. Diese stellen wegen der einfachen Struktur eine interessante Alternative zu den sehr komplexen Kurzkanal-MOSFETs dar. Im Vordergrund steht die Einführung in die Funktionsweise und die wesentlichen Eigenschaften von TFETs [43, 44]. Eine ausführliche Analyse zur Optimierung dieser Bauelemente wird im vierten Kapitel vorgenommen. Abschließend wird der sogenannte Zener-Feldeffekttransistor (ZFET) betrachtet, der im Wesentlichen eine Kombination aus MOSFET und Zenerdiode darstellt.

3.1.1 Esaki-Tunneldiode

Bislang war es nicht gelungen, das Betriebsverhalten von Esaki-Tunneldioden mit einem konventionellen (das heißt auf der Koppelung von Ladungsträgerbilanzgleichungen und Poissongleichung aufbauenden) Bauelementsimulator realistisch darzustellen. Dies lag

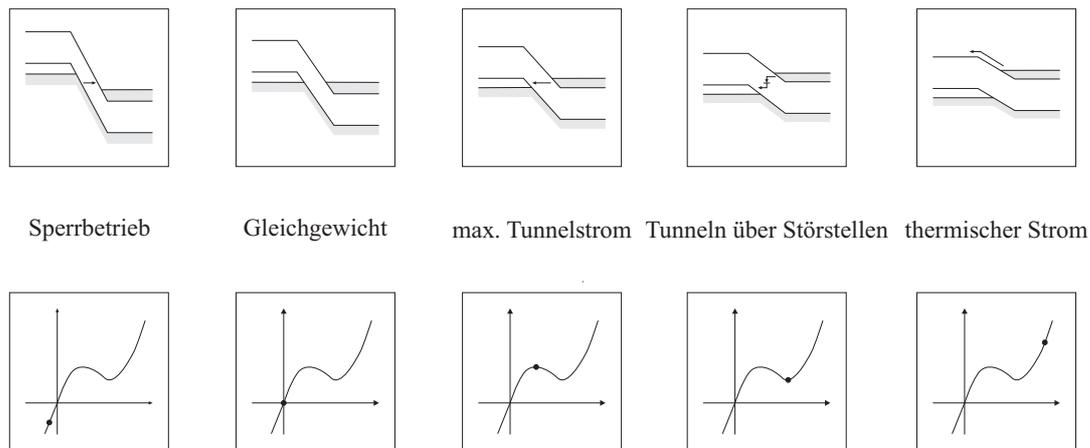


Abb. 3.1: Schematisches Banddiagramm und dazugehörige Strom-Spannungs-Kennlinie einer Esaki-Tunnel diode bei verschiedenen Betriebszuständen.

unter anderem daran, dass das zumeist zugrunde gelegte vereinfachte Kane-Modell [38] die Esaki-Tunnelströme (bzw. die hierzu äquivalenten Rekombinationsraten) nicht korrekt nachbilden kann. Hier konnte nun ein wesentlicher Fortschritt erzielt werden. Durch die in den Kapiteln 2.3 und 2.4 ausführlich besprochene Kombination einer Erweiterung des Kane-Modells nach Schenk [17] für das phononenunterstützte Band-zu-Band-Tunneln mit einer Erweiterung des Shockley-Read-Hall-Störstellenrekombinationsmodells auf den Fall des defektunterstützten Tunnelns gelangt man zu einer Darstellung der jeweiligen Tunnelrekombinationsraten, die strukturell ohne Schwierigkeiten in einen herkömmlichen Bauelementsimulator zu implementieren ist und die Kennlinien von Esaki-Tunnel dioden realitätsgetreu wiedergibt [45].

In Abbildung 3.1 ist in einer Folge von schematischen Energie-Ort Banddiagrammen die Funktionsweise einer Tunnel diode dargestellt. Den Ausgangspunkt bildet hierbei ein hoch dotierter und möglichst abrupter pn-Übergang. Die Dotierstoffkonzentrationen sollten so hoch sein, dass der Halbleiter mindestens in einer der beiden Regionen entartet ist. Auf diese Weise resultiert im thermodynamischen Gleichgewicht (ohne angelegte Spannungen an den Kontakten) der im zweiten Teilbild von links gezeigte Bandkantenverlauf und die angedeutete Elektronenverteilung in den Energiebändern. Eine positive Spannung (d.h. in Vorwärtsrichtung der Diode) führt zu den rechts davon gezeigten Bandverläufen. Im Fall niedriger angelegter Spannungen gibt es für Elektronen aus dem Leitungsband immer auch passende freie Stellen im Valenzband. Da die Barriere wegen der hohen und abrupten Dotierungen sehr gering ist, können die Elektronen durch die Bandlücke tunneln und damit einen Strom tragen. Diese Situation wird solange zu einem Anstieg des Stromes führen, bis die Elektronen im Leitungsband in der n-dotierten Region und die freien Plätze im Valenzband in der p-dotierten Region sich vollständig gegenüberstehen. Anschließend wird der durch das phononenunterstützte Band-zu-Band Tunneln getragene Strom wieder abnehmen. Allerdings besteht dann die Möglichkeit, dass die Elektronen über Störstellen in der Bandlücke (meist Dotieratome) die Barriere überwinden. Ist letztlich die angelegte Spannung hoch genug, dann wird der Hauptanteil des durch die Tunnel diode fließenden

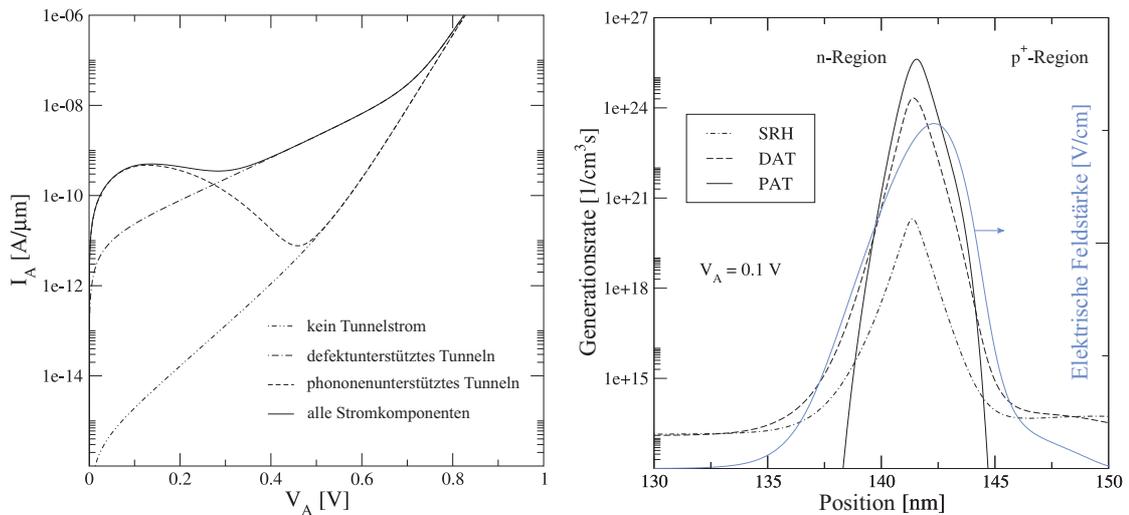


Abb. 3.2: Strom-Spannungs Charakteristik mit den unterschiedlichen Stromanteilen im Vorwärtsbetrieb (links) und Generationsraten (rechts) einer Tunneldiode.

Stromes mittels thermischer Anregung über die immer niedriger werdende Barriere im Leitungsband transportiert. Auf diese Weise resultiert in diesem Kennlinienbereich die Strom-Spannungs-Charakteristik einer herkömmlichen pn-Diode. Insgesamt folgt somit das für Esaki-Tunnelioden charakteristische Maximum in der Kennlinie. In der unteren Bilderreihe in Abbildung 3.1 ist dieses deutlich zu erkennen. Auch im Rückwärtsbetrieb können Elektronen durch die Bandlücke tunneln. Dadurch entsteht der als Zener-Tunneln bezeichnete Strom in Sperrichtung. Da im Fall immer größerer negativer Spannungen die Barriere immer schmaler wird, steigt die Wahrscheinlichkeit für das Band-zu-Band Tunneln immer weiter an und der Zener-Strom nimmt betragsmäßig zu. Dieser Effekt ist im Wesentlichen auch eines der Funktionsprinzipien des Tunnelfeldeffekttransistors.

Die unterschiedlichen simulierten Stromanteile in einer Esaki-Tunnelioden, bestehend aus einer $3 \cdot 10^{20} \text{ cm}^{-3}$ bor- und einer $2 \cdot 10^{19} \text{ cm}^{-3}$ arsen-dotierten Region mit einem 5 nm Übergangsbereich, sind in Abbildung 3.2 dargestellt. Die jeweiligen Beiträge der beiden Tunnelmechanismen sind hierbei klar zu unterscheiden. Damit ist demonstriert, dass die maßgeblichen Tunneleffekte auf Bauelementebene auch im Rahmen der semiklassischen Halbleitergleichungen realistisch simuliert werden können. Die Generationsraten und die elektrische Feldstärke am pn-Übergang bei einer angelegten Spannung von 0.1 V sind in dieser Abbildung auf der rechten Seite gezeigt. Man erkennt deutlich die Korrelation von Feldstärke und Generationsraten. Der Unterschied in der Position der Maxima resultiert aus den Ladungsträgerverteilungen, die die Raten maßgeblich beeinflussen.

3.1.2 MOSFET

Eines der zentralen Halbleiterbauelemente auf Siliziumbasis ist der sogenannte MOSFET (*metal-oxide-semiconductor field-effect transistor*). Die vielleicht wichtigste Anwendung dieses Bauelements liegt in komplementären Logikschaltungen. Hierbei ist es von großem

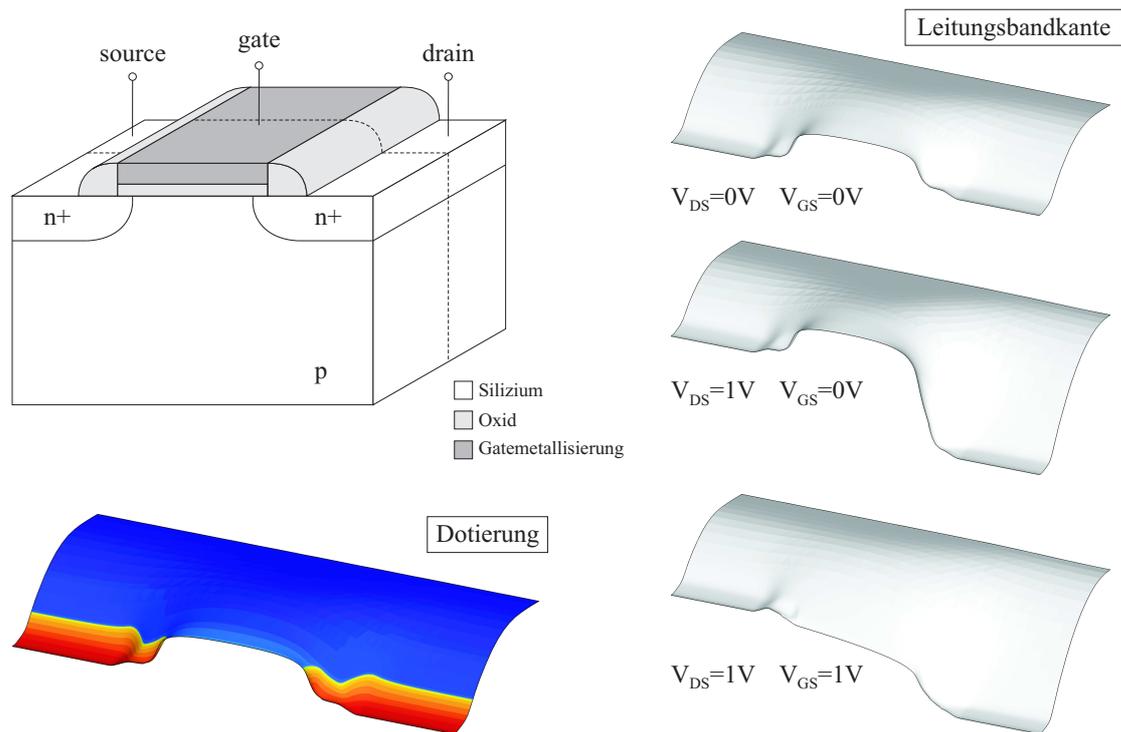


Abb. 3.3: Schemazeichnung und dazugehörige Dotierstoffkonzentration des simulierten MOSFETs mit 100 nm Gatelänge (links) und Profil der Leitungsbandkante im Bauelement bei verschiedenen angelegten Spannungen (rechts).

Vorteil, dass sowohl n- wie auch p-Kanal MOSFETs ohne große technologische Probleme auf Siliziumbasis in einem Chip integriert werden können. Bei diesen stellen entweder die Elektronen (n) oder die Löcher (p) die den Strom tragenden Majoritätsladungsträger im Kanal. Aus der fortschreitenden Miniaturisierung der Bauelementabmessungen ergeben sich allerdings einige Herausforderungen: So treten vielfältige Kurzkanaleffekte auf, die das Bauelementverhalten in erheblichem Maß negativ beeinflussen und deshalb durch komplexe Dotierprofile ausgeglichen werden müssen.

Ein konventioneller MOSFET besteht zumeist aus fünf verschiedenen Regionen: dem Substratmaterial, dem Kanal, der Source- und Drainwanne und dem Gatestapel. Dieser Aufbau ist in Abbildung 3.3 links schematisch dargestellt. Das Gate ist hierbei aus einem Stapel, bestehend aus einer dünnen Schicht an isolierendem Oxidmaterial (z.B. SiO_2) und einem hochdotierten Polysiliziumgatekontakt, aufgebaut. Diese Gatekontaktregion wird allerdings in modernen nanostrukturierten MOSFETs vielfach auch durch ein Metall wie Aluminium ersetzt. Links unten ist in dieser Abbildung die Dotierstoffkonzentration des im Folgenden untersuchten Bauelements gezeigt. Sowohl die Source- wie die Drainwanne sind dabei deutlich zu erkennen. Diese sind jeweils in zwei Gebiete aufgeteilt: Eine hoch dotierte Wanne bei dem Kontakt und eine schwächer dotierte und weniger tiefe Wanne nahe dem Gatestapel. Durch die hohe Dotierstoffkonzentration nahe der Kontakte soll der Widerstand gering gehalten und mittels der flachen, niedriger dotierten Wanne können Kurzkanaleffekte verringert werden. Auf der rechten Seite dieser Abbildung ist das Profil

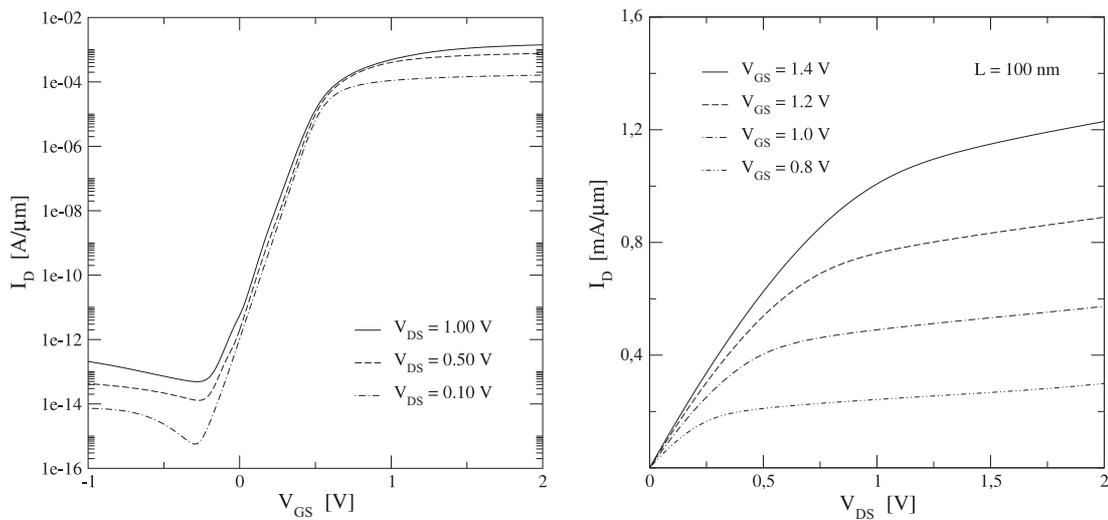


Abb. 3.4: Simulierte Transfer- und Ausgangscharakteristik eines MOSFETs mit einer physikalischen Gatelänge von 100 nm.

der Leitungsbandkante bei verschiedenen Betriebszuständen dargestellt. Sind zwischen den drei Kontakten keine Spannungen angelegt, so resultiert eine Energiebarriere für die Elektronen. Auch von einer hohen Spannungsdifferenz zwischen Drain und Source wird diese Barriere nicht signifikant beeinflusst. Fällt zwischen Gate und Source eine Spannung ab, so reduziert sich die Barrierenhöhe und ein leitender Kanal ist die Folge. In Abbildung 3.4 sind die dazugehörigen Transfer- und Ausgangscharakteristiken dargestellt. Aus den links gezeigten Transferkennlinien ist das sehr gute Schaltverhalten dieses MOSFETs abzulesen. Bei niedrigen Gatespannungen sperrt das Bauelement und nur ein geringer Strom fließt zwischen Source und Drain. Wird die Gatespannung erhöht, folgt ein abrupter Anstieg des Stroms, der dann nahezu in Sättigung geht. Auf diese Weise ist ein An- und ein Aus-Zustand klar definiert. Die entsprechende Ausgangscharakteristik (rechte Seite in Abbildung 3.4) bestätigt dann die guten Bauelementeigenschaften. Allerdings wäre für ein optimales Design des MOSFETs ein besseres Sättigungsverhalten des Drainstroms bei hohen Drain-Source-Spannungen erstrebenswert.

Das Ziel eine höhere Packungsdichte und schnellere Bauelemente und Schaltungen zu fertigen, erfordert die Verkleinerung der Bauelementabmessungen. Eine Verringerung der Kanallänge hat allerdings eine Reihe von Effekten zur Folge, die unter dem Begriff Kurzkanaleffekte zusammengefasst werden. Ebenso wirft eine Verkleinerung der Kanalregion die Frage auf, inwieweit die Annahme einer kontinuierlichen Dotierstoffkonzentration überhaupt noch zulässig ist. So beträgt bei einer Dotierkonzentration von 10^{18} cm^{-3} im Mittel der Abstand zwischen zwei Dotierstoffatomen 10 nm. Diese Situation führt damit zwangsläufig zu statistischen Schwankungen in den Bauelementcharakteristiken, da die einzelnen Dotieratome fertigungsbedingt nicht immer exakt gleichverteilt sind. Ebenso ist eine klare Einteilung in Regionen nicht mehr ohne weiteres möglich. Insgesamt bleibt festzuhalten, dass extrem kurze MOSFETs (Gatelänge $< 5 \text{ nm}$) nur noch eingeschränkt im Rahmen eines Kontinuumsmodells zuverlässig beschrieben werden können. Hierfür

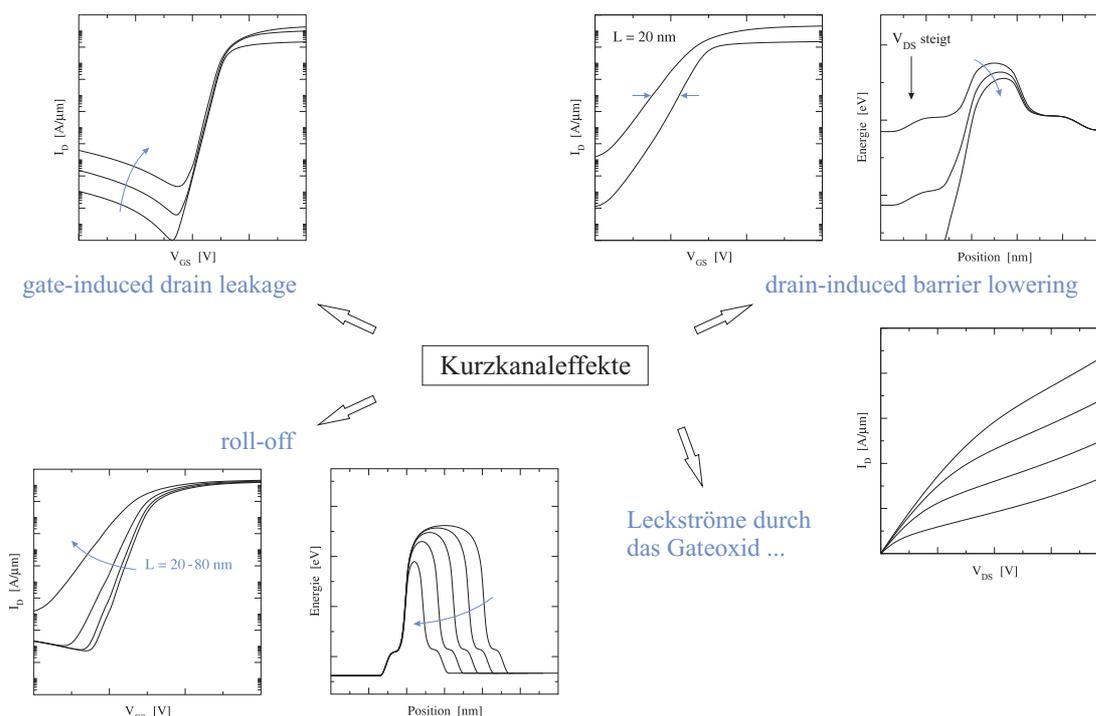


Abb. 3.5: Darstellung einiger Kurzkanaleffekte in lateralen MOSFETs: gateinduzierter Drain-Leckstrom, draininduzierte Barrierenverringern und Verschieben der Einsatzspannung bei unterschiedlichen Kanallängen.

ist es mindestens notwendig in der Bauelementsimulation realistische (d.h. atomistische) Dotierstoffverteilungen einzubeziehen. Eine Auswahl der wichtigsten Kurzkanaleffekte ist in Abbildung 3.5 anhand der Charakteristiken und Leitungsbänder gezeigt. Durch die Verkürzung der Kanallänge folgt das sogenannte *drain-induced barrier lowering* (DIBL) nahezu zwangsläufig. Wie der Name bereits verrät, ist die Barrierenhöhe abhängig von der angelegten Drainspannung. Dadurch ergeben sich die in der Abbildung illustrierten Folgen für die Transfer- und Ausgangskennlinien. So wird das Sättigungsverhalten in der Ausgangskennlinie signifikant schlechter und zudem folgt eine erhebliche Abhängigkeit der Einsatzspannung von der angelegten Drainspannung. Um diesen Effekt zu vermeiden werden sogenannte Halo-Implantationen eingebaut. Im Prinzip handelt es sich hierbei um zusätzliche p-Dotierstoffregionen in der Kanalregion. Allerdings führen diese zu einer Verstärkung des *gate-induced drain-leakage* (GIDL). Dieser Leckstrom im Sperrbereich wird durch phononenunterstütztes Band-zu-Band Tunneln hervorgerufen. Eine weitere Auswirkung kurzer Gate-längen führt neben kleineren Breiten auch zu einer geringeren Höhe der Barriere (*roll-off*). Dadurch folgt eine Verschiebung der Einsatzspannung und ein weniger deutlich ausgeprägtes Schaltverhalten. Da bei kleineren Bauelementen auch das Gateoxid immer dünner wird, steigt die Wahrscheinlichkeit, dass Elektronen durch dieses hindurchtunneln und es auf diese Weise zu Leckströmen zwischen Kanal und Gate kommt. Auch wenn sich viele dieser Kurzkanaleffekte durch komplexere Dotierprofile in ihrer Ausprägung dämpfen lassen, können dennoch nicht alle Probleme bei beliebig kleinen Gate-längen vermieden werden.

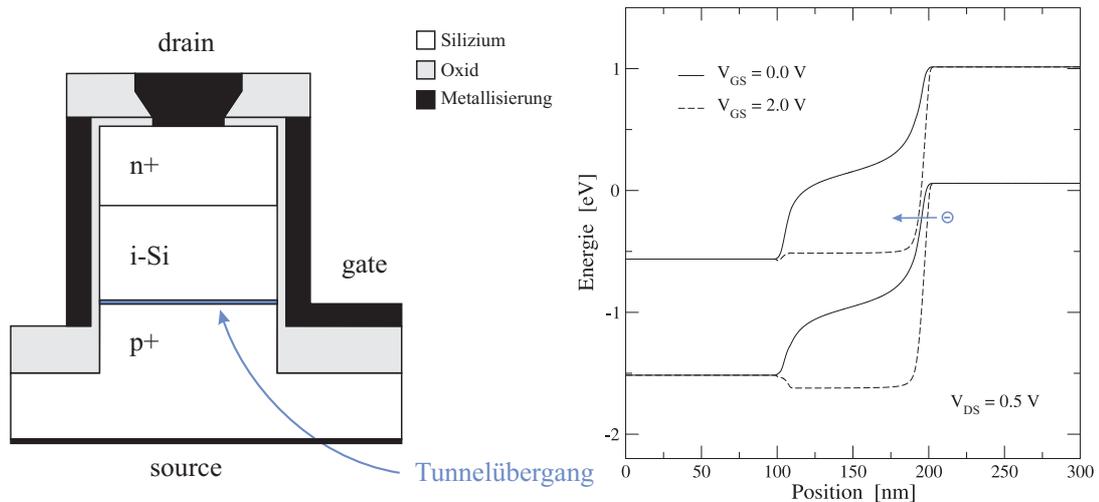


Abb. 3.6: Aufbau eines vertikalen Tunnelfeldeffekttransistors (links) und Banddiagramm bei unterschiedlichen angelegten Gatespannungen (rechts).

3.1.3 Tunneltransistor

Ein alternatives Bauelement, das im Gegensatz zum MOSFET keine Kurzkanaleffekte aufweist, ist der sogenannte Tunnelfeldeffekttransistor (TFET). In diesem Bauelement stellt das Band-zu-Band Tunneln keinen parasitären Effekt dar. Es bestimmt vielmehr das Funktionsprinzip in erheblichen Maß. Der schematische Aufbau des Tunneltransistors in vertikaler Ausführung ist in Abbildung 3.6 gezeigt. Im Prinzip handelt es sich um eine pin-Diode mit einem Gatekontakt über der intrinsischen Region. Wird keine Spannung an das Gate angelegt, so wird auch die Funktionsweise des TFETs von dieser pin-Diode bestimmt. An der Sourceseite werden positive Spannungen angelegt und die Diode wird daher in Sperrichtung betrieben, weshalb extrem geringe Sperrströme resultieren. Den Grund hierfür kann man auf der rechten Seite der Abbildung erkennen: Die Barriere ist so hoch, dass keine thermischen Elektronen über diese transportiert werden können. Wird nun eine positive Gatespannung angelegt, werden das Leitungs- und das Valenzband so stark verbogen, dass unter dem Gatekontakt ein Elektronenkanal entsteht und Elektronen durch die hieraus resultierende schmale Barriere vom Valenz- in das Leitungsband tunneln können.

Um die erwarteten guten Schalteigenschaften zu verifizieren, wird ein Tunneltransistor mit 100 nm Gatelänge und 2 nm Gateoxiddicke untersucht. Der Gatestapel besteht hierbei aus Siliziumdioxid und Aluminium als Kontaktmaterial. Das Siliziumsubstrat ist dabei auf der Sourceseite 10^{20} cm^{-3} bordotiert und auf der Drainseite 10^{20} cm^{-3} arsen-dotiert. Der Kanal ist mit 10^{16} cm^{-3} nur schwach bordotiert. Der Gradient des Dotierprofils beträgt auf der Source- und Drainseite jeweils 2 nm/dec und folgt hierbei jeweils einem Gaußprofil. In Abbildung 3.7 ist die aus der Bauelementsimulation folgende Transfercharakteristik des Tunneltransistors abgebildet. Hieraus sind die sehr vielversprechenden Eigenschaften des TFETs abzulesen. Der Sperrstrom selbst wird durch oberflächennahe Shockley-Read-Hall Rekombination dominiert und bleibt sehr klein. Bei steigender Gatespannung fließt sehr

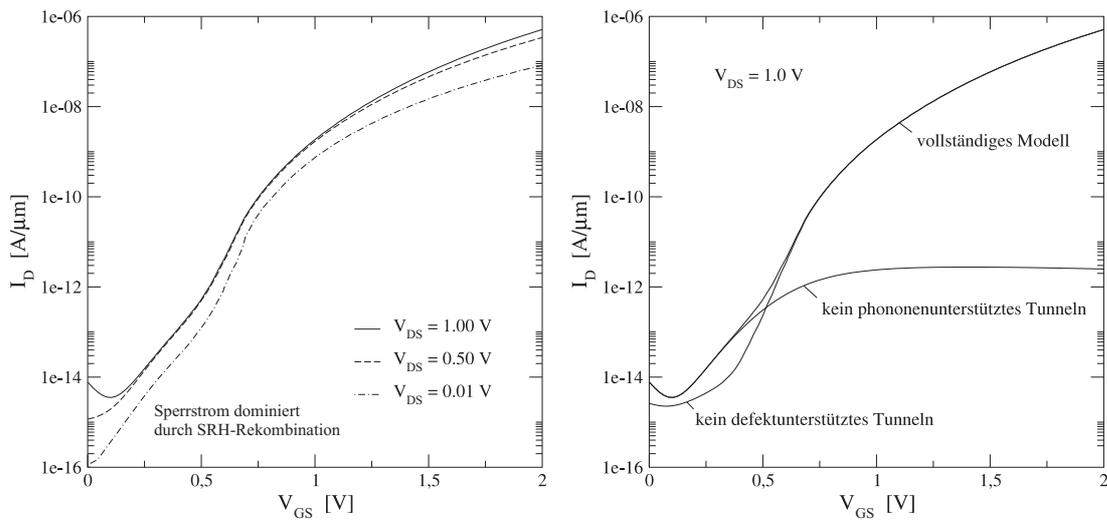


Abb. 3.7: Transfercharakteristik des untersuchten Tunnelfeldeffekttransistors (links) und jeweilige Tunnelstromanteile (rechts).

schnell ein annehmbar hoher Strom. Allerdings ist dieser noch deutlich niedriger als bei einem vergleichbaren MOSFET und damit besteht noch einiger Optimierungsbedarf. Auf der rechten Seite dieser Abbildung sind die, durch verschiedene nacheinander ausgeführte Simulationen identifizierten, einzelnen Stromkomponenten separat eingezeichnet. Dabei ist festzustellen, dass bei niedrigen Gatespannungen erst das defektunterstützte Tunneln dominiert und erst anschließend das reine phononenunterstützte Tunneln an Bedeutung gewinnt. Wird der Einfluss des defektunterstützten Band-zu-Band Tunnelns reduziert, so folgt auch eine abruptere Kennlinie. Auf welche Weise dieses Ziel zu erreichen ist, wird im vierten Kapitel genauer untersucht. Dort werden auch Möglichkeiten zur weiteren Optimierung der Bauelementcharakteristiken vorgestellt.

Nachfolgend wird das Funktionsprinzip des Tunnelfeldeffekttransistors noch ein wenig genauer untersucht. In Abbildung 3.8 sind Konturbilder der Tunnelgenerationsraten unter dem Gateapfel bei verschiedenen angelegten Gatespannungen und die damit verbundene Transferkennlinie bei einer Drain-Source-Spannung von 1.0 V dargestellt. Die einzelnen Bildausschnitte zeigen immer nur kleine Regionen unter dem Gateapfel ($10 \text{ nm} \times 15 \text{ nm}$ im Siliziumsubstrat und, zur besseren Positionsbestimmung, zusätzlich der entsprechende Ausschnitt des Gateapfels). Die Darstellung der Tunnelraten erfolgt im logarithmischen Maßstab und hohe Werte werden durch dunkle Farbtöne beschrieben. Liegt zwischen Gate- und Source keine Spannung an, kommt es am Übergang von der Drainregion zum schwach dotierten Kanal zu einem phononenunterstützten Tunneln, da durch die hohe Drainspannung die Energiebänder ausreichend stark verbogen werden, um die Barriere klein genug werden zu lassen. Wird die Gatespannung sukzessive erhöht, resultiert ein zunehmend stärker ausgeprägter Elektronenkanal unter dem Gate und das Band-zu-Band Tunneln findet dann am Übergang zwischen Sourceregion und dem Kanal statt. Hierbei ist klar zu erkennen, dass die Tunnelregion insgesamt nur einen kleinen Bereich unter dem Gateapfel ausmacht und eine Variation der Gatelänge keinen wesentlichen Einfluss auf

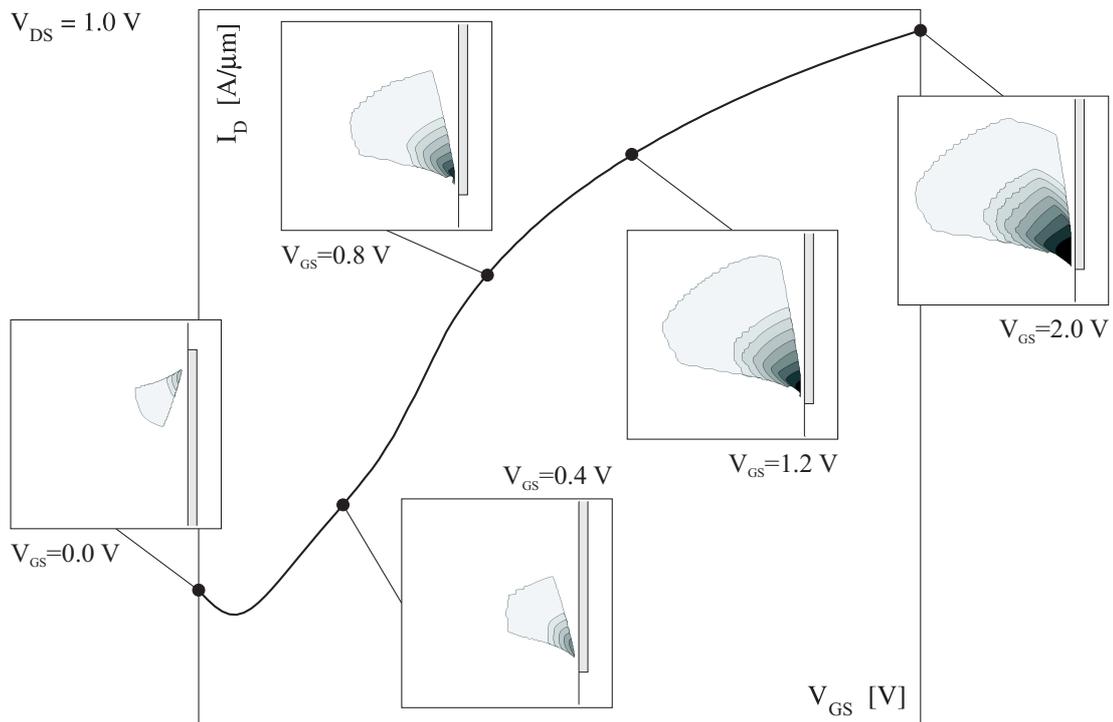


Abb. 3.8: Konturbilder der Tunnelraten des Band-zu-Band Tunnelns bei verschiedenen angelegten Gatespannungen und Einfluss auf die Transfercharakteristik. Die Darstellung erfolgt im logarithmischen Maßstab und hohe Werte werden durch dunkle Farbtöne beschrieben.

die Bauelementeigenschaften haben wird. Im Hinblick auf die im nachfolgenden Kapitel untersuchten Skalierungseigenschaften des Tunneltransistors ist diese Beobachtung von zentraler Bedeutung. Außerdem verschiebt sich die Position der maximalen Tunnelrate mit steigender Gatespannung in Richtung der Ecke des Gatestapels und weist in größerer Tiefe einen Anstieg auf. Wegen der höheren Feldstärke nimmt die Löcherkonzentration unter dem Gate ab und es verschiebt sich der pn-Übergang in Richtung Sourcekontakt. Da die Band-zu-Band Tunnelrate ihr Maximum am pn-Übergang hat, folgt das beschriebene Verhalten. Ein Blick auf die Ausgangscharakteristik in Abbildung 3.9 offenbart das gute Sättigungsverhalten des Tunnelfeldeffekttransistors.

Da ein großer Vorzug des MOSFETs darin besteht mittels n- und p-Kanal Bauelementen komplementäre Logikschaltungen aufzubauen, stellt sich die Frage, ob diese Möglichkeit auch für Tunnelfeldeffekttransistoren existiert. Einen ersten Ansatzpunkt liefert die im vorangehenden Absatz beschriebene Situation bei kleinen Gatespannungswerten. So kann prinzipiell auf beiden Seiten des Tunneltransistors phononenunterstütztes Band-zu-Band Tunneln auftreten. Allerdings ist für eine sinnvolle Anwendung eine etwas modifizierte Beschaltung des Tunneltransistors erforderlich: Bei einem n-Kanal TFET werden sowohl am Drain- wie am Gatekontakt positive Spannungen angelegt und der Sourcekontakt stellt den Referenznullpunkt dar. Dieser Sachverhalt ist im mittleren Teilbild der Abbildung 3.9 schematisch illustriert. Im Gegensatz zum MOSFET ist keine Anpassung der Dotierung

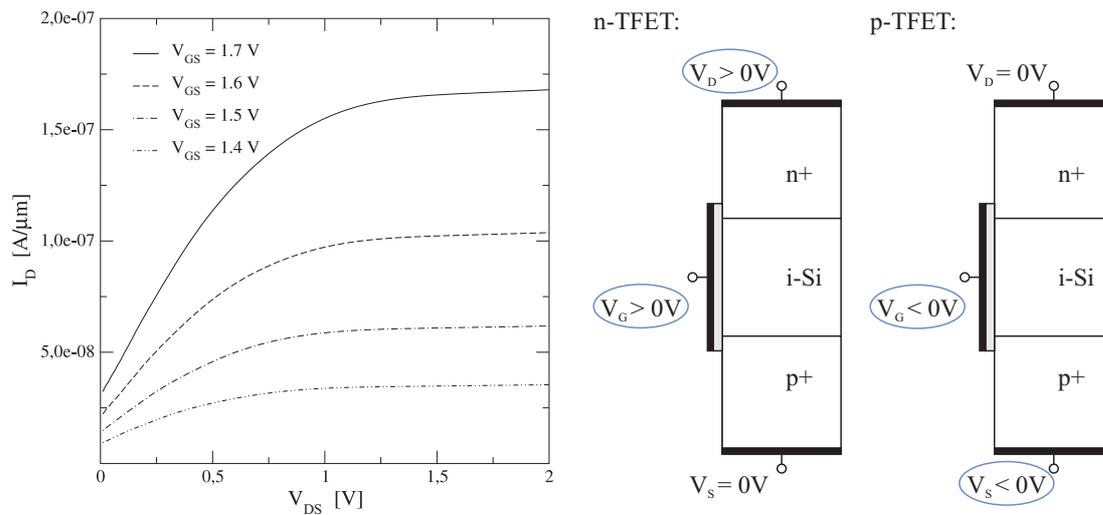


Abb. 3.9: Ausgangskennlinien des untersuchten Tunnelfeldeffekttransistors (links) und Beschaltung im p- und n-Betrieb für komplementäre Logik (rechts).

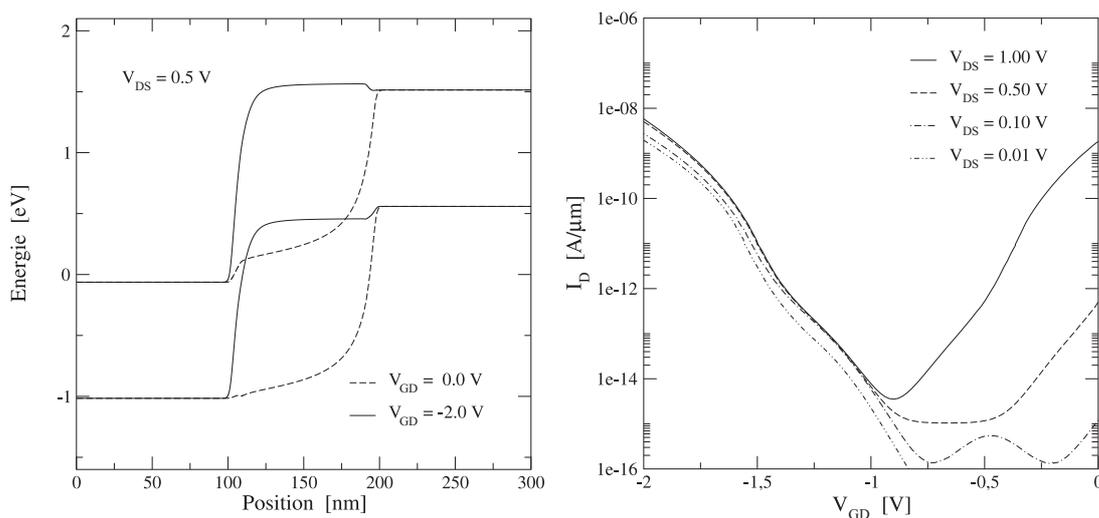


Abb. 3.10: Banddiagramm des Tunneltransistors im p-Betrieb (links) und resultierende Transfercharakteristik (rechts). Hierbei wurde das gleiche Bauelement wie in Abbildung 3.7 untersucht.

erforderlich, um einen p-Kanal Tunneltransistor zu erzeugen. Hierfür wird lediglich an Source- und Gatekontakt eine negative Spannung angelegt und der Drainkontakt stellt nun den Referenznullpunkt dar. In Abbildung 3.10 ist das Banddiagramm des p-TFET für verschiedene Gatespannungen gezeigt. Aus diesem ist die Bandverbiegung und die daraus folgende Reduktion der Tunnelbarrieren zu erkennen. Bereits aus dieser Darstellung ist abzusehen, was in der Transfercharakteristik deutlich zutage tritt: Bei kleinen Spannungen zwischen Gate und Drain ist die Tunnelwegstrecke zwischen Valenz- und Leitungsband nicht besonders groß. Daher ist das parasitäre Tunneln auf der Sourceseite sehr ausgeprägt und führt zu einem hohen Strom. Auf welche Weise dieser Leckstrom vermieden werden

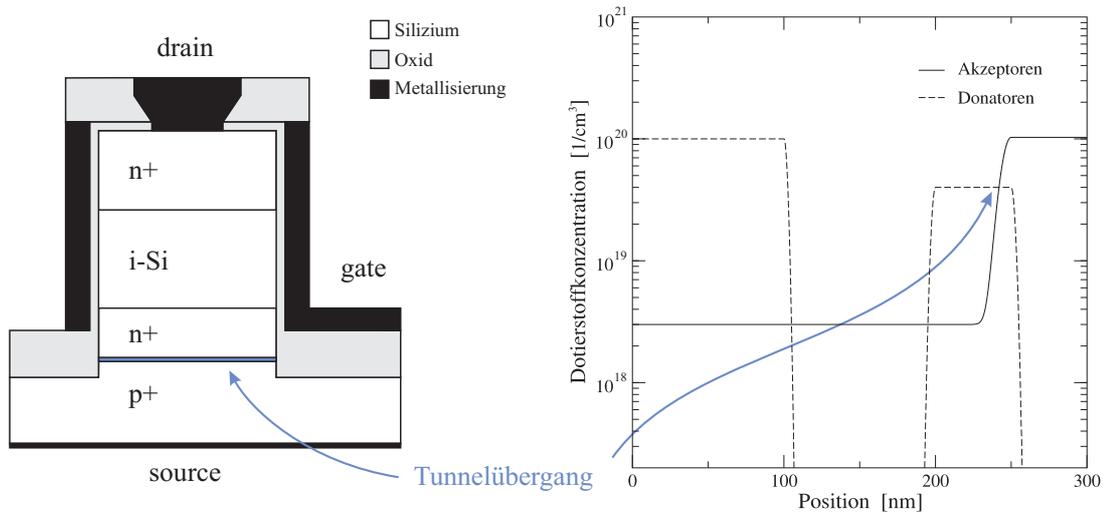


Abb. 3.11: Schematischer Aufbau eines Zener-Feldeffekttransistors (ZFET) mit 100 nm Gatelänge in vertikaler Realisierung (links) und Dotierstoffkonzentrationen in einem Schnitt 1 nm unter dem Gateoxid (rechts). Das Gate liegt zwischen den Punkten bei 100 und 200 nm.

kann, wird im nächsten Kapitel besprochen. Entscheidend ist jedoch, dass bei höheren Spannungsdifferenzen zwischen Gate und Drain ein eindeutiges Schaltverhalten vorliegt und somit p-TFETs realisierbar sind.

Insgesamt bleibt festzuhalten, dass die Simulationen des Tunneltransistors die erwarteten guten Schalteigenschaften bestätigen konnten. Besonders vielversprechend sind hierbei die sehr niedrigen Sperrströme und die große Steigung der Transferkennlinien. Vor dem Hintergrund, dass die in diesem Abschnitt untersuchte Tunneltransistorgeometrie und die dazugehörigen Dotierprofile denkbar einfach sind, kann zudem eine weitere erhebliche Verbesserung der Bauelementeigenschaften erwartet werden. Eine positive Eigenschaft ist zudem die unproblematische Herstellung von Bauelementen im n- und p-Betrieb.

3.1.4 Zener-Feldeffekttransistor

Ein letztes in diesem Kapitel betrachtetes Bauelement ist der Zener-Feldeffekttransistor (ZFET). Das Ausgangsbaulement ist ein MOSFET, der auf der Source- und Drainseite aus zwei unterschiedlich tiefen Dotierwannen besteht (siehe auch Abschnitt 3.1.2). Wird nun auf der Sourceseite die tiefere Wanne nicht mehr arsen- sondern bordotiert, folgt ein Zener-Feldeffekttransistor. Für die Simulation wird eine vertikale Realisierung dieses Bauelements untersucht: In Abbildung 3.11 ist der schematische Aufbau des ZFETs und das Dotierprofil skizziert. Der Zener-Feldeffekttransistor hat 100 nm Gatelänge und 2 nm Gateoxiddicke. Der Gatestapel besteht wie beim Tunneltransistor aus Siliziumdioxid und Aluminium als Kontaktmaterial. Das Substrat ist auf der Sourceseite 10^{20} cm^{-3} bordotiert und auf der Drainseite 10^{20} cm^{-3} arsen-dotiert. Der Kanal ist mit $3 \cdot 10^{18} \text{ cm}^{-3}$ nur schwach bordotiert. Allerdings ist die Sourcewanne 50 nm vom Gatestapel entfernt und zwischen

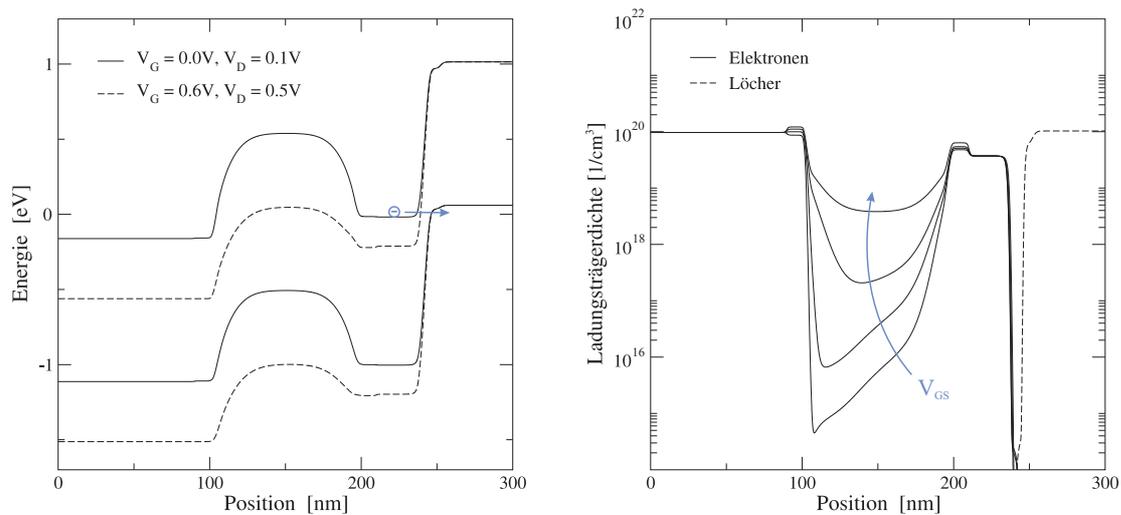


Abb. 3.12: Banddiagramm des ZFETs bei verschiedenen angelegten Spannungen (links) und Elektronen- und Löcherdichteverteilungen in einem Schnitt 1 nm unter dem Gateoxid bei der Variation der Gatespannung (rechts).

beiden befindet sich ein mit $4 \cdot 10^{19} \text{cm}^{-3}$ arsendotiertes Gebiet. Der Tunnelübergang hat damit circa 35 nm Abstand zu dem Gatestack. Der Gradient der Dotierprofile beträgt mit Ausnahme der Source-seite jeweils 2 nm/dec. Die Source-dotierung weist einen Gradienten von 8 nm/dec auf.

In Abbildung 3.12 ist das resultierende Banddiagramm des Zener-Feldeffekttransistors bei unterschiedlichen Betriebszuständen dargestellt. Zusätzlich sind die Elektronen- und Löcherdichten bei verschiedenen Gatespannungen abgebildet. Es zeigt sich, dass an dem Tunnelübergang nur geringe Auswirkungen des Gatekontaktes zu beobachten sind. Die Elektronendichte weist den für MOSFETs typischen Verlauf in der Kanalregion bei einer Variation der Gatespannung auf. Die simulierten Transfer- und die Ausgangskennlinien des ZFETs sind schließlich in Abbildung 3.13 dargestellt. Besonders ist der dreistufige Verlauf der Transferkennlinien hervorzuheben: Im Fall niedriger Gatespannungen baut sich ein Elektronenkanal unter dem Gate auf und das Einschalten des Bauelements kann beobachtet werden. Der Großteil der angelegten Source-Drain-Spannung fällt hierbei über der Kanalregion ab. Wird die Gatespannung weiter erhöht, leitet der Kanal immer besser und an der Zenerdiode fällt mehr Spannung ab. Damit steigt der Strom durch den ZFET weiter an. Ab einer Grenzspannung wird der Strom dann nur noch durch die Zenerdiode begrenzt und es folgt das deutlich sichtbare Sättigungsverhalten [45, 46].

3.2 Modellkalibrierung und Parametervariation

Für eine verlässliche Analyse und die anschließende Optimierung von Bauelementen ist es erforderlich die eingesetzten Modellparameter an die jeweiligen Fertigungsprozesse zu kalibrieren. Dazu werden allerdings nicht nur Messungen der Kennlinien, sondern auch

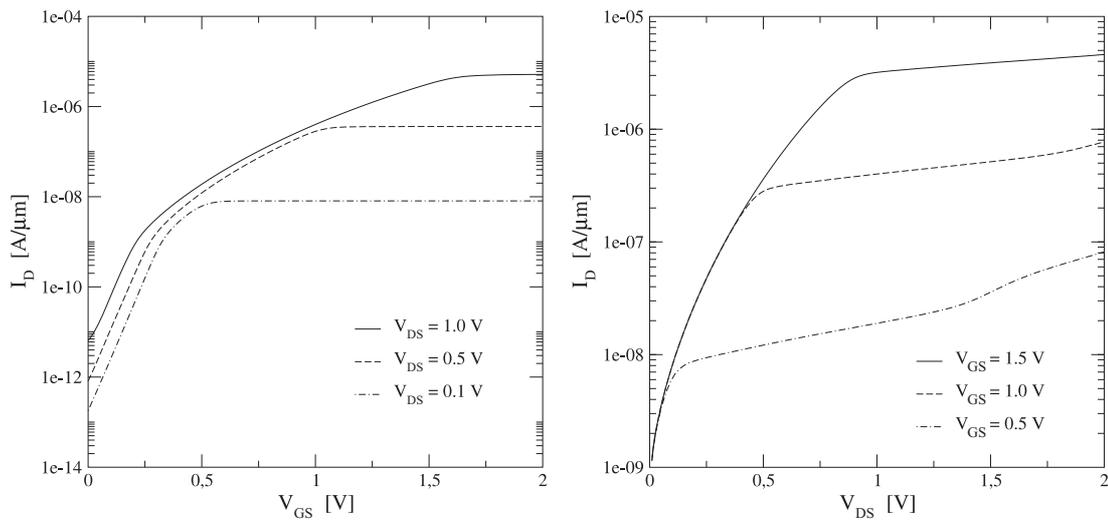


Abb. 3.13: Simulierte Transfer- (links) und Ausgangskennlinien (rechts) des untersuchten Zener-Feldeffekttransistors.

möglichst genaue Informationen über die Dotierprofile und die Bauelementabmessungen benötigt. Diese Daten sollten hierbei nicht ausschließlich auf reinen Prozesssimulationen basieren. Vielfach ist es allerdings von den einzelnen Halbleiterbauelementherstellern aus Wettbewerbsgründen nicht gewünscht, diese Daten der Öffentlichkeit zur Verfügung zu stellen. Aus diesen Gründen werden in diesem Abschnitt nur allgemeine Hinweise zur Kalibrierung der Modelle gegeben. Von zentraler Bedeutung ist dabei die Kenntnis der Einflüsse der Modellparameter auf die Bauelementcharakteristiken. Bei der Einführung in das Modell zur Beschreibung von phononenunterstützten Band-zu-Band Tunnelprozessen in Abschnitt 2.4 wurde bereits kurz die Notwendigkeit einer lokalen Dichtekorrektur der Generationsraten im Hinblick auf die Elektronen- und Löcherdichte thematisiert. Wird der im Abschnitt 3.1.3 betrachtete Tunneltransistor erneut untersucht, so ergibt sich die in Abbildung 3.14 gezeigte Korrektur der Tunnelraten und der Transfercharakteristiken. Die zu extrem ausgeprägte Sättigung der Generationsraten kann auf diese Weise wirkungsvoll vermieden werden. Eine Anpassung der Messresultate an die Simulationsergebnisse kann nun durch eine Dämpfung der lokalen Dichtekorrektur erzielt werden.

3.2.1 Parameter des Band-zu-Band Tunnelmodells

In Gleichung (2.68) finden sich mit dem Vorfaktor B und der kritischen Feldstärke F_c die zwei wesentlichen Kalibrierungsparameter des Tunnelmodells. Da der Vorfaktor einigen Näherungen unterworfen ist, kann durch eine Variation von B das aus den Simulationen berechnete Bauelementverhalten an die Messungen angepasst werden. Auf diese Weise ist es möglich die Ströme in der Transferkennlinie zu höheren oder niedrigeren Werten zu verschieben. Dieser Effekt kann in Abbildung 3.15 deutlich erkannt werden. So ist die lineare Abhängigkeit der Tunnelrate von B ebenso im Drainstrom sichtbar. Die kritische Feldstärke hat im Vergleich dazu einen noch größeren Einfluss. Eine Änderung führt zu

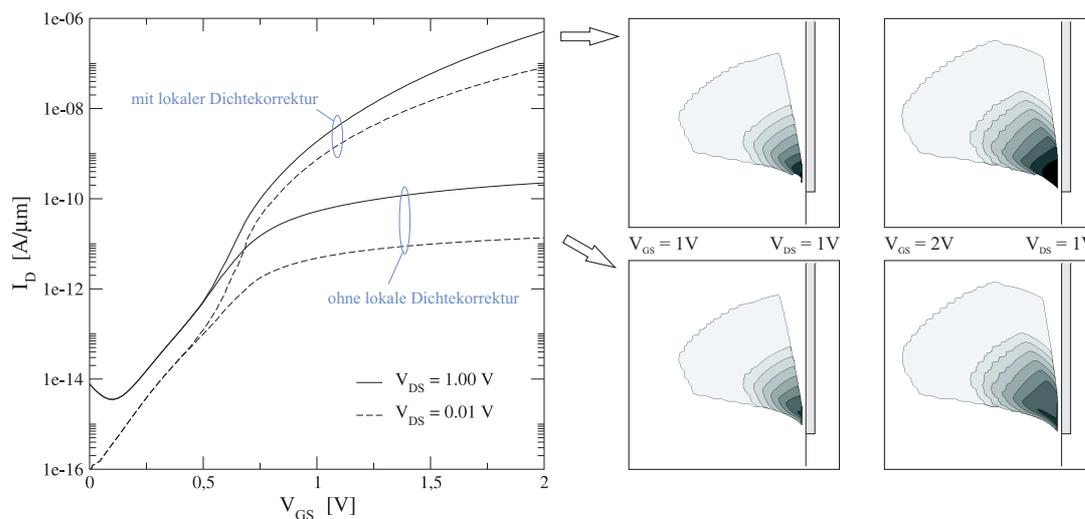


Abb. 3.14: Einfluss der lokalen Dichtekorrektur auf die simulierte Transfercharakteristik eines TFETs (links) und dazugehörige Generationsraten (rechts).

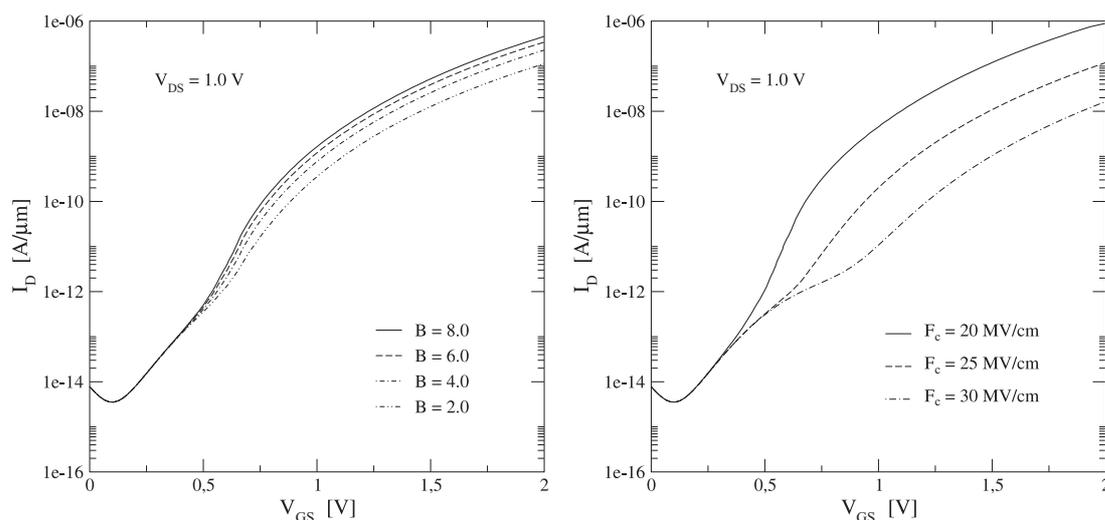


Abb. 3.15: Einfluss des Vorfaktors B und der kritischen elektrischen Feldstärke F_c aus Gleichung (2.68) auf die Transferkennlinien eines Tunneltransistors.

einer Verschiebung der Einsatzspannung des phononenunterstützten Tunnelstroms und damit zu erheblichen Auswirkungen auf die Kennlinie. Zusätzlich nimmt die Steilheit der Transferkennlinie des untersuchten Tunneltransistors zu. Im Fall beider Modellparameter kann eine Variation durch die Abhängigkeit von den, zur Approximation der Bandverläufe eingeführten, effektiven Massen und die damit verbundene Unsicherheit in physikalisch sinnvollen Werten begründet werden.

Da das Tunneln der Elektronen entlang der elektrischen Feldlinien stattfindet und für die Simulationen ein lokales Beschreibungsmodell eingesetzt wird, kann es vorkommen, dass der Start- oder der Endpunkt der Tunnelstrecke im Gateoxid liegt und damit physikalisch

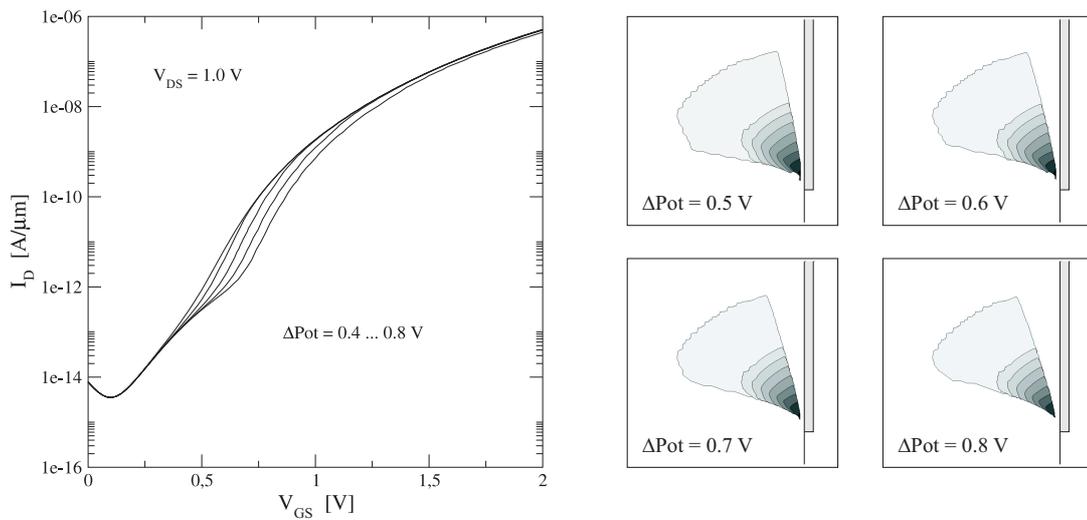


Abb. 3.16: Auswirkung der Variation der Potentialdifferenz ΔPot auf die simulierten Transferkennlinien eines Tunneltransistors und dazugehörige Generationsrate des phononenunterstützten Band-zu-Band Tunnelns.

kein Tunneln auftreten kann. Dennoch wird das Modell zu einer Generationsrate in der Mitte der Bandlücke führen und damit einen unphysikalischen Strom hervorrufen. Um diesen Effekt zu vermeiden, wird daher angenommen, dass das elektrische Potential in Feldrichtung um einen minimalen Betrag abfallen muss. Da eine Suchrichtung für die Potentialänderung vorgegeben und nicht den realen Feldlinien gefolgt wird, tritt jedoch letztlich eine gewisse Unsicherheit hinsichtlich der richtigen Wahl des Parameterwerts auf. In Abbildung 3.16 sind die Transferkennlinien und die Raten bei unterschiedlichen Potentialdifferenzen ΔPot dargestellt. Es ist hierbei zu erkennen, dass die vornehmliche Auswirkung auf die Kennlinie der Einsatzpunkt des Tunnelstroms ist und zusätzlich auch die Steigung zunimmt. Der maximale Tunnelstrom bei einer Gatespannung von 2 V bleibt dagegen nahezu unverändert. Auf der rechten Seite der Abbildung ist die entsprechende Band-zu-Band Tunnelgenerationsrate dargestellt. Aus den Abbildungen geht hervor, dass in einem etwa keilförmigen Gebiet unter dem Gateapfel kein Tunneln stattfindet. Der Öffnungswinkel ist dabei linear von dem Parameter ΔPot abhängig. Den Grund hierfür stellt der näherungsweise kreisförmige Verlauf der elektrischen Feldstärke dar.

3.2.2 Kritische Betrachtung der Stoßionisation in Tunneltransistoren

In Abschnitt 2.3.3 wurde mit Gleichung (2.39) ein Generationsterm zur Beschreibung der Stoßionisation eingeführt. Die jeweiligen Ionisationskoeffizienten für die Elektronen und Löcher des Modells nach van Overstraeten und de Man [36] sind allerdings für Dioden mit sehr weiten Raumladungszonen bestimmt worden. Die in dieser Arbeit untersuchten Tunneltransistoren weisen zumeist sehr kurze Raumladungszonen auf und es ist deshalb erforderlich zu klären, ob das Modell für die Stoßionisation überhaupt eingesetzt werden

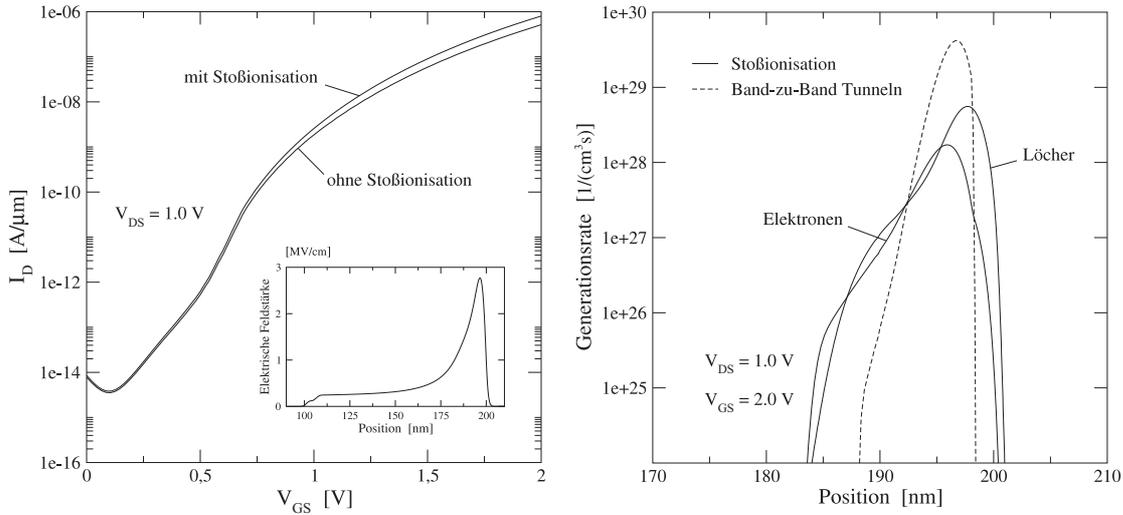


Abb. 3.17: Simulierte Transfercharakteristik des betrachteten Tunnelfeldeffekttransistors unter Einbeziehung der Stoßionisation (links) und Ionisationsraten in einem Schnitt 2 nm unter dem Gatestack (rechts).

kann. In einem ersten Schritt wird der Tunneltransistor aus dem vorangehenden Abschnitt einmal mit und einmal ohne Generationsterm für die Stoßionisation simuliert. Die hieraus folgenden Transferkennlinien sind in Abbildung 3.17 zu sehen: Durch die Einbindung der Stoßionisation resultiert ein signifikant höherer Strom im Durchlassbereich. Auf der rechten Seite der Abbildung sind die Raten der Stoßionisation und die Band-zu-Band Tunnelrate dargestellt. Da die Generationsmodelle lokale Modelle sind, ist die hohe Rate ursächliche Folge der hohen elektrischen Feldstärke am Tunnelübergang (eingebettetes Bild auf der linken Seite). Eine Bewertung dieser Ergebnisse ist mittels einer Integration über die Strompfade möglich. Hierbei muss näherungsweise gelten

$$q \int_{S_c} F(s) ds \geq E_{\text{ion},c} \quad \text{mit } c \in \{n, p\} \quad (3.1)$$

Dabei wird die Energiezunahme der Elektronen beziehungsweise Löcher im elektrischen Feld F mit der jeweiligen Ionisationsenergie $E_{\text{ion},c}$ verglichen. Die Integration erfolgt entlang der Stromdichten (Strompfade S_c). Grundsätzlich ist zu beachten, dass der auf diese Weise bestimmte Energiezuwachs der Ladungsträger bei der Bewegung durch das elektrische Feld wegen Elektron-Phonon-Stößen gedämpft wird, da die freie Weglänge der Elektron-Phonon-Wechselwirkung kleiner als die der Stoßionisation ist. Für eine erste Abschätzung ist diese Formel aber ausreichend. Typische Werte für die Ionisationsenergie sind nach Moll und van Overstraeten [47]: 1.8 eV (Elektronen) und 2.4 eV (Löcher). Aus numerischen Berechnungen, basierend auf realistischen Bandstrukturen, folgen ähnliche Resultate [48].

Angewandt auf das konkrete Beispiel aus Abbildung 3.17 müssen zwei Integrationen durchgeführt werden. Der Elektronenpfad verläuft ausgehend vom Drainkontakt zum Tunnelübergang. Dieser liegt etwa bei dem Maximum der elektrischen Feldstärke. Die

Löcher bewegen sich entsprechend zwischen Sourcekontakt und Tunnelübergang. Im Fall der Elektronen ist die Bedingung (3.1) erfüllt. Wegen der kleineren Raumladungszone wird die Bedingung bei den Löchern allerdings verletzt. Daher wird die Generationsrate der Löcher und somit der Strom überschätzt. Dieser Sachverhalt ist problematisch, da auf diese Weise auch zusätzlich die simulierte Tunnelrate wegen der höheren Ladungsdichten übermäßig ansteigt und zu einer weiteren Steigerung der Ströme führt. In den folgenden Analysen wird die Stoßionisation vernachlässigt, da vornehmlich Tunneltransistoren mit erheblich kürzeren Gatelängen untersucht werden. Bei solchen Kurzkanal-TFETs sind die Raumladungszonen der Elektronen und Löcher sehr klein und damit wird das Modell der Stoßionisation immer zu unphysikalisch hohen Generationsraten führen.

4. Optimierung von Tunnelbauelementen

Nachdem im vorangehenden Kapitel der grundlegende Aufbau, das Funktionsprinzip und die zentralen Eigenschaften von Tunnelbauelementen behandelt wurden, steht in diesem Kapitel die Optimierung von Tunneltransistoren im Vordergrund. Da die fortschreitende Reduzierung der Abmessungen von Halbleiterbauelementen außerdem zu einer verstärkt dreidimensionalen Strukturierung führt, werden nachfolgend vornehmlich zylindrische Nanodraht-Tunneltransistoren untersucht. Eine solche Beschränkung wird gewählt, da die wesentlichen Eigenschaften und Effekte auch mit zylindrischen Bauelementen analysiert werden können und auf diese Weise zusätzlich der Rechenaufwand signifikant kleiner gehalten werden kann.

Im ersten Abschnitt werden die Effekte ausführlich beschrieben und erklärt, die aus der quantenmechanischen Korrektur der Simulationen resultieren. Mit Hilfe der im zweiten Kapitel eingeführten Dichtegradientenmethode wird diese Korrektur vorgenommen. Als gemeinsame Basis der durchgeführten Simulationen dient das Drift-Diffusions-Modell und alle in diesem Kapitel vorgestellten Resultate wurden mittels Sentaurus Device [38] gewonnen. Im zweiten Abschnitt werden die jeweiligen Auswirkungen einer Variation der Dotierprofile des Tunneltransistors behandelt. Hierzu werden die Konzentrationen und die Gradienten der Profile in der Kanalregion des Tunneltransistors verändert und das daraus folgende Bauelementverhalten analysiert. Zusätzlich wird der Einfluss der Position der Sourcereion in Relation zum Gate untersucht. Da die Minimierung der Abmessungen ein entscheidendes Kriterium für die Eignung von Tunneltransistoren darstellt, ist die Analyse der Einflüsse einer Verringerung der Gatelänge und des Durchmessers der Nanodrähte von zentraler Bedeutung. Anschließend werden unterschiedliche Materialien im Kanalgebiet und im Gatestapel betrachtet. So kommen high-k Materialien als Gateoxid, Metalle als Gatekontakte und SiGe im Kanal zum Einsatz. Die hieraus gewonnenen Resultate werden im fünften Abschnitt zusammengefasst und zudem wird ein optimiertes Bauelement mit 5 nm Gatelänge vorgestellt. Den Abschluss bildet dann die Analyse von dreidimensional strukturierten Tunneltransistoren. In Abbildung 4.1 sind die einzelnen Untersuchungen bei zylindrischen Nanodraht-Tunneltransistoren schematisch zusammengefasst.

4.1 Einfluss von Quanteneffekten auf die TFET-Charakteristiken

Damit die Vergleichbarkeit der einzelnen Simulationsergebnisse gewährleistet ist, wird für alle in diesem Kapitel vorgestellten Resultate ein gemeinsames Referenzbauelement

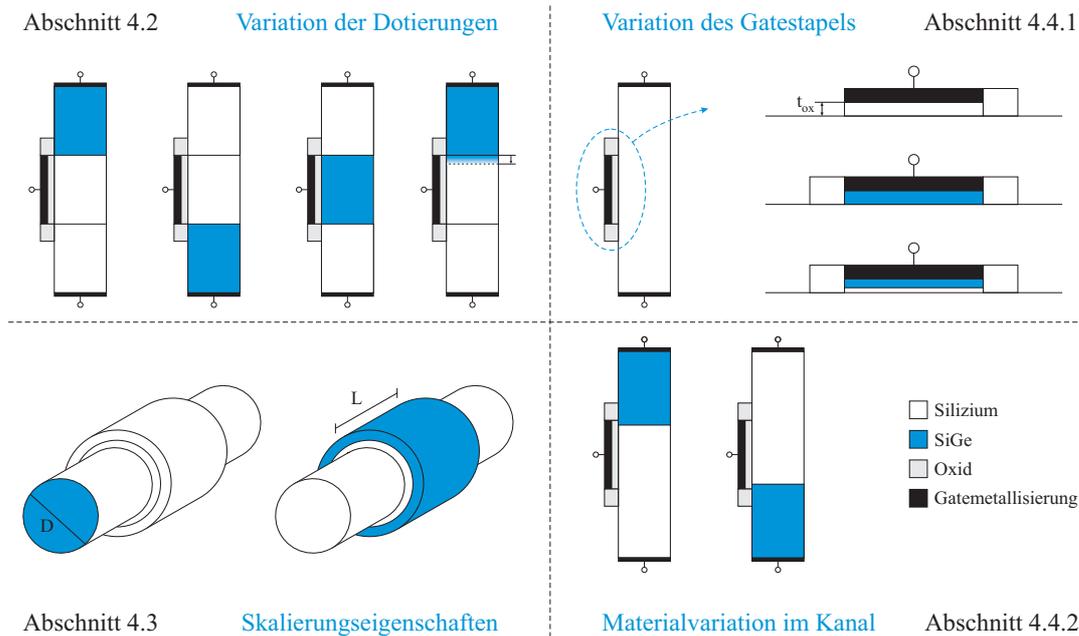


Abb. 4.1: Schematische Darstellung der in diesem Kapitel untersuchten Variationen von zylindrischen Nanodraht-Tunneltransistoren.

eingesetzt. Dieses ist in Abbildung 4.2 schematisch dargestellt und besteht aus einem zylindrischen Nanodraht-Tunneltransistor, der durch Ausnutzen der Rotationssymmetrie zweidimensional simuliert werden kann. Die Gatelänge beträgt 40 nm und das Gateoxid ist 2 nm dick. Der Gatestapel besteht hierbei aus Siliziumdioxid und als Kontaktmaterial wird TiN eingesetzt. An den Seiten des Gatestapels sind Seitenoxide eingebaut, damit die elektrischen Streufelder und die aus diesen resultierenden Effekte sinnvoll berücksichtigt werden können. Das Siliziumsubstrat ist dabei auf der Sourceseite 10^{20}cm^{-3} bordotiert und auf der Drainseite 10^{20}cm^{-3} arsen-dotiert. Der Kanal ist mit 10^{16}cm^{-3} nur schwach bordotiert und der Gradient der Dotierprofile beträgt auf der Source- und der Drainseite jeweils 2 nm/dec und folgt hierbei jeweils einem Gaußprofil. Zudem beträgt der Abstand der Dotierprofilmaxima zum Gatestapel 5 nm. Dieser sogenannte Spacer dient zum Schutz des Gatestapels vor Defekten bei der Implantierung der Source- und Draindotierungen. In den meisten Fällen wird ein Durchmesser des Nanodrahts von 30 nm gewählt. Dabei sind der Source- und der Drainkontakt an den Stirnseiten des Zylinders angebracht. Mögliche Abweichungen von den hier angegebenen Werten werden im Einzelfall in den jeweiligen Abschnitten beschrieben.

Auf der rechten Seite von Abbildung 4.2 ist das simulierte Orts-Energie Banddiagramm bei unterschiedlichen Gatespannungen gezeigt. Bereits aus diesem Bild kann das gute Schaltverhalten des Tunneltransistors herausgelesen werden: So ist die Barrierenbreite für das indirekte Tunneln vom Valenz- zum Leitungsband bei hohen Gatespannungen nur noch gering, wodurch eine hohe Tunnelwahrscheinlichkeit und damit auch ein großer Strom folgt. Ein weiterer Effekt, der aus diesem Banddiagramm herausgelesen werden kann, ist die Verringerung der Bandlücke durch hohe Dotierstoffkonzentrationen. Aus den

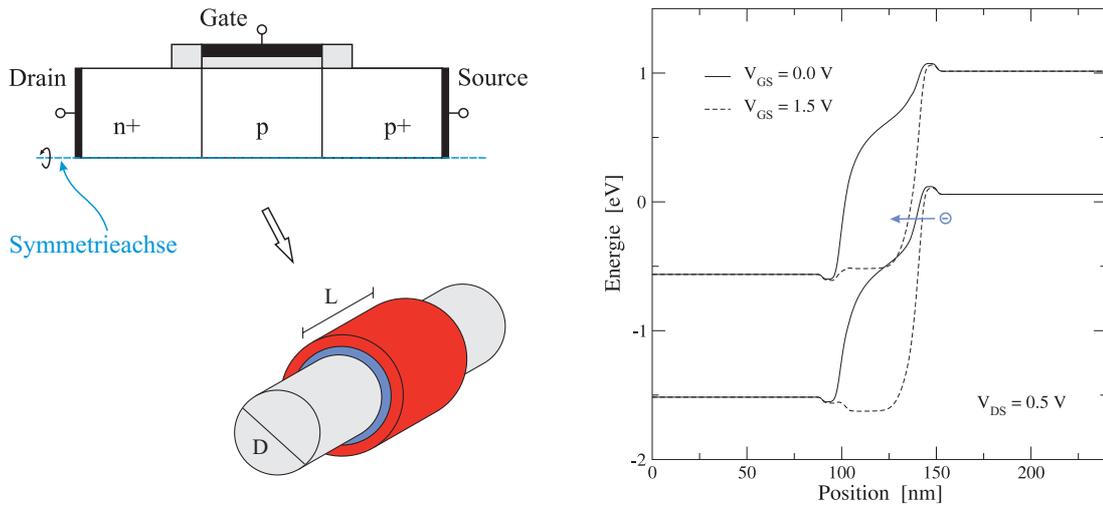


Abb. 4.2: Grundlegender Aufbau zylindrischer Nanodraht-Tunnelfeldeffekttransistoren und Bandverlauf eines TFETs mit 30 nm Durchmesser und 40 nm Gatelänge.

Modellgleichungen in Abschnitt 2.4.2 folgt, dass die Wahrscheinlichkeit für das indirekte Band-zu-Band Tunneln mit kleinerer Bandlücke zunimmt. Reduziert sich demnach die Bandlücke aufgrund von hohen Dotierkonzentrationen, so ergibt sich einerseits durch die größere Anzahl an Ladungsträgern und andererseits durch die verringerte Bandlücke eine signifikante Steigerung der Ströme zwischen Source- und Drainkontakt.

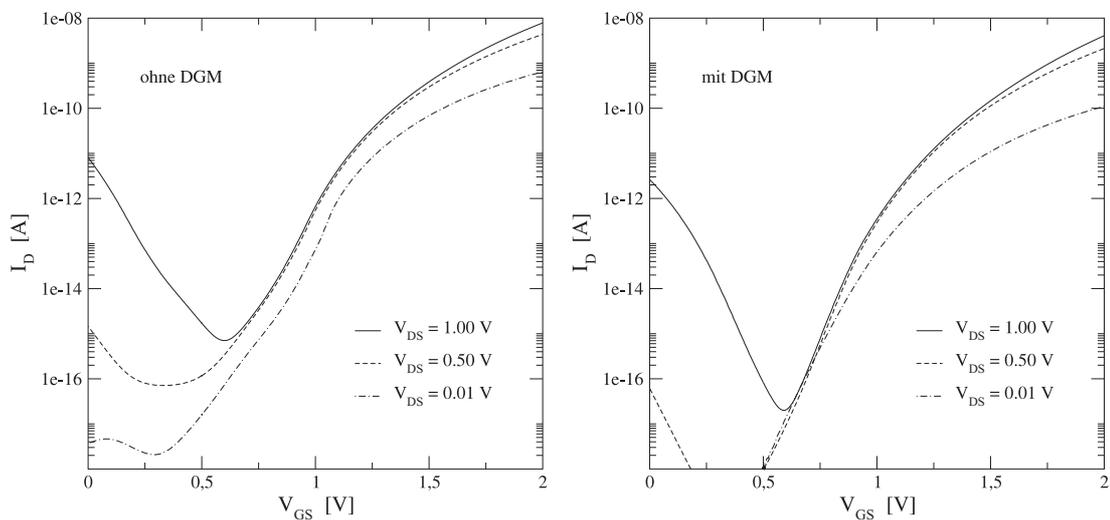


Abb. 4.3: Simulierte Transfercharakteristik eines zylindrischen Tunneltransistors mit 50 nm Durchmesser und 40 nm Gatelänge sowohl ohne (links), als auch mit quantenmechanischer Dichtekorrektur (rechts).

Die Simulation eines zylindrischen Nanodraht-Tunneltransistors mit 50 nm Durchmesser und 40 nm Gatelänge liefert die in Abbildung 4.3 gezeigten Transferkennlinien. Hierzu ist einerseits eine semiklassische und andererseits eine, mit der in Abschnitt 2.5 eingeführten

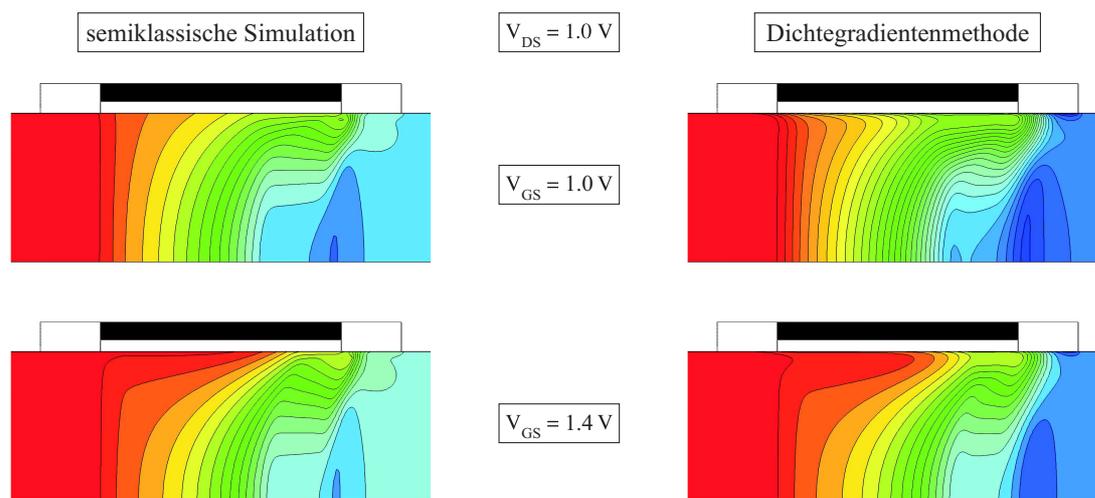


Abb. 4.4: Vergleich der Elektronendichteverteilung bei zwei unterschiedlich angelegten Gatespannungen unter Vernachlässigung (links) und unter Berücksichtigung der quantenmechanischen Dichtekorrektur (rechts).

Dichtegradientenmethode, quantenmechanisch-korrigierte Simulationsreihe durchgeführt worden. Da es sich bei dem untersuchten Tunneltransistor um ein bipolares Bauelement handelt, erfordert die Verwendung der Dichtegradientenmethode jeweils eine zusätzliche mit den Ladungsträgerdichten und dem elektrostatischen Potential gekoppelte partielle Differentialgleichung für die Elektronen und die Löcher. Wegen der daraus folgenden deutlichen Steigerung des rechentechnischen Aufwands ist es notwendig in einem ersten Schritt zu überprüfen, inwieweit überhaupt ein Effekt auf das simulierte Verhalten des Bauelements festzustellen ist. Hierbei zeigt sich, dass sowohl im Sperrbereich des Tunneltransistors, wie auch bei hohen Gatespannungen, erhebliche Unterschiede auszumachen sind. Des Weiteren ist auch ohne anliegende Gatespannung ein relativ großer Stromfluss durch das Bauelement zu beobachten. Dieser wird allerdings durch einen Band-zu-Band Tunnelstrom auf der Drainseite des Tunneltransistors hervorgerufen. Besonders prägnant ist der Vergleich der Sperrströme (d.h. bei minimalem Stromflusses zwischen Drain- und Sourcekontakt): So sinkt durch die quantenmechanische Korrektur der Ladungsdichten der Strom deutlich. Der Grund hierfür ist die Tatsache, dass der Strom vornehmlich durch Grenzflächenrekombination und defektunterstützte Tunnelprozesse getragen wird. Weil nun beide Effekte signifikant von der Elektronen- und Löcherdichte abhängen, ist eine Veränderung der Dichteverteilungen auch im Stromfluss zu beobachten. In Abbildung 4.4 ist beispielhaft die Elektronendichte mit und ohne quantenmechanische Dichtekorrektur für verschiedene Gatespannungen dargestellt. Hier ist die Reduktion der Elektronendichte direkt unter dem Gate zu erkennen. Außerdem hängt zudem die phononengestützte Band-zu-Band Tunnelrate von den Dichteverteilungen ab, da sowohl freie Elektronen zu Beginn des Tunnelpfades wie auch freie Stellen am Ende erforderlich sind. Daher ist auch die Veränderung der Ströme bei hohen Gatespannungen zu erklären [49].

Die Transferkennlinien für die beiden untersuchten Fälle sind in Abbildung 4.5 (links) in einem gemeinsamen Bild gezeigt. Hierbei ist die generelle Reduktion der Ströme zu

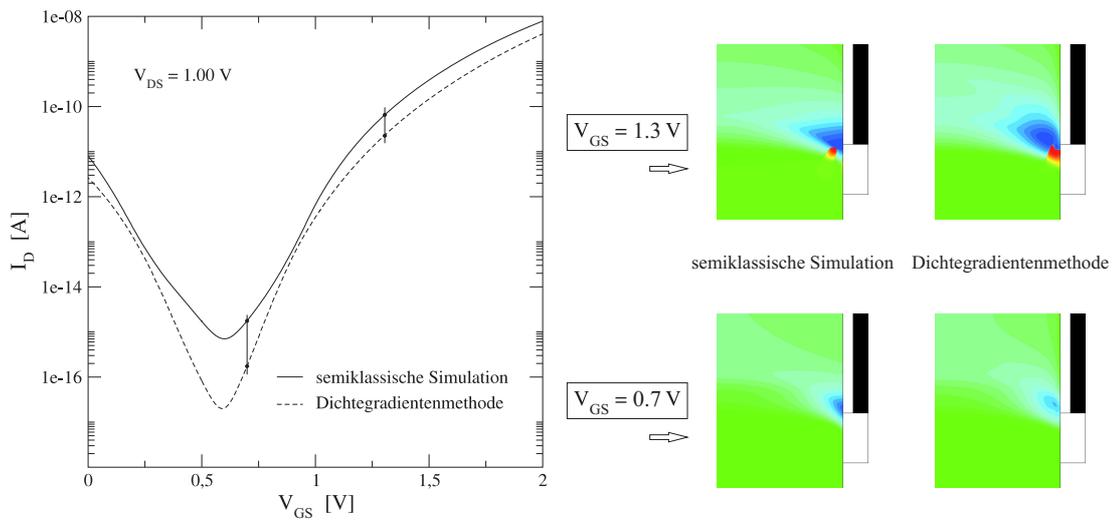


Abb. 4.5: Vergleich der Transferkennlinien (links) und der Shockley-Read-Hall-Raten (rechts) des untersuchten zylindrischen Nanodraht-Tunneltransistors im Fall der semiklassischen Simulation und im Fall der Dichtegradientenmethode.

bemerkten. Auf der rechten Seite der Abbildung sind zusätzlich die Shockley-Read-Hall-Rekombinationsraten bei ausgewählten Gatespannungen dargestellt. Zusammenfassend bleibt festzuhalten, dass für eine sinnvolle Simulation, und damit auch die aussagekräftige Bewertung von Tunnelfeldeffekttransistoren, Quantisierungseffekte unter dem Gateoxid nicht vernachlässigt werden können. Wegen der vergleichsweise einfachen Implementierbarkeit ist die Dichtegradientenmethode hervorragend für diesen Zweck geeignet. Auf die Grenzen dieses Verfahrens wird im Abschnitt 4.3 über die Skalierungseigenschaften von zylindrischen Tunneltransistoren noch genauer eingegangen.

4.2 Auswirkungen der Dotierprofile auf das Bauelementverhalten

Eine der wichtigsten Methoden das Verhalten von Halbleiterbauelementen zu optimieren ist die Dotierung in der Kanalregion. Um dieses Ziel zu erreichen, gibt es eine Vielzahl an Möglichkeiten. Allerdings ist zu beachten, dass nicht alle denkbaren Dotierprofile mit vertretbarem Aufwand zu fertigen sind und eine Abwägung zwischen dem Mehrwert einer komplexen Dotierung und den dabei entstehenden Kosten erforderlich ist.

In diesem Abschnitt werden zuerst die Auswirkungen der Dotierstoffkonzentrationen auf der Source- und der Drainseite untersucht. Anschließend wird die Dotierung unter dem Gatestack variert. Es werden sowohl bor- als auch arsen-dotierte Profile betrachtet. Ein weiterer, wichtiger Punkt ist der Gradient der Dotierprofile auf der Sourcesseite. Außerdem ist für eine sinnvolle Bewertung auch der Einfluss des Abstandes zwischen Gatestack und Sourcedotierstoffmaximum zu berücksichtigen.

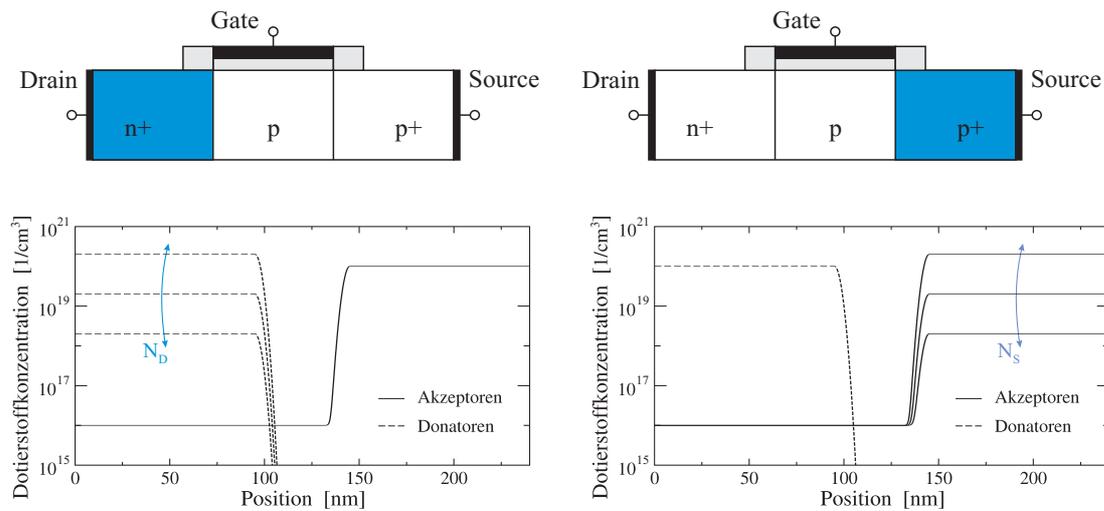


Abb. 4.6: Variation der Dotierstoffkonzentrationen auf der Source- und der Drainseite des Referenztunneltransistors mit 30 nm Durchmesser und 40 nm Gatelänge.

4.2.1 Drain- und Sourcekonzentration

Für alle Simulationen in diesem Abschnitt wird als Basis der im vorangehenden Abschnitt eingeführte zylindrische Referenztunnelfeldeffekttransistor mit einem Durchmesser von 30 nm und einer Gatelänge von 40 nm herangezogen. Eine schematische Darstellung der nachfolgend untersuchten Variationen ist in Abbildung 4.6 gegeben und zusätzlich sind beispielhaft einige Dotierprofile in einem Schnitt durch den Tunneltransistor entlang der Symmetrieachse gezeigt. Es ist hierbei anzumerken, dass die Dotierstoffkonzentrationen eine Ortsabhängigkeit nur in Transportrichtung aufweisen.

Wird in einem ersten Schritt das Maximum der Arsenkonzentration auf der Drainseite des zylindrischen Tunneltransistors variiert, so ergeben sich die auf der linken Seite von Abbildung 4.7 gezeigten Auswirkungen auf die Transferkennlinien. Es kann beobachtet werden, dass vornehmlich der Sperrbereich der Charakteristiken beeinflusst wird. Auf der rechten Seite der Abbildung ist der Sperrstrom in doppeltlogarithmischer Darstellung in Abhängigkeit von der maximalen Draindotierung dargestellt. Deutlich zu erkennen ist hierbei der exponentielle Anstieg des Stromes mit ansteigender Arsenkonzentration. Bei größeren Werten setzt allerdings ein Sättigungseffekt ein. Um nachvollziehen zu können auf welche Weise dieses Verhalten zustande kommt, lohnt sich ein Blick auf den Verlauf der elektrischen Feldstärke. In Abbildung 4.8 ist der Betrag dieser in einem Schnitt 2 nm unter dem Gateoxid für drei unterschiedliche Arsenkonzentrationen gezeigt. Zusätzlich ist die Band-zu-Band Tunnelrate an dem Übergang von Drainregion zum Kanal entlang derselben Schnittlinie abgebildet. Die elektrische Feldstärke ist ab der Position bei 130 nm im Wesentlichen unverändert, wohingegen am Übergang von Drain- zur Kanalregion eine signifikante Änderung zu beobachten ist. Einerseits weist das Maximum geringere Werte auf und andererseits wandert dieses in die Drainregion. Diese beiden Effekte sind direkt miteinander verbunden, da bei einer geringeren Dotierung der Drainregion auch weniger Ladungsträger verfügbar sind, und deshalb die Raumladungszonen größer werden. Auf

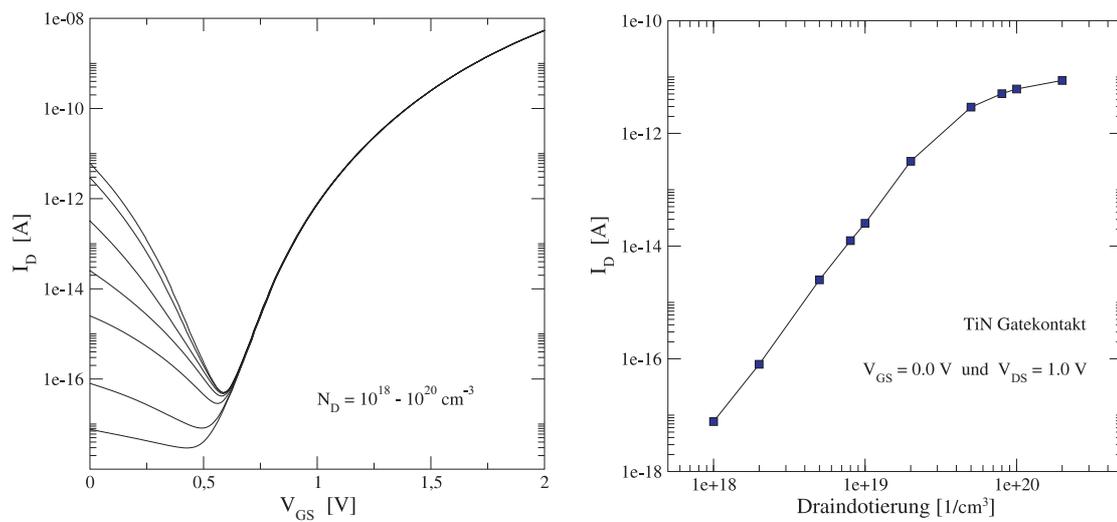


Abb. 4.7: Einfluss der Dotierstoffkonzentration der arsendotierten Drainseite auf die Transfercharakteristik (links) und den Sperrstrom des CNW-TFETs (rechts).

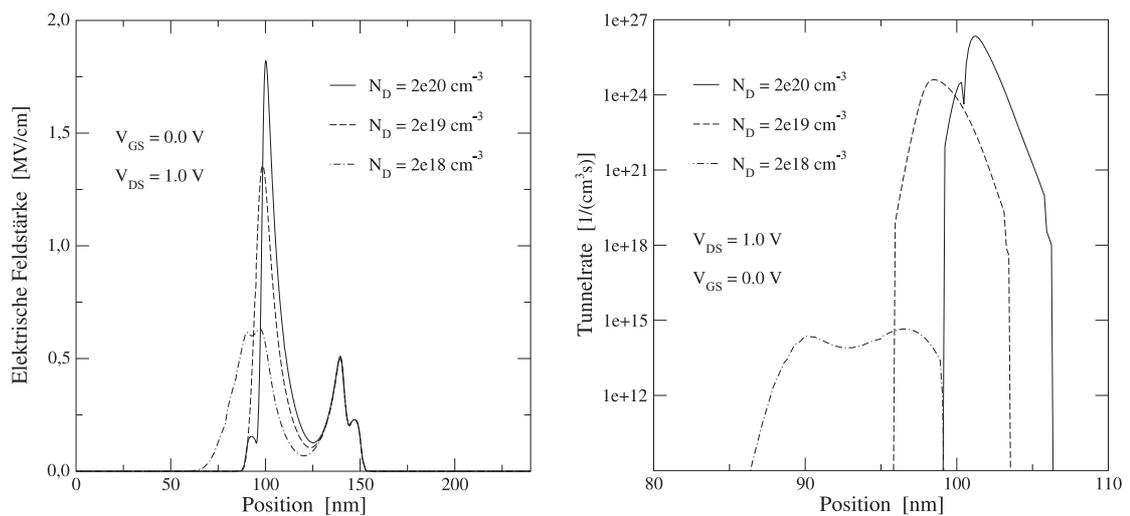


Abb. 4.8: Verlauf der elektrischen Feldstärke (links) und der Band-zu-Band Tunnelrate (rechts) bei verschiedenen Dotierstoffkonzentrationen auf der Drainseite des Tunneltransistors entlang einer Schnittlinie 2 nm unter dem Gateoxid.

diese Weise resultieren weniger steile Feldstärkeverläufe und es steigt die Feldstärke unter dem Gatestapel auf weniger hohe Werte an. Da die Band-zu-Band Tunnelrate wesentlich von der elektrischen Feldstärke abhängt und zudem ein leichtes Sättigungsverhalten mit steigender Feldstärke aufweist (siehe Abbildung 2.5), ergeben sich die auf der rechten Seite von Abbildung 4.8 dargestellten Generationsraten und mit diesen ist letztlich der in Abbildung 4.7 beobachtete Sättigungseffekt der Drainströme erklärbar.

Eine Variation der Borkonzentration auf der Sourceseite des Tunneltransistors resultiert in den in Abbildung 4.9 links gezeigten Transfercharakteristiken. Hierbei ist prinzipiell ein

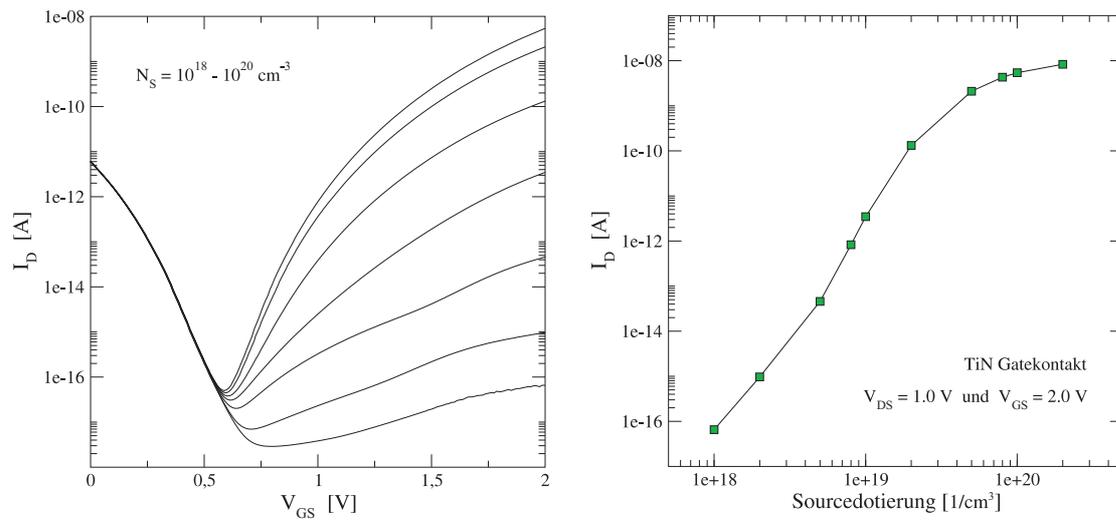


Abb. 4.9: Einfluss der Dotierstoffkonzentration der bordotierten Sourceseite auf die Transfercharakteristik (links) und den On-Strom des CNW-TFETs (rechts).

äquivalentes Verhalten der Drainströme im Vergleich zur Variation der Draindotierungen zu beobachten. Allerdings ist der Strom im Durchlassbetrieb betroffen. So fällt besonders bei schwachen Dotierungen der Anteil des defektunterstützten Tunnels am Gesamtstrom auf: Dieser ist bei einer Sourcedotierstoffkonzentration von $4 \cdot 10^{18} \text{ cm}^{-3}$ (d.h. der dritten Kennlinie von unten) sehr gut zu erkennen. Bei mittleren Gatespannungen dominiert das defektunterstützte Tunneln und sättigt bei einer weiteren Erhöhung der Gatespannung. Im weiteren Verlauf der Kennlinien wird der entsprechende Anteil des phononenunterstützten Band-zu-Band Tunnelns größer. Werden die Ströme im Durchlassfall in Abhängigkeit von der Konzentration der Sourcedotierung aufgetragen, so ergibt sich ein Verhalten, das sehr große Ähnlichkeiten zum Kennlinienverlauf in Abbildung 4.7 rechts aufweist. Der Unterschied bei niedrigen Dotierungen ist einerseits durch die Bor-Grunddotierung im Kanal und andererseits durch die unterschiedlichen Differenzen der Gatespannungen im Vergleich zur Spannung bei minimalen Stromfluss zu erklären [50].

In Abbildung 4.10 ist der Betrag der elektrischen Feldstärke in einem Schnitt 2 nm unter dem Gateoxid für drei unterschiedliche Bordotierstoffkonzentrationen der Sourcereion gezeigt. Zusätzlich ist die Band-zu-Band Tunnelrate an dem Übergang zwischen Kanal und Sourcereion entlang derselben Schnittlinie abgebildet. Hier zeigen sich die bereits besprochenen Effekte. Der Gatekontakt ist zwischen den Positionen bei 100 und 140 nm angebracht. Damit entfernt sich das Maximum der Tunnelrate im Fall einer sinkenden Sourcedotierstoffkonzentration von dem Gatestapel und die elektrische Feldstärke wird in der Tunnelregion vornehmlich durch Streufelder durch das Spaceroxid beeinflusst. Auf die Auswirkungen unterschiedlicher Materialien des Spaceroxids wird in Abschnitt 4.4.1 genauer eingegangen. Für die in diesem Kapitel angestrebte Entwicklung von geeigneten Optimierungsstrategien zylindrischer Tunneltransistoren bleibt damit festzuhalten, dass die Sperrströme durch niedrige Draindotierungen reduziert werden können. Ein Problem einer extremen Verringerung der Dotierungen auf der Drainseite liegt allerdings in dem

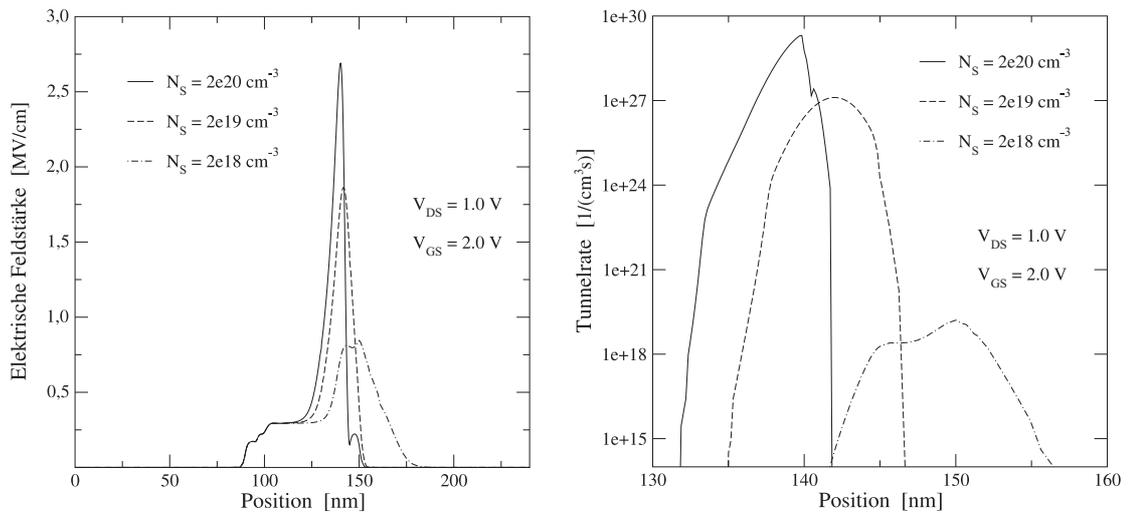


Abb. 4.10: Verlauf der elektrischen Feldstärke (links) und der Band-zu-Band Tunnelrate (rechts) bei verschiedenen Dotierstoffkonzentrationen auf der Sourceseite des Tunneltransistors entlang einer Schnittlinie 2 nm unter dem Gateoxid.

Anstieg der Widerstände beim Übergang zur Drainkontaktmetallisierung. Eine mögliche Lösung des Problems liegt in der stufenweisen Erhöhung der Draindotierung. So kann durch einen größeren Abstand des hochdotierten Drainbereichs zum Gatestack ein niedriger Kontaktwiderstand, andererseits eine Verringerung des Sperrstroms erzielt werden. Wird zudem die Konzentration der Sourcedotierung möglichst hoch gewählt, so werden auch große Ströme im Durchlassbereich realisierbar.

4.2.2 Kanalkonzentration

Nachdem im letzten Teilabschnitt zu erkennen war, dass die Dotierung im Kanalgebiet auch Auswirkungen auf das Bauelementverhalten hat, werden diese nachfolgend detailliert untersucht. Zuerst wird die Konzentration der homogenen Bor-Grunddotierung des Tunneltransistors variiert. Damit bleibt der Kanal immer p-dotiert. Die resultierenden Transferkennlinien sind auf der linken Seite von Abbildung 4.11 dargestellt. Dabei sind unterschiedliche Effekte zu beobachten: Zuerst einmal verschiebt sich das Minimum des Stroms durch den Tunneltransistor, wenn die Dotierstoffkonzentration erhöht wird. Dies entspricht einer Verschiebung der Einsatzspannung. Zudem steigt der Sperrstrom an. Auf der rechten Seite der Abbildung ist dieser in Abhängigkeit von der Dotierkonzentration dargestellt. Hier zeigt sich der mehr als exponentielle Anstieg der Sperrströme.

Ein Grund für dieses Verhalten liegt in der Verringerung der Barrierenbreite zwischen Valenz- und Leitungsband am Übergang von der Drainregion zum Kanal. Auf Grund der höheren Bordotierung des Kanals wird der pn-Übergang ohne angelegte Gatespannung in Richtung des Drainkontakts wandern und damit steigt auch die Tunnelwahrscheinlichkeit an diesem Übergang. Zusätzlich wird eine höhere Gatespannung benötigt, um den Kanal unter dem Gate zu verarmen. In der Folge verschiebt sich die Einsatzspannung. Wegen

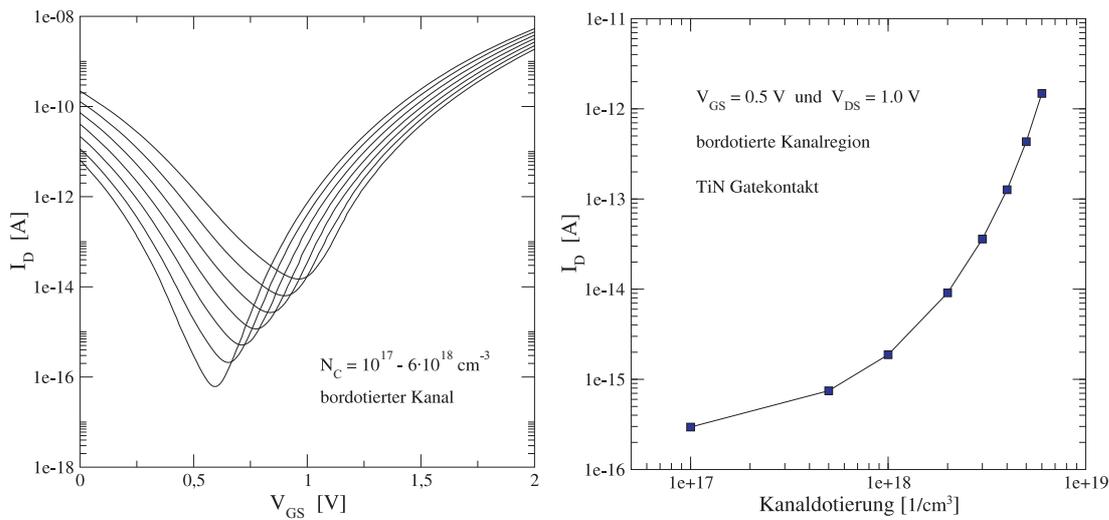


Abb. 4.11: Einfluss der Dotierstoffkonzentration einer bordotierten Kanalregion auf die Transfercharakteristik (links) und den Sperrstrom des CNW-TFETs (rechts).

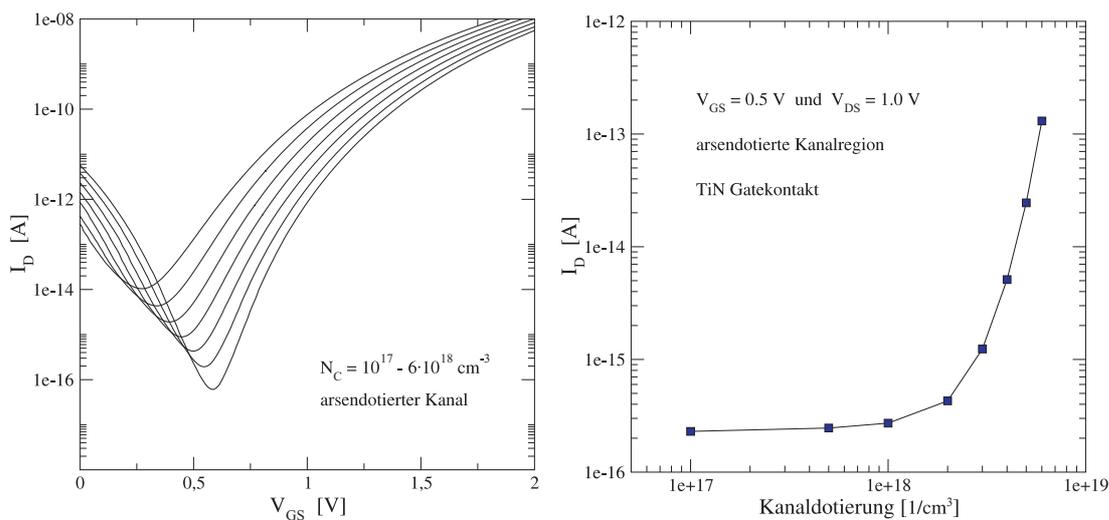


Abb. 4.12: Einfluss der Dotierstoffkonzentration eines arsendotierten Kanals auf die Transfercharakteristik (links) und den Sperrstrom des CNW-TFETs (rechts).

der höheren Kanaldotierung ist der Übergang zwischen Drain und Kanal abrupt und der zwischen Kanal und Source weniger abrupt. Daher nehmen auch die Sperrströme zu, da der unerwünschte Tunnelstromanteil auf der Drainseite mehr und mehr das Verhalten des Bauelements dominiert. Eine dritte, wichtige Erkenntnis ist die Invarianz der Steigung des Stromes in Bezug auf eine Variation der Bordotierung des Kanals.

Eine äquivalente Variation der Kanaldotierstoffkonzentrationen wurde auch für den Fall einer Arsendotierung durchgeführt. Hierbei resultieren die in Abbildung 4.12 gezeigten Abhängigkeiten der Transferkennlinien und der Sperrströme von der Arsenkonzentration in der Kanalregion. Prinzipiell gelten alle oben getroffenen Aussagen, wenn die Begriffe

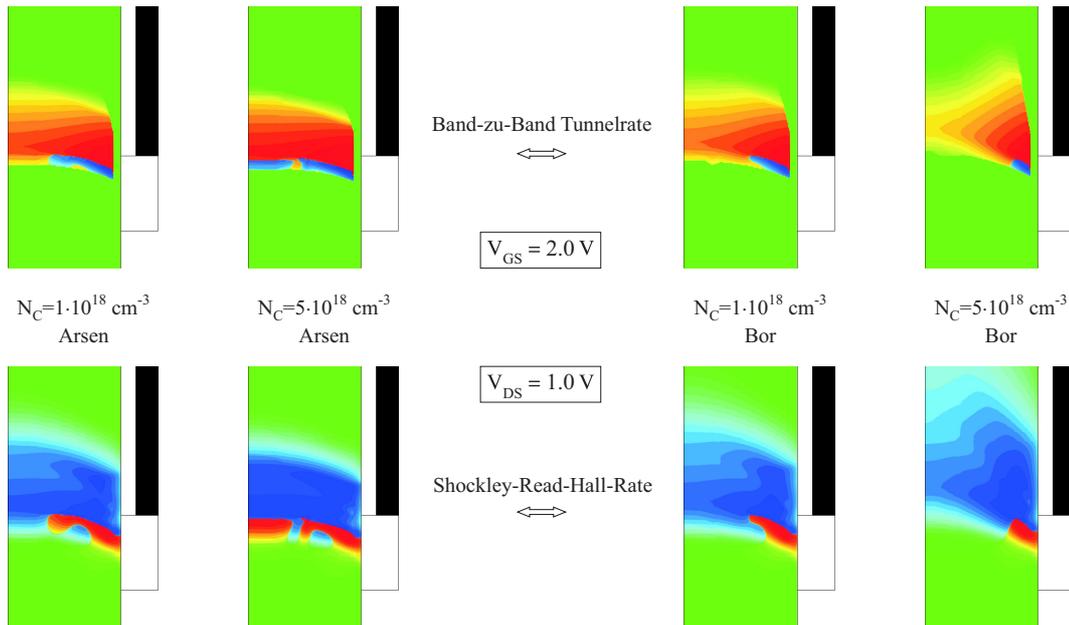


Abb. 4.13: Vergleich der Band-zu-Band Tunnelraten (oben) und der Shockley-Read-Hall-Rekombinationsraten (unten) bei unterschiedlichen Kanaldotierungen.

Source und Drain vertauscht werden. Die wichtigsten Auswirkungen sind die Zunahme der Ströme im Durchlassbereich und die starke Abnahme der Steigung der Ströme bei steigender Arsenkonzentration.

In Abbildung 4.13 sind die Band-zu-Band Tunnelraten wie auch die Shockley-Read-Hall-Rekombinationsraten im Durchlassfall für verschiedene Dotierungen der Kanalregion dargestellt. Mit deren Hilfe kann somit eine abschließende Bewertung der Variationen hinsichtlich der Optimierung des Bauelementverhaltens vorgenommen werden. Bei einer Arsendotierung sind die Tunnelraten stärker lokalisiert. Daraus folgt die Erwartung von grundsätzlich besseren Skalierungseigenschaften, da die Abhängigkeit von der Gatelänge geringer sein sollte. Daher scheint die Wahl auf eine hoch arsen-dotierte Kanalregion zu fallen, weil dann zusätzlich die Ströme im Durchlassfall höher als bei einer Bordotierung sind. Allerdings ist die Abnahme der Steigung ein kritischer Effekt, wodurch letztlich niedrige Bordotierungen empfehlenswert sind. Zudem kann durch kleine Variationen der Borkonzentration die Einsatzspannung optimiert werden, ohne die anderen Kenngrößen der Charakteristik signifikant zu beeinflussen. Hohe Ströme im Durchlassbereich sollten besser durch hohe Dotierstoffkonzentrationen der Sourceregion erzielt werden.

4.2.3 Variation der Spacerweite

In den vorangehenden Abschnitten wurde der Einfluss von Dotierstoffkonzentrationen untersucht. Hierbei konnte festgestellt werden, dass eine hohe Sourcedotierung für das Verhalten der Tunneltransistoren von Vorteil ist. Allerdings fehlen noch einige relevante

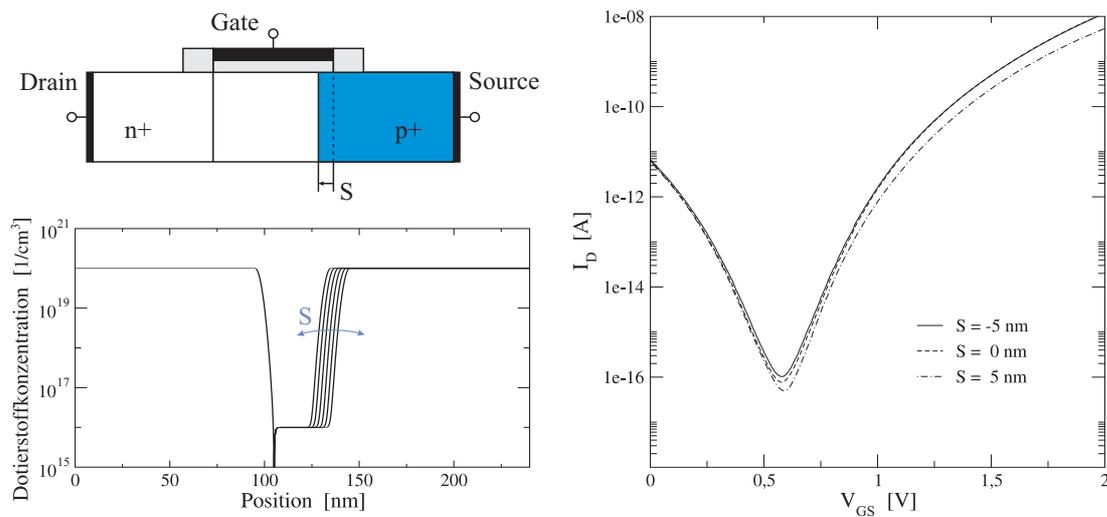


Abb. 4.14: Schematische Darstellung und Dotierprofil bei der Variation des Überlapps (negative Werte) beziehungsweise Abstands (positive Werte) von Gatestapel und Sourcegebiet (links) und die resultierenden Transferkennlinien (rechts).

Parameter der Dotierprofile. So ist es ebenso von Bedeutung, an welcher Position sich das Maximum der Sourcedotierung befindet. Dieses wird durch die Weite der Spaceroxide und die Diffusion der Dotierstoffatome unter diese definiert. Die nachfolgend analysierten Dotierprofile sind in Abbildung 4.14 links dargestellt. Auf der rechten Seite sind zudem die Transfercharakteristiken bei drei unterschiedlichen Varianten gezeigt. Positive Werte von S beschreiben den Abstand zwischen dem Maximum des Sourcedotierprofils und dem Gatestapel und negative Werte einen entsprechenden Überlapp. Hierbei fällt auf, dass nur dann ein Unterschied im Bauelementverhalten zu beobachten ist, wenn ein Abstand zwischen Gatestapel und dem Maximum der Sourcedotierung besteht.

Ein erheblich genaueres Bild erhält man beim Blick auf die Ströme zwischen Source- und Drainkontakt des Tunneltransistors im Durchlass- und im Sperrbetrieb. In Abbildung 4.15 sind diese in Abhängigkeit von der Position des Gatestapels in Bezug auf die Sourcereion des zylindrischen Nanodraht-Tunneltransistors zu sehen. Die Koordinatenachsen beider Kennlinien sind mit einer linearen Skala dargestellt und es bestätigt sich der bereits zuvor geschilderte Eindruck, dass im Fall eines Überlapps von Sourcereion und Gatestapel das Bauelementverhalten nahezu unabhängig von der genauen Position ist. Der Strom ist im Durchlassbetrieb tatsächlich identisch. Besteht allerdings ein lateraler Abstand zwischen Sourcereion und Gatestapel, so fällt im Durchlassbereich mit steigendem Abstand eine signifikante Abnahme des Stroms auf.

Eine Begründung des beobachteten Sachverhalts kann man in Abbildung 4.16 sehen. So sind in diesem Bild die Band-zu-Band Tunnelraten, die elektrischen Feldlinien und die elektrische Feldstärke in zwei unterschiedlichen Tunneltransistoren dargestellt. Im oben abgebildeten Fall handelt es sich um ein Bauelement mit 5 nm Spacerabstand und im unten gezeigten Fall um eines mit 5 nm Überlapp von Sourcereion und Gatestapel. Auf diese Weise werden die beiden untersuchten Extremfälle miteinander verglichen. Hierbei

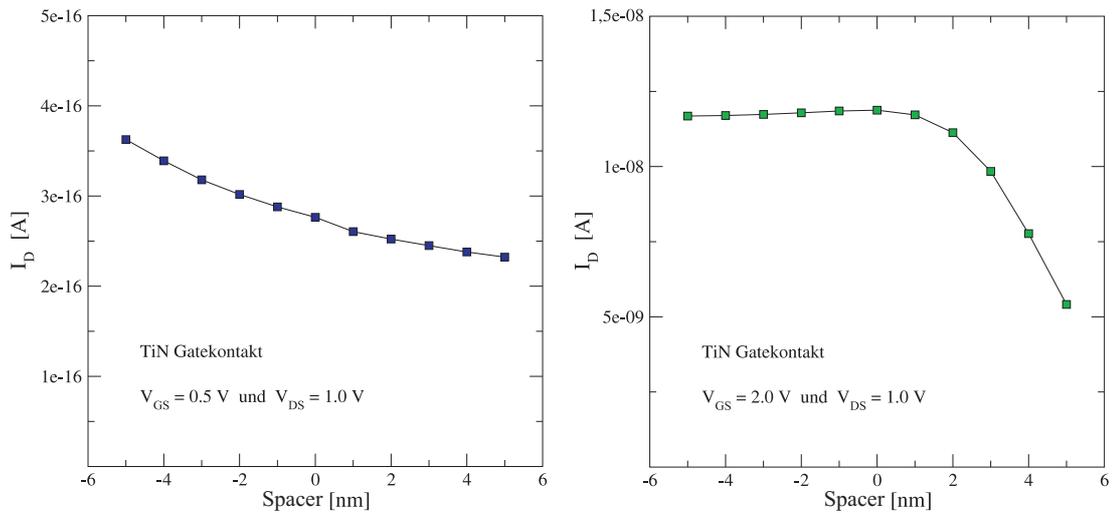


Abb. 4.15: Sperr- (links) und Durchlass-Ströme (rechts) bei unterschiedlichen Positionen des Gatestapels in Bezug auf die Sourcereion des CNW-TFETs.

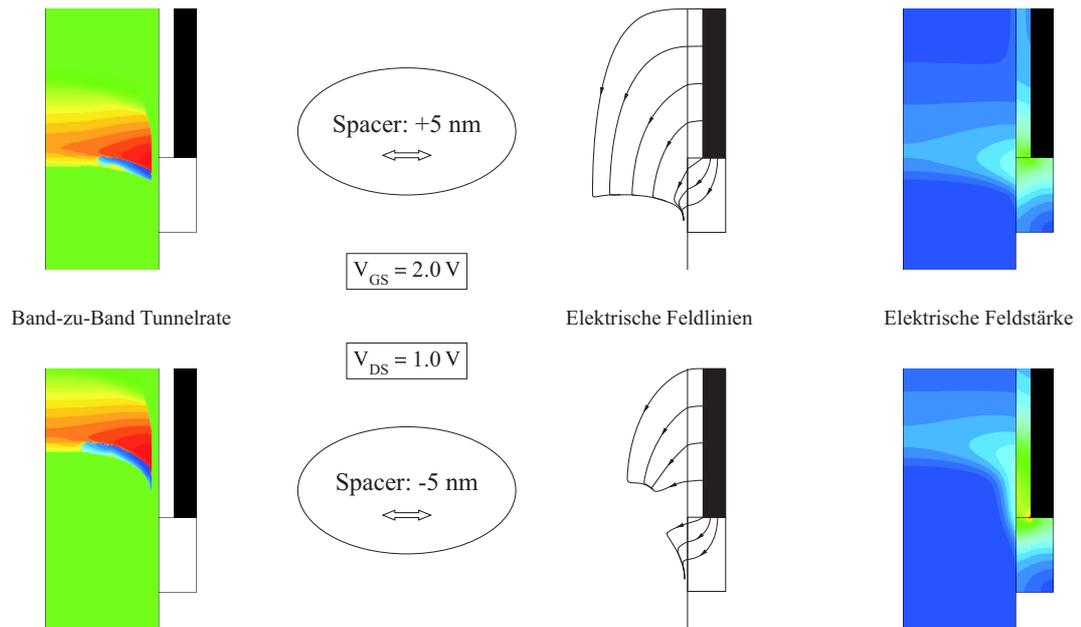


Abb. 4.16: Vergleich der Band-zu-Band Generationsraten (links) und der dazugehörigen elektrischen Feldlinien und -stärke (rechts) bei verschiedenen Positionen des Gatestapels in Bezug auf die Sourcereion des CNW-TFETs.

liegt im ersten Fall die Tunnelregion unter der Ecke des Gatestapels und im zweiten Fall ist die Tunnelregion vollständig unter dem Gate lokalisiert. Die elektrische Feldstärke ist bei einer Überlappung höher und das Band-zu-Band Tunneln wird daher nahezu allein durch das elektrische Feld, das durch das Gateoxid verläuft, getragen. Demgegenüber wird die Tunnelwahrscheinlichkeit im Fall eines großen Spacerabstandes wesentlich von

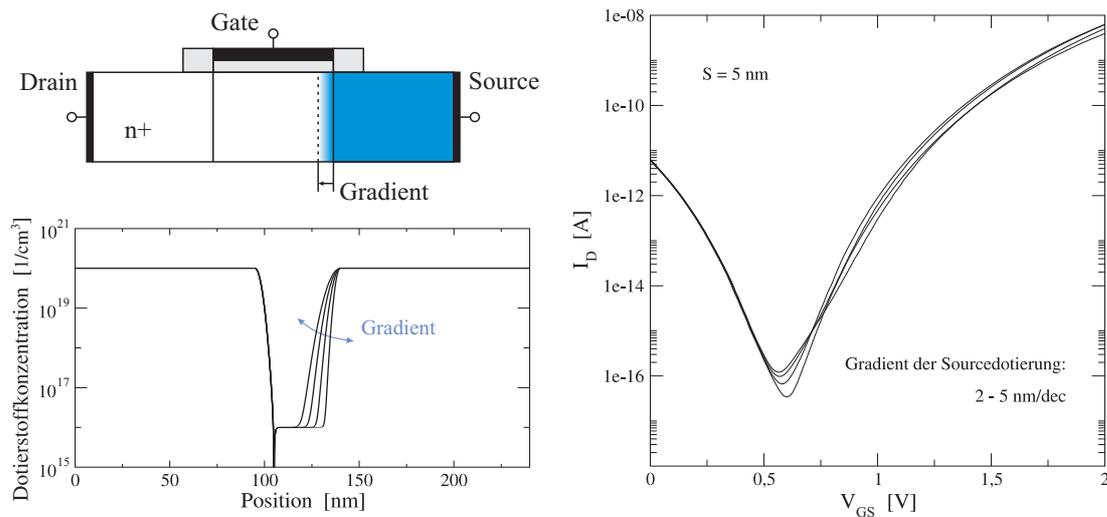


Abb. 4.17: Schematische Darstellung und Dotierprofile der Variation des Gradienten der Sourcedotierung (links) und die dazugehörigen Transferkennlinien (rechts).

dem durch das Spaceroxid verlaufende elektrische Feld beeinflusst. Sind auch die Ströme bei einer Überlappung höher, so ist im Hinblick auf die Skalierungseigenschaften der Mittelweg zwischen den zwei gezeigten Varianten vorzuziehen. Diese Entscheidung wird im nächsten Abschnitt untermauert und auch bei der Analyse der Einflüsse verschiedener Spacermaterialien in Abschnitt 4.4.1 an weiterer Bedeutung gewinnen.

4.2.4 Gradient der Sourcedotierung

Der letzte zu untersuchende Aspekt der Auswirkungen der Dotierprofile auf das Verhalten der Nanodraht-Tunneltransistoren ist der Gradient der Sourcedotierung. Hierbei kann auf eine Beschreibung der Effekte, die durch eine Variation der Draindotierstoffgradienten folgen, verzichtet werden, da nach Möglichkeit die Draindotierkonzentrationen gering gehalten werden sollten. Nachfolgend werden zwei verschiedene Parametervariationen durchgeführt: Zuerst mit einem 5 nm Spacer zwischen Gatestack und Source region und anschließend ohne Abstand. In Abbildung 4.17 sind die Dotierprofile in einem Schnitt durch den Tunneltransistor für einige unterschiedliche Gradienten der Sourcedotierung aufgezeigt. Zudem sind auf der rechten Seite dieser Abbildung die daraus resultierenden Transferkennlinien im Fall eines Spacerabstandes zu sehen. Allerdings kann aus dieser Form der Darstellung nur die sehr indifferente Situation festgestellt werden, da sich kein genereller Trend aus einer Änderung der Steilheit der Sourceprofile ableiten lässt.

Um eine sinnvolle Bewertung der einzelnen simulierten Bauelementvarianten vornehmen zu können, sind in Abbildung 4.18 sowohl die Sperrströme (links) als auch die Ströme im Durchlassbereich (rechts), in Abhängigkeit von den Gradienten der Sourcedotierung für die Tunneltransistoren mit Spacerabstand dargestellt. Aus diesen Charakteristiken kann einerseits abgelesen werden, dass die Sperrströme bei flacher werdenden Dotierprofilen zunehmen. Allerdings ist die verschobene, linear skalierte Ordinate zu beachten. Zudem

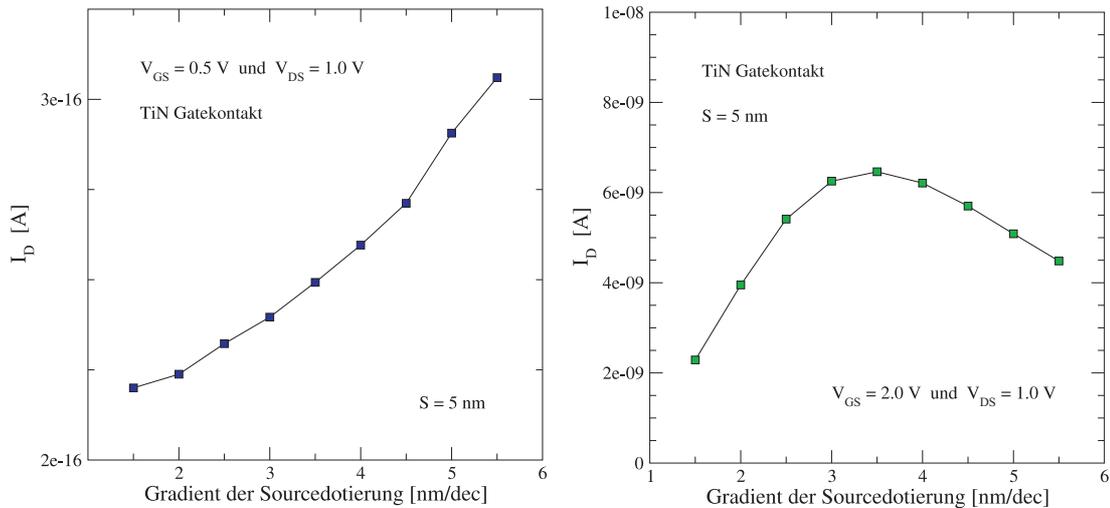


Abb. 4.18: Sperr- (links) und Durchlass-Ströme (rechts) bei unterschiedlichen Gradienten der Sourcedotierung des CNW-TFETs mit einem 5 nm Spacerabstand.

resultiert ein Maximum der Ströme im Durchlassbereich bei einem Gradienten von 3.5 nm pro Dekade. Diese Situation ergibt sich, da bei steileren Dotierprofilen prinzipiell derselbe Effekt wie bei einer Erhöhung der Spacerabstände auftritt: Weil die elektrische Feldstärke in der immer weiter von dem Gatekontakt entfernten Tunnelregion geringer wird, nimmt die Wahrscheinlichkeit für Band-zu-Band Tunnelprozesse ab und daher nehmen auch die Ströme im Fall steilerer Profile ab. Werden dagegen flache Dotierprofile betrachtet, spielt ein anderer Effekt eine Rolle.

Zu diesem Zweck werden nun Tunneltransistoren ohne Spacerabstand untersucht. Dann ergeben sich die in Abbildung 4.19 dargestellten Transferkennlinien für eine Variation der Gradienten. Allein aus diesen Charakteristiken kann ein deutlich klareres Bild gewonnen werden: Die Ströme im Durchlassbereich nehmen bei einem steileren Dotierprofil auf der Sourceseite zu, die Sperrströme werden geringer und die Steigung der Kennlinien wird ebenfalls größer. Auf der rechten Seite sind zusätzlich die elektrischen Feldstärken in einem Schnitt 2 nm unter dem Gateoxid für drei verschiedene Gradienten abgebildet. Es ist zu erkennen, dass die Maxima der Feldstärke mit steigendem Gradienten in Richtung des Sourcekontaktes verschoben sind und zudem auch größer werden. Im Fall abrupterer Dotierprofile werden bei hohen Gatespannungen die dazugehörigen Raumladungszonen kleiner und ausgeprägter und dadurch entstehen die schmaleren und gleichzeitig höheren Feldspitzen. Werden nur kleine Gatespannungen angelegt, dann folgt wegen der steileren Sourcedotierung eine längere Kanalregion und damit ein besseres Sperrverhalten der dann vorliegenden pin-Diode.

Die jeweiligen Durchlass- und Sperrströme des untersuchten Tunneltransistors ohne Spacerabstand sind in Abbildung 4.20 noch einmal zusammengefasst. So zeigt sich eine nur geringe Veränderung der Sperrströme mit dem Gradienten der Sourcedotierung und dagegen eine erhebliche Auswirkung auf die Durchlassströme. Einen Sonderfall stellt die Situation bei einem Dotierprofilgradienten unter 2 nm pro Dekade dar. Hierbei sättigt der

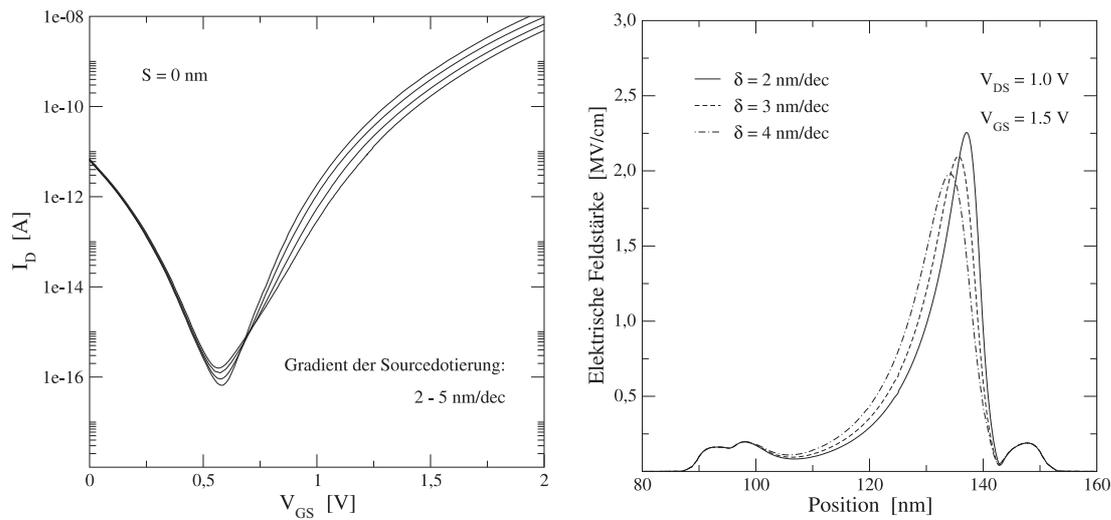


Abb. 4.19: Transferkennlinien (links) des CNW-TFETs ohne Spacerabstand und Verlauf der elektrischen Feldstärke entlang einer Schnittlinie 2 nm unter dem Gateoxid (rechts) bei einer Variation des Sourcegradienten.

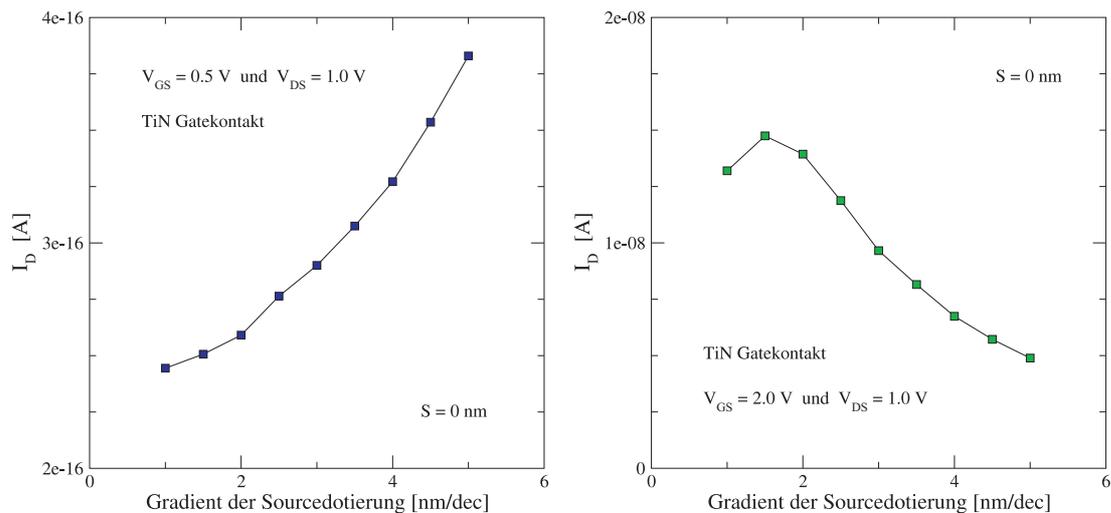


Abb. 4.20: Sperr- (links) und Durchlass-Ströme (rechts) bei unterschiedlichen Gradienten der Sourcedotierung des CNW-TFETs ohne Spacerabstand.

Strom zuerst einmal und nimmt für Gradienten unter 1.5 nm pro Dekade sogar ab. Der Grund dafür liegt in der Verkleinerung der Abmessungen der Raumladungszonen und der daraus resultierenden geringeren Breite der Feldspitzen. Nun sind zwar die Feldstärken höher und damit auch die dazugehörigen Band-zu-Band Tunnelraten, jedoch wird in einer kleineren Region eine relevante Tunnelrate auftreten. Dadurch sättigt zuerst der Strom und anschließend nimmt er bei einer weiteren Erhöhung der Steigung sogar ab. Einige Beispiele an Band-zu-Band Tunnelraten und Shockley-Read-Hall-Rekombinationsraten entlang einer Schnittlinie 2 nm unter dem Gateoxid sind in Abbildung 4.21 illustriert. Es muss dabei allerdings bedacht werden, dass sowieso keine beliebig abrupten Dotierprofile

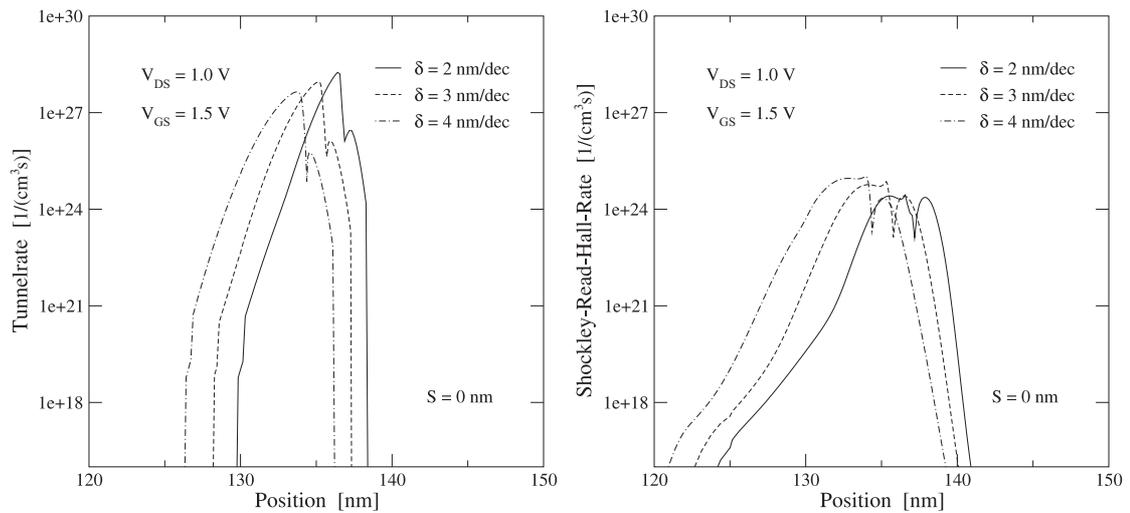


Abb. 4.21: Verlauf der Band-zu-Band Tunnelrate (links) und der Shockley-Read-Hall-Rate (rechts) bei verschiedenen Gradienten der Sourcedotierung des Tunneltransistors entlang einer Schnittlinie 2 nm unter dem Gateoxid.

möglich sind. Schließlich handelt es sich um diskret verteilte Dotierstoffatome und damit nicht um ein kontinuierliches Profil. Zusammen mit den Ergebnissen aus dem letzten Teilabschnitt folgt für die angestrebte Optimierung der Eigenschaften von zylindrischen Tunneltransistoren, dass einerseits keine Spacerabstände, beziehungsweise maximal ein geringer Überlapp, anzustreben sind und andererseits ein möglichst steiler Gradient der Sourcedotierung wünschenswert ist.

4.3 Skalierungseigenschaften bei zylindrischen Nanodraht-Tunneltransistoren

Weil Tunnelfeldeffekttransistoren im Hinblick auf die Skalierbarkeit vielversprechende Eigenschaften aufweisen, wird in diesem Abschnitt untersucht inwieweit dies auch auf zylindrische Nanodraht-Tunneltransistoren zutrifft. Unter dem Begriff Skalierbarkeit wird im Allgemeinen verstanden, dass die Abmessungen der Bauelemente verringert werden können ohne die grundlegenden Schalteigenschaften wesentlich zu verändern. Zunächst wird die Gatelänge der Tunneltransistoren und anschließend der Durchmesser variiert, da hiermit der Platzbedarf der Tunneltransistoren verkleinert werden kann.

4.3.1 Gatelänge

Zuerst wird der Einfluss einer Verringerung der Gatelänge auf das Bauelementverhalten des Tunneltransistors untersucht. Dazu wird in einem ersten Schritt das in Abschnitt 4.1 eingeführte Referenzbauelement als Ausgangspunkt benutzt und hierbei die Gatelänge

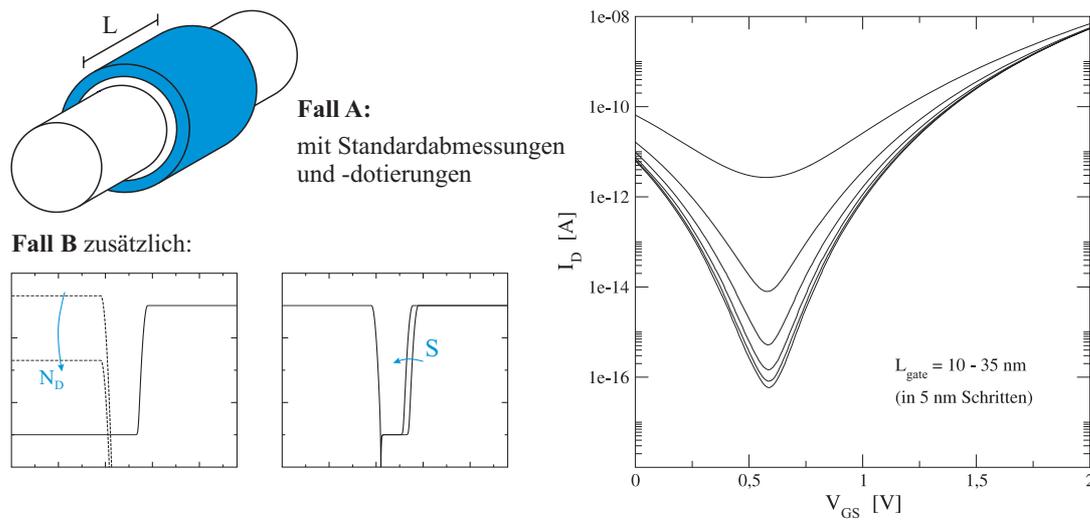


Abb. 4.22: Schematische Darstellung der untersuchten TFET-Varianten zur Analyse der Auswirkungen einer Gatelängenänderung (links) und die dazugehörigen Transferkennlinien im Fall A (rechts).

variiert. Anschließend wird dieselbe Gruppe an Variationen für einen leicht modifizierten Tunneltransistor betrachtet. Hierzu wird die Dotierstoffkonzentration der Drainseite auf $5 \cdot 10^{18} \text{ cm}^{-3}$ reduziert und der Spacerabstand von 5 nm auf Null gesetzt. Beide Varianten sind in Abbildung 4.22 (links) schematisch dargestellt. Rechts sind die Transferkennlinien für den ersten Fall bei einer Änderung der Gatellänge gezeigt.

Einige Skalierungseigenschaften können bereits aus dieser dargestellten Kennlinienschar abgelesen werden: Über einen weiten Bereich zeigt die Transfercharakteristik nur geringe Veränderungen. Unter 20 nm Gatellänge wird allerdings besonders das Sperrverhalten der zylindrischen Nanodraht-Tunneltransistoren sehr schlecht. Mit dem Sperrstrom wird das Minimum der Transferkennlinie bezeichnet. Ein genaueres Bild dieser Situation liefert ein Blick auf die Durchlass- und Sperrströme in Abhängigkeit von der Gatellänge. Diese sind in Abbildung 4.23 gezeigt. So ist die Invarianz der Ströme im Durchlassbereich klar zu erkennen. Demgegenüber steigt der Sperrstrom mehr als exponentiell an. Im Fall eines Tunneltransistors mit 10 nm Gatellänge ist dieser um vier Größenordnungen höher als bei dem Referenzbauelement. Tunneltransistoren mit noch kleineren Abmessungen weisen überhaupt keine Schalteigenschaften auf, da dann zwischen Sperr- und Durchlassbereich letztlich nicht mehr unterschieden werden kann [51].

Um das beobachtete Verhalten der Tunnelfeldeffekttransistoren bei einer Änderung der Gatellänge zu erklären, lohnt sich ein Blick auf die elektrischen Feldstärken innerhalb der Bauelemente. In Abbildung 4.24 sind auf der linken Seite die elektrischen Feldstärken von drei Tunneltransistoren mit unterschiedlicher Gatellänge in einem Schnitt 2 nm unter dem Gateoxid dargestellt. Dabei wird der Sperrbereich bei einer angelegten Gatespannung von 0,5 V betrachtet. Auf der rechten Seite finden sich die dazugehörigen Konturbilder der Tunnelraten und der Elektronendichten. Im Fall des 30 nm-Bauelements ist in der Transferkennlinie wegen des sehr niedrigen Sperrstroms ein ausgeprägtes Schaltverhalten

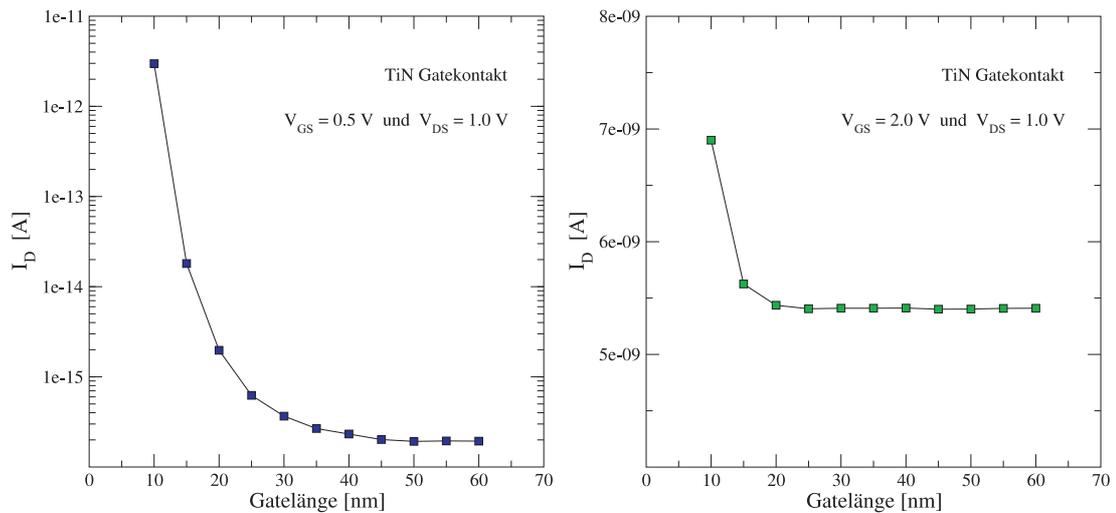


Abb. 4.23: Sperr- (links) und Durchlass-Ströme (rechts) bei verschiedenen Gatelängen für CNW-TFETs im Fall A (siehe Abb. 4.22).

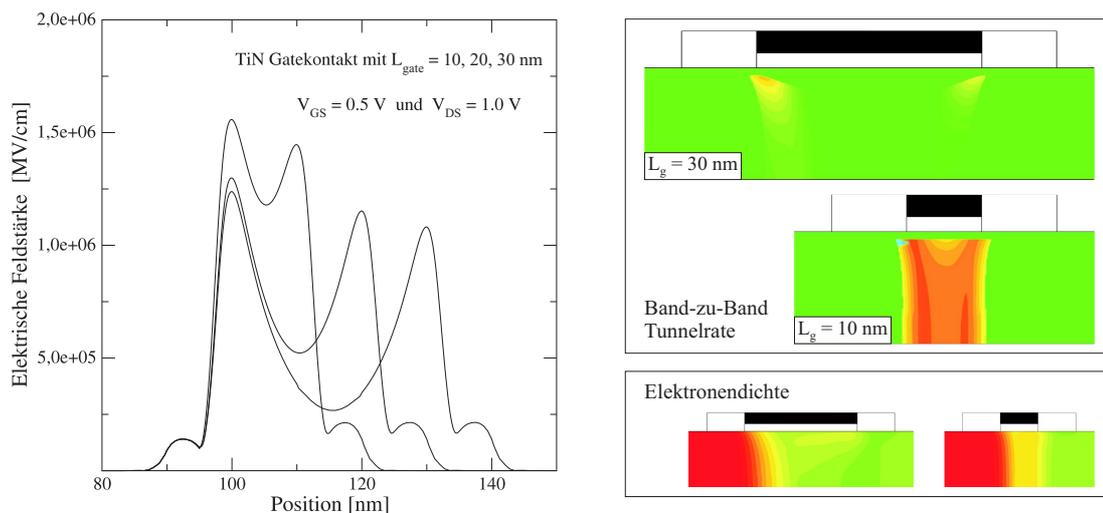


Abb. 4.24: Verlauf der elektrischen Feldstärke entlang einer Schnittlinie 2 nm unter dem Gateoxid (links) und zugehörige Konturbilder der Band-zu-Band Tunnelraten und Elektronendichten (rechts) im Sperrbetrieb bei verschiedenen Gatelängen des CNW-TFETs aus Fall A (siehe Abb. 4.22).

zu beobachten. Einen Grund liefert die verhältnismäßig geringe elektrische Feldstärke im Bauelement. So finden sich im untersuchten Betriebszustand grundsätzlich zwei Maxima an den jeweiligen Regionengrenzen Source/Kanal und Drain/Kanal. Diese sind allerdings räumlich deutlich voneinander getrennt und nicht besonders hoch. In der Folge ist auch die Tunnelwahrscheinlichkeit gering und nur ein relativ geringer Sperrstrom fließt durch das Bauelement. Wird nun die Gatelänge verringert, so werden sich die Maxima mehr und mehr überlagern. Auf diese Weise steigen dann auch die Feldstärkewerte an und

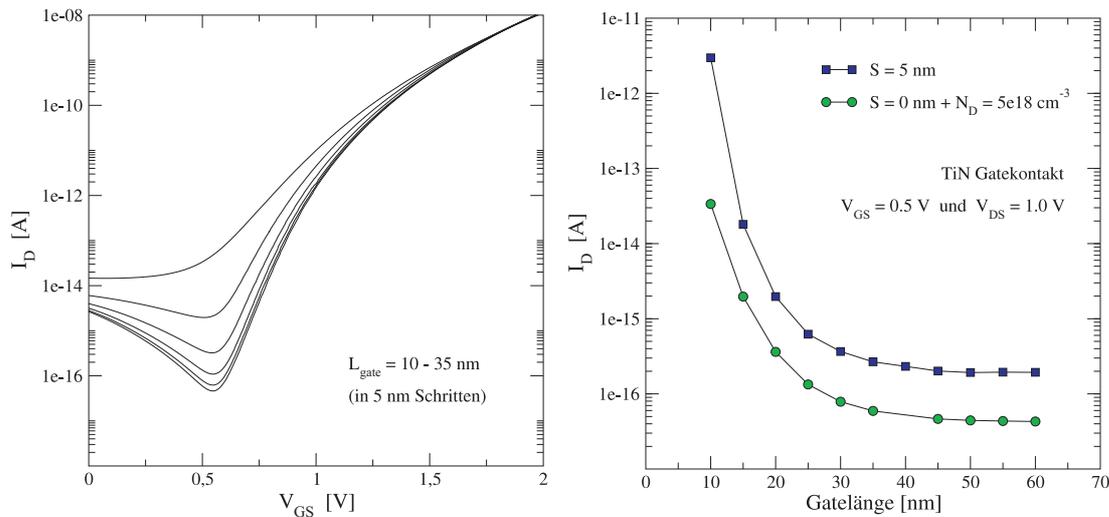


Abb. 4.25: Transfercharakteristiken der CNW-TFETs im Fall B (links) und Vergleich der Sperrströme in den beiden untersuchten Fällen (rechts) bei einer Variation der Gatelänge.

die Tunnelraten werden signifikant höher. Ist schließlich die Gatelänge kleiner als 10 nm wird nur noch ein sehr hohes Feldstärkemaximum vorliegen und dann auch der Einfluss des Gatekontaktes auf das Bauelementverhalten verschwinden.

Der wiederum hinter dieser Beobachtung stehende Grund ist durch die Dotierprofile und die hieraus folgenden Elektronendichteverteilungen gegeben. Das Funktionsprinzip des Tunneltransistors basiert auf der Erzeugung eines durch den Gatekontakt steuerbaren Elektronenkanals. Ist allerdings die Gatelänge so gering wie hier beschrieben, werden sich die Source- und die Draindotierungen nahezu überlappen, so dass unabhängig von der angelegten Gatespannung konstruktionsbedingt ein abrupter hochdotierter pn-Übergang vorliegt und damit das Bauelementverhalten durch diese Zenerdiode dominiert wird. Um diesen Effekt zu vermeiden, wird die bereits in Abschnitt 4.2.1 beschriebene Reduktion der Draindotierung vorgenommen. Wird zudem der Spacerabstand auf der Sourceseite von 5 nm auf Null gesetzt, so folgt die oben beschriebene Variante B.

Die für diese Bauelementvariante resultierenden simulierten Transferkennlinien sind in Abbildung 4.25 links für verschiedene Gatelängen gezeigt. Aus dieser Darstellung ist der wesentlich geringere Einfluss der Gatelänge auf das Schaltverhalten des Tunneltransistors zu erkennen. So sind die Ströme im Durchlassbereich auch weiterhin von der Gatelänge unbeeinflusst und bei niedrigen angelegten Gatespannungen ist der Anstieg des Stroms weniger ausgeprägt. Diese Situation ist auf der rechten Seite von Abbildung 4.25 noch besser herauszulesen. Hier sind die Sperrströme bei einer angelegten Gatespannung von 0.5 V in Abhängigkeit von der Gatelänge dargestellt. Dabei zeigt sich zusätzlich, dass die Sperrströme insgesamt kleiner als bei Variante A sind. Darüberhinaus werden die oben getroffenen Aussagen bestätigt. Auch für diese Simulationsreihe ist der Blick auf die elektrische Feldstärkeverteilung innerhalb des Bauelements sinnvoll. In Abbildung 4.26 sind auf der linken Seite die elektrischen Feldstärken von einem 10 nm Tunneltransistor

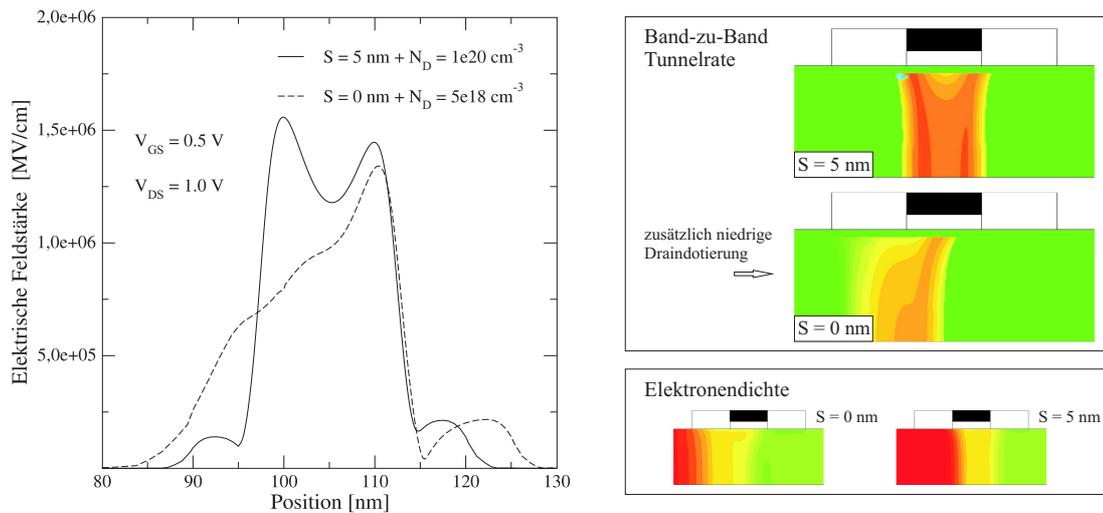


Abb. 4.26: Vergleich des Verlaufs der elektrischen Feldstärke entlang einer Schnittlinie 2 nm unter dem Gateoxid (links) und der zugehörigen Elektronendichten und Band-zu-Band Tunnelraten (rechts) bei CNW-TFETs mit 10 nm Gatelänge in den beiden untersuchten Fällen (siehe Abb. 4.22).

im Fall A und B in einem Schnitt 2 nm unter dem Gateoxid dargestellt. Hierbei ist ein grundsätzlich unterschiedlicher Verlauf der Feldstärke zu beobachten: So weist die zweite Variante nur noch ein Maximum der elektrischen Feldstärke auf. Dieses ist zudem auch weniger hoch. Somit ist zu erwarten, dass die Tunnelwahrscheinlichkeit nur noch an dem Übergang von Kanal- zur Sourcereion signifikant hoch ist. Die aus diesem Feldverlauf folgenden Auswirkungen auf die Tunnelraten sind wiederum auf der rechten Seite der Abbildung 4.26 illustriert. Es ist deutlich zu erkennen, dass einerseits die Region, in der die Tunnelraten hoch sind, kleiner wird und andererseits auch die Tunnelraten insgesamt geringer werden. Auf diese Weise resultiert dann der um mehr als zwei Größenordnungen niedrigere Sperrstrom. Allerdings fällt bei beiden Varianten auf, dass die Tunnelregion sich über den gesamten Durchmesser des Tunneltransistors erstreckt. Im Vergleich zu der in Abbildung 4.24 dargestellten Tunnelgenerationsrate des 30 nm-Bauelements liegt damit ein erheblicher Unterschied vor. Einen Grund stellt die geringere Gatefläche bei Kurzkanaltransistoren dar, die in lokal höheren elektrischen Feldstärken resultiert. Dabei werden Elektronen aus dem Draingebiet weniger effizient von dem pn-Übergang nahe der Sourceseite ferngehalten. Eine weitere Verbesserung des Sperrverhaltens kann letztlich durch eine größere Spacerweite auf der Drainseite erzielt werden.

4.3.2 Durchmesser

Ein weiterer Weg die Abmessungen von zylindrischen Tunneltransistoren zu reduzieren besteht darin den Durchmesser der Kanalregion zu variieren. Dabei wird dann wie bei der Verkleinerung der Gatelänge die wirksame Gatefläche linear abnehmen. Grundsätzlich steht zu erwarten, dass keine signifikanten Auswirkungen auf das Bauelementverhalten

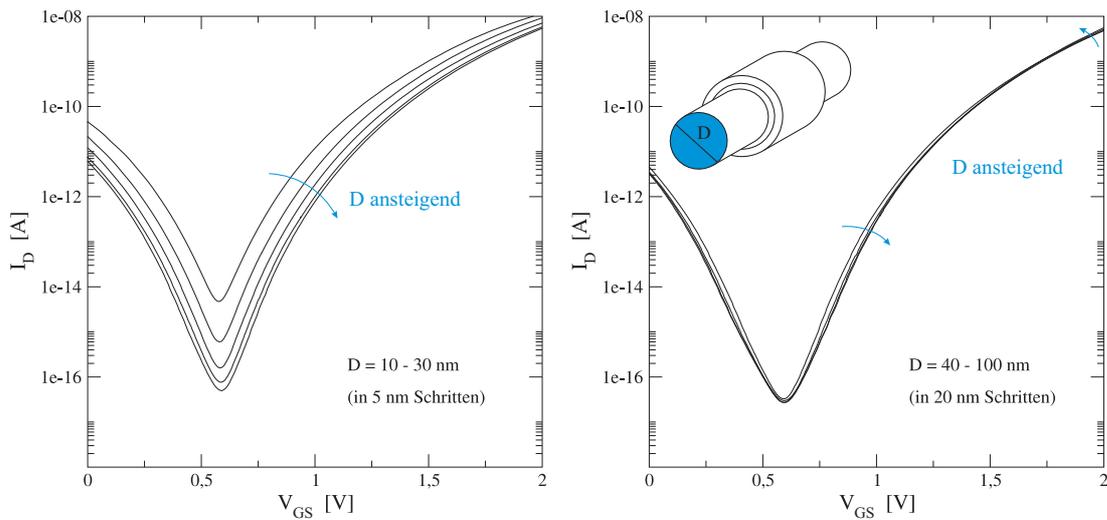


Abb. 4.27: Transferkennlinien des zylindrischen Nanodraht-TFETs im Fall kleiner (links) und im Fall großer Werte (rechts) des Bauelementdurchmessers.

auftreten werden, da das Band-zu-Band Tunneln vornehmlich in einer verhältnismäßig kleinen Region am pn-Übergang zwischen Source- und Kanalregion unter dem Gateapfel stattfindet. Auch für diese Variationsreihe wird das Referenzbauelement aus Abschnitt 4.1 herangezogen. Die Änderung des Bauelementdurchmessers von 10 bis 100 nm und deren Auswirkungen auf die Transferkennlinien des Tunneltransistors sind in Abbildung 4.27 dargestellt. Hierbei sind auf der linken Seite die simulierten Kennlinien von Bauelementen mit kleinen Kanaldurchmessern gezeigt. Es fällt auf, dass die Ströme zwischen Drain- und Sourcekontakt durch den Tunneltransistor mit fallenden Durchmessern über den gesamten betrachteten Bereich der Gatespannung ansteigen. Allerdings gibt es im Fall größerer Durchmesser (ab 50 nm) keinen wirklich relevanten Unterschied der Transferkennlinien der verschiedenen Tunneltransistoren. Diese Kennlinien sind auf der rechten Seite der Abbildung gezeigt.

Eine etwas genauere Untersuchung ist wieder möglich, wenn die Ströme im Sperr- und im Durchlassbereich in Abhängigkeit von dem Durchmesser aufgetragen werden. Dies ist in Abbildung 4.28 durchgeführt. Der Durchlassbereich ist zuerst einmal etwas weniger auffällig, da hierbei die Ströme nicht übermäßig voneinander abweichen (man beachte die lineare Skalierung der Stromachse). Im links gezeigten Fall der Sperrströme (bei einer angelegten Gatespannung von 0.5 V) ändert sich der Strom nicht für Tunneltransistoren zwischen 40 bis 100 nm Durchmesser. Eine weitere Verringerung des Kanaldurchmessers führt allerdings zu dramatischen Konsequenzen. So resultiert ein erheblicher, das heißt mehr als exponentieller, Anstieg der Sperrströme durch den Tunneltransistor, wenn der Durchmesser kleiner als 30 nm wird. Jedoch muss an dieser Stelle darauf hingewiesen werden, dass selbst bei einem TFET mit 10 nm Durchmesser noch immer ein sehr gutes Schaltverhalten und ein ausreichend niedriger Sperrstrom zu beobachten sind. Wird nun auch der Durchlassstrom bei einer angelegten Gatespannung von 2 V betrachtet, so zeigt sich zuerst einmal eine etwas überraschende Situation: Die Charakteristik zylindrischer

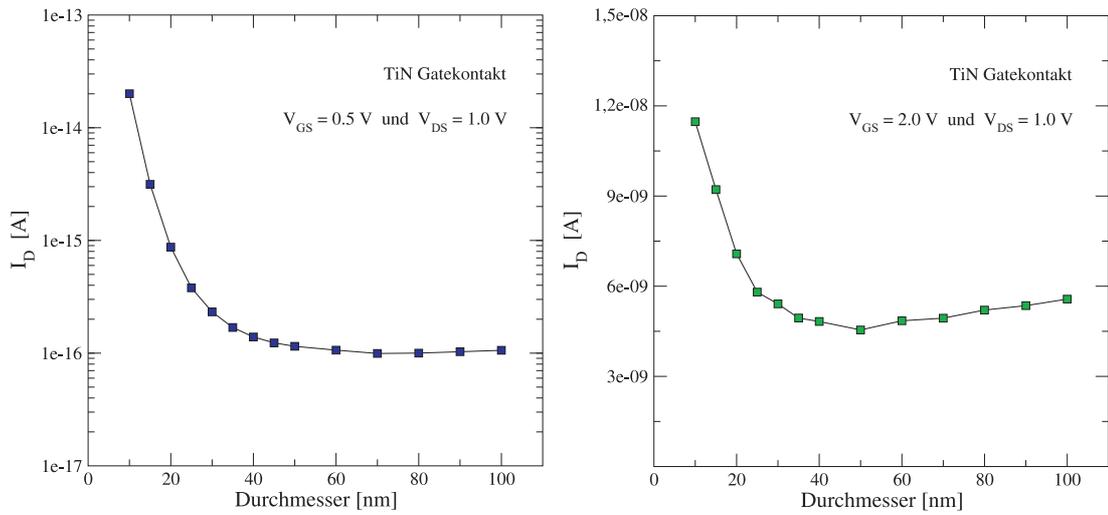


Abb. 4.28: Sperr- (links) beziehungsweise Durchlass-Ströme (rechts) der CNW-TFETs in Abhängigkeit vom Durchmesser der Kanalregion.

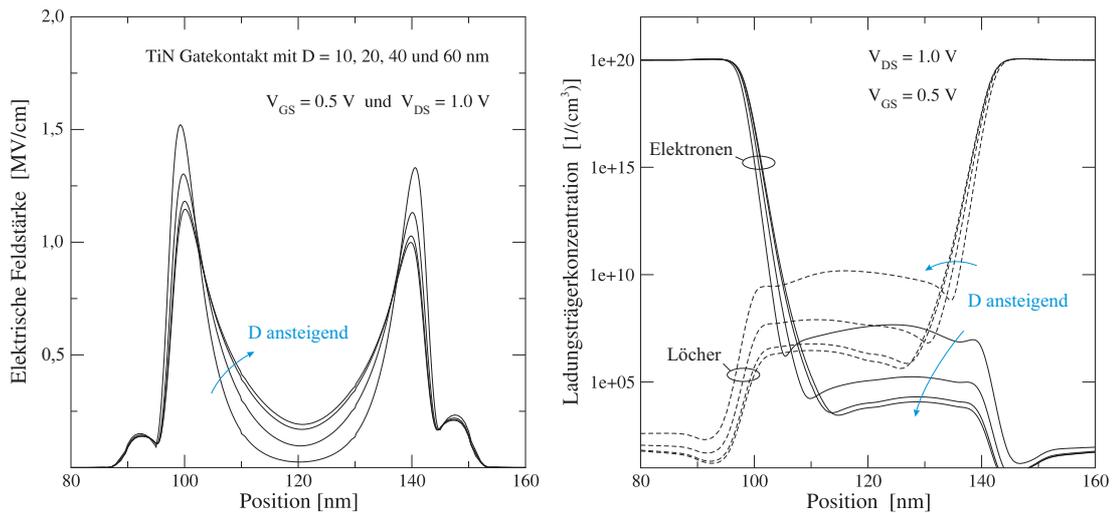


Abb. 4.29: Vergleich der elektrischen Feldstärken (links) und der Ladungsdichten (rechts) von CNW-TFETs bei einer Variation des Durchmessers in einer Schnittlinie 2 nm unter dem Gateoxid.

Nanodraht-Tunneltransistoren weist bei einem Durchmesser von 50 nm ein Minimum im Durchlassstrom auf. Besonders ausgeprägt ist dabei der Anstieg bei einer fortgesetzten Reduktion der Abmessungen.

Um zu beantworten, weshalb die beobachteten Effekte auftreten, lohnt sich ein Blick auf die elektrischen Feldstärken und die Ladungsträgerdichten. In Abbildung 4.29 sind der elektrische Feldstärkeverlauf (links) und die Elektronen- und Löcherverteilung (rechts) in einer Schnittlinie 2 nm unter dem Gateoxid abgebildet. Dabei werden vier verschiedene Tunneltransistoren im Sperrbetrieb bei einer Gatespannung von 0.5 V betrachtet. Mit dem

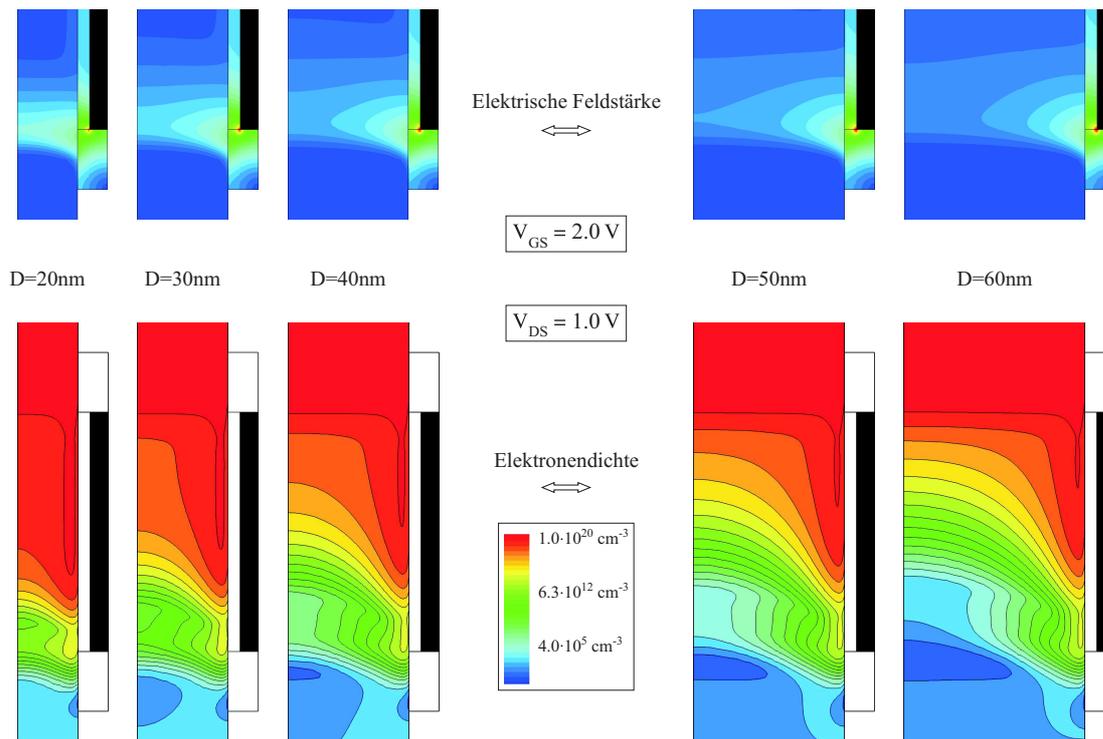


Abb. 4.30: Konturbilder der elektrischen Feldstärken (oben) und der Elektronendichten (unten) von CNW-TFETs bei einer Variation des Durchmessers.

Anstieg des Durchmessers nimmt zwar die elektrische Feldstärke im Kanalgebiet zu, doch sinkt diese zunächst signifikant an den beiden Ecken des Gatestapels ab. Eine weitere Erhöhung führt aber nicht zu einer weiteren Verkleinerung der Feldstärkemaxima und somit zu einer Sättigung der Tunnelraten und dementsprechend der Sperrströme. Einen Grund stellt die stärkere wechselseitige Beeinflussung der Gatekontakte im Inneren des Kanals dar. So wird bei großen Durchmessern das Tunneln und damit auch der Stromfluss nur in einem Ring unter dem Gatestapel stattfinden. Im Fall eines dünnen Kanals wird die Tunnelregion mehr und mehr zu einer Kreisscheibe, da nun die elektrische Feldstärke über den gesamten Durchmesser sehr hoch wird.

Dieser Sachverhalt führt dann auch zu den beschriebenen Effekten im Durchlassbereich des Tunneltransistors. So sind in den Konturbildern der Abbildung 4.30 die elektrischen Feldstärken (oben) und die Elektronendichten (unten) gezeigt. Dabei bestätigen sich die oben getroffenen Aussagen: Im Fall großer Durchmesser ist die Elektronendichte nur in einer schmalen Region unter dem Gatestapel sehr hoch und es sind zudem nur geringe Unterschiede, sowohl der Elektronendichte, als auch der elektrischen Feldstärke zwischen einem TFET mit 50nm und einem mit 60nm Durchmesser festzustellen. Wird nun der Durchmesser verringert, so nimmt die Feldstärke zu und das Maximum konzentriert sich zunehmend am gesamten pn-Übergang. Der Elektronenkanal dehnt sich nun über den vollständigen Durchmesser des Tunneltransistors aus. Die aus dieser Situation folgenden Band-zu-Band Generationsraten und die Shockley-Read-Hall-Rekombinationsraten der

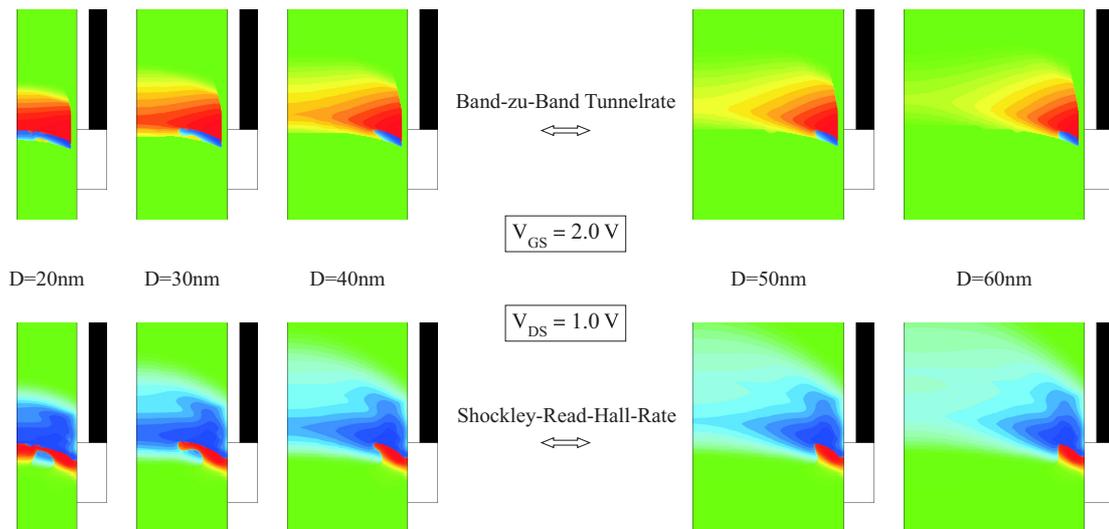


Abb. 4.31: Vergleich der Band-zu-Band Tunnelraten (oben) und der Shockley-Read-Hall-Rekombinationsraten (unten) bei einer Variation des Durchmessers.

fünf betrachteten TFETs im Durchlassbetrieb sind in Abbildung 4.31 dargestellt. Hierbei ist die Zunahme der Raten bei kleinen Durchmessern und die gleichzeitige Ausdehnung der Region hoher Tunnelwahrscheinlichkeit auf die gesamte Querschnittsfläche im Gebiet des pn-Übergangs zu beobachten. Letztlich kann somit das Verhalten im Durchlassbetrieb erklärt werden: Bei hohen Durchmessern findet Tunneln nur in einem Ring unter dem Gatestapel statt. Dadurch steigt der Strom linear mit dem Durchmesser an. Im Fall kleiner Durchmesser nimmt der Strom dann wieder zu, da nun in der gesamten Querschnittsfläche ein Strom fließen wird. Wenn schließlich der Durchmesser weiter reduziert wird, dann erfolgt zunächst eine Sättigung des Stromes bis hin zu einer rapiden Abnahme. Allerdings ist dieser Bereich (unter 10 nm) nicht mehr im Rahmen des eingesetzten semiklassischen Modells sinnvoll simulierbar und nur mit Hilfe einer vollständigen quantenmechanischen Simulation können physikalisch vernünftige Ergebnisse erhalten werden.

4.4 Auswirkungen unterschiedlicher Materialien auf das Bauelementverhalten

Um eine möglichst gute Optimierung der Eigenschaften von Tunnelfeldeffekttransistoren erreichen zu können, ist es notwendig, die Auswirkungen verschiedener Materialien zu untersuchen. In diesem Rahmen wird zuerst der Gatestapel betrachtet. Die Analyse der Gateoxidstärke ist hierbei als Vorstufe zur Variation der Oxidmaterialien zu sehen. Da die Bandlücke des eingesetzten Kanalmaterials einen nicht zu vernachlässigenden Einfluss auf die Tunnelwahrscheinlichkeit hat, ist es zudem sinnvoll die Auswirkungen von SiGe anstelle von Silizium im Kanal zu untersuchen. Wichtig ist nicht allein die Konzentration an Germanium, sondern auch das Gebiet in dem SiGe anstelle von Si eingesetzt wird.

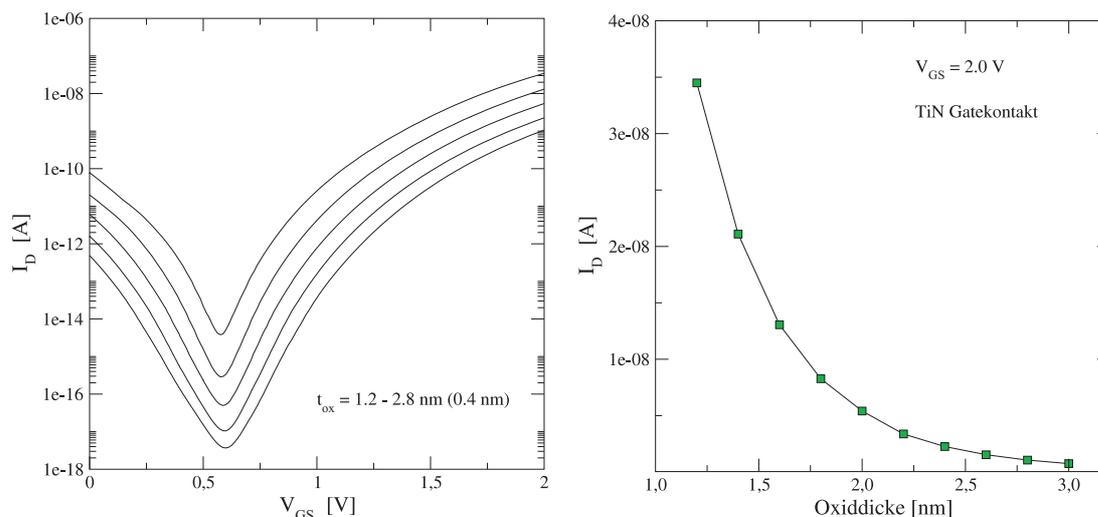


Abb. 4.32: Transfercharakteristiken (links) und Vergleich der Durchlass-Ströme (rechts) von CNW-TFETs bei einer Variation der Gateoxidicke.

4.4.1 Variation des Gate-Stapels (Material und Dicke)

Als Basis der folgenden Untersuchungen dient der in den vorangehenden Abschnitten eingesetzte zylindrische Referenztunneltransistor mit einem Durchmesser von 30 nm und einer Gatelänge von 40 nm. Das Gateoxid ist aus Siliziumdioxid und 2 nm dick. Erneut wird als Gatemetallisierung TiN mit einer Elektronenaustrittsarbeit von 4.7 eV eingesetzt und die Spaceroxide sind an den beiden Seiten des Gatestapels 15 nm breit, wobei die in Abschnitt 4.2.3 eingeführte Spacerweite (das heißt der Abstand der Dotierprofilmaxima zum Gatestapel) weiterhin 5 nm beträgt. Diese Seitenoxide erlauben die Ausbreitung von Streufeldern ausgehend von dem Gatemetall und beeinflussen somit das Verhalten der Tunneltransistoren in einem nicht zu vernachlässigendem Maß. Wird die Gateoxidicke zwischen 1.2 und 3 nm variiert ohne die sonstigen Bauelementgrößen zu verändern, so folgen die in Abbildung 4.32 dargestellten Transfercharakteristiken. Auf den ersten Blick sieht es aus, als ob die Kennlinien bei einer Reduktion der Oxidicke einfach zu höheren Strömen verschoben sind. Allerdings ist diese Vermutung nicht ganz zutreffend, da durch die Verringerung der Oxidicke die Feldstärken im Siliziumkanal erhöht werden. Somit setzt der Band-zu-Band Tunnelstrom bereits bei kleineren Gatespannungen ein und die beiden Zweige der Transferkennlinien rücken näher zusammen, wodurch dann letztlich der steilere Verlauf nahe des Minimums und die höheren Ströme im Durchlassbereich resultieren. Auf der rechten Seite dieser Abbildung sind die Ströme zwischen Source und Drain im Durchlassbereich in Abhängigkeit von der Gateoxidicke aufgetragen. Die exponentielle Zunahme des Stroms bei einer Gatespannung von 2 V legt nahe, dass ein optimaler Tunneltransistor ein möglichst dünnes Gateoxid aufweisen sollte. Denn zum einen nimmt der Strom im Durchlassfall signifikant zu und zum anderen ist der Strom im Sperrfall bei einer Gatespannung von 0.6 V trotz der gleichermaßen starken Zunahme immer noch hinreichend gering. Allerdings werden mit dieser Zielsetzung einige andere Probleme immer größer: So ist es extrem schwierig dünne Siliziumdioxidschichten zu

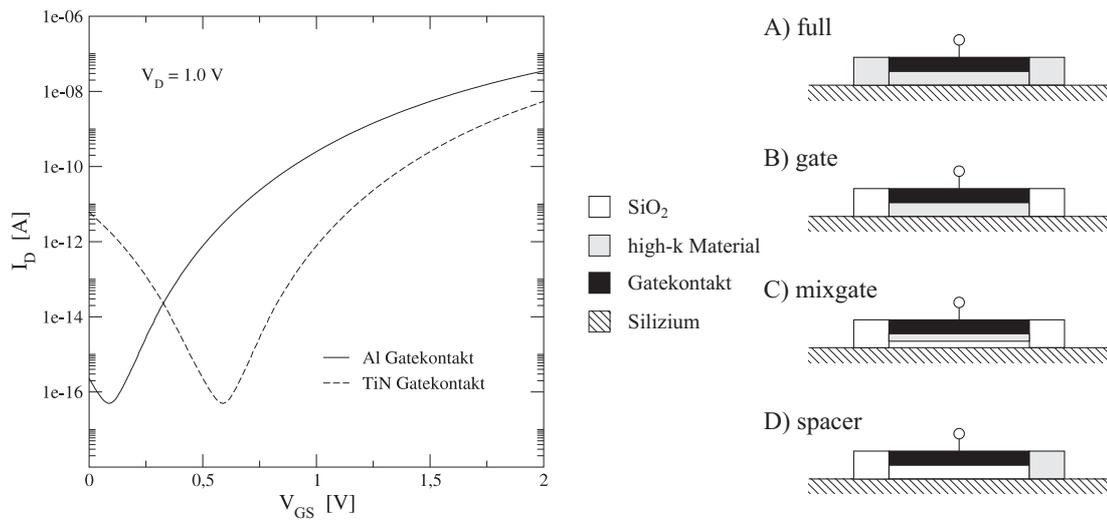


Abb. 4.33: Transferkennlinien von zylindrischen Nanodraht-TFETs bei einer Variation des Gatekontaktmaterials (links) und Bauelementvarianten bei der Analyse von high-k Materialien als Gateoxid (rechts).

fertigen. Als Untergrenze fungiert ein Quasimonolayer aus SiO₂ der circa 0.16 nm dick ist und die Erkenntnis, dass erst ab einer Dicke von 0.7 nm von einer funktionierenden Isolationsschicht gesprochen werden kann [52]. Darüberhinaus wird im Fall sehr dünner Gateoxide die Wahrscheinlichkeit für ein Tunneln von Elektronen durch das Gateoxid immer höher und es resultiert ein nicht zu vernachlässigender Gateleckstrom. Auf welche Weise die beschriebenen Verbesserungen der Bauelementeigenschaften dennoch genutzt werden können, wird auf den nächsten Seiten beschrieben. Zu diesem Zweck werden dann alternative Oxidmaterialien eingesetzt.

Bevor allerdings die Auswirkungen unterschiedlicher Materialien in den Oxidregionen ausführlich diskutiert werden, sollte auf den Einfluss der Wahl des Gatekontaktmaterials eingegangen werden. Bei nanostrukturierten Bauelementen werden typischerweise die Gatekontakte aus Polysilizium oder Metallen gefertigt [53]. In der bisherigen Simulation zylindrischer Nanodraht-Tunneltransistoren wurde als Gatekontaktmaterial TiN mit einer Elektronenaustrittsarbeit von 4.7 eV eingesetzt. Übliche Werte der Austrittsarbeit von TiN schwanken zwischen 4.2 und 5.3 eV [54, 55]. Hierbei zeigt sich, dass das Minimum des Drainstromes bei einer Gatespannung von 0.6 V liegt. Durch die Wahl von Aluminium als alternatives Gatekontaktmaterial kann die in Abbildung 4.33 dargestellte Verbesserung der Transfercharakteristik erreicht werden. In der Simulation wird für Aluminium eine Elektronenaustrittsarbeit von 4.2 eV verwandt. Je nach Fertigungsprozess variiert diese typischerweise zwischen 4.06 und 4.26 eV [56]. Auf diese Weise folgt eine Verschiebung der Transferkennlinie um 0.5 V und letztlich ein erheblich besseres Schaltverhalten des Tunneltransistors. Es bleibt festzuhalten, dass durch die Wahl des Gatekontaktmaterials nur die Kennlinien verschoben werden und damit eine relativ einfache Möglichkeit zur Auswahl eines optimierten Materials besteht. Abschließend ist noch der folgende Hinweis angebracht: So werden bei modernen nanostrukturierten MOS-Bauelementen inzwischen

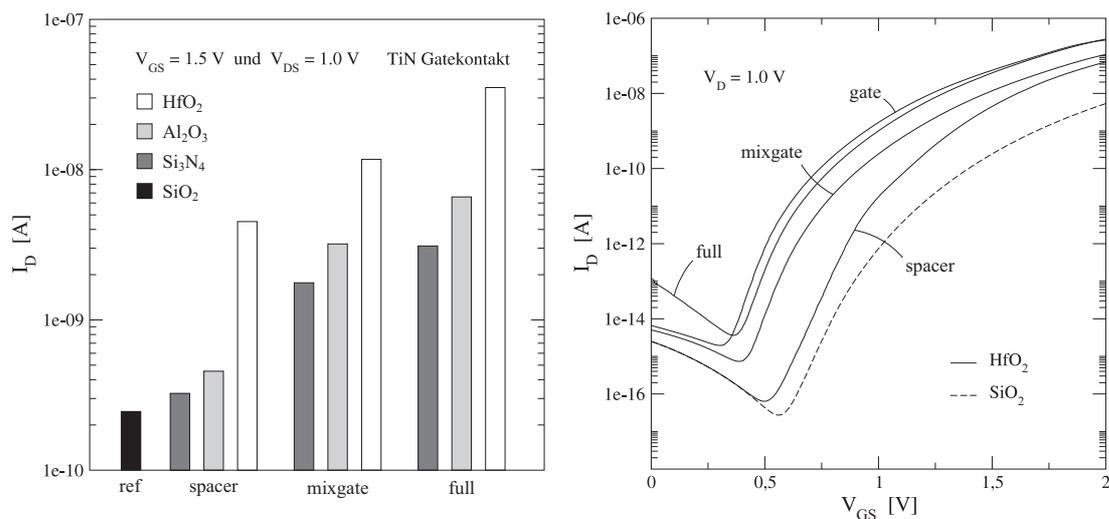


Abb. 4.34: Durchlass-Ströme der CNW-TFETs mit verschiedenen Gateoxidmaterialien in unterschiedlichen Realisierungen nach Abb. 4.33 (links) und Vergleich der Transferkennlinien bei Verwendung von HfO_2 als Gateoxid (rechts).

vornehmlich Metallgatekontakte eingesetzt, da bei dotiertem Polysilizium als Gatekontakt und gleichzeitiger Verwendung von high-k Gateoxiden unter anderem die Beweglichkeit der Ladungsträger im Kanal stark abnimmt. Durch die hohe Zahl an Ladungsträgern im Metall kann diese Abnahme der Beweglichkeit jedoch kompensiert werden [53].

Auf der rechten Seite von Abbildung 4.33 sind die betrachteten Tunneltransistorvarianten für die Analyse von sogenannten high-k Oxiden dargestellt. Unter einem high-k Material versteht man Oxide mit einer hohen relativen Dielektrizitätszahl. Als Bezugsgröße dient hierbei Siliziumdioxid mit $\epsilon = 3.9$. Um nun eine gute Steuerwirkung des Gatekontaktes auf den Kanal im Silizium zu erhalten, können, wie bereits ausgeführt, die Oxiddicken verkleinert oder Materialien mit hohen Dielektrizitätszahlen eingesetzt werden. Aus dem Vergleich mit Silizium ($\epsilon = 11.7$) folgt, dass das natürliche Oxid einen extrem geringen Wert aufweist. Allerdings kann das SiO_2 ohne größere Schwierigkeiten in der Fertigung von MOS-Bauelementen eingesetzt werden. Andere Materialien müssen eine Vielzahl von Anforderungen erfüllen, um eine sinnvolle Alternative darzustellen. Beispielsweise muss eine gute Gitteranpassung zwischen Oxid und Silizium gegeben sein, damit es nicht zu unerwünschten Verspannungen im Siliziumkanal kommt [52]. Die einzelnen Varianten aus der Abbildung unterscheiden sich im Hinblick auf die Gebiete im Gatestapel in denen ein high-k Material eingesetzt wird: Das Gate- und die beiden Seitenoxide (A), allein das Gateoxid (B) und nur das sourceseitige Seitenoxid (D) werden aus dem high-k Material gefertigt. Im Fall der dritten Variante besteht das Gateoxid aus einem Schichtstapel mit einer 0.4 nm dünnen Schicht SiO_2 direkt auf dem Silizium und einem high-k Oxid. Auf diese Weise werden die Verspannungen im Kanal reduziert.

Die Simulationsergebnisse der entsprechenden Tunnelfeldeffekttransistorvarianten sind in Abbildung 4.34 dargestellt. Rechts finden sich die Transferkennlinien, wenn als high-k Material HfO_2 mit der relativen Dielektrizitätszahl $\epsilon = 22$ verwandt wird. Zusätzlich ist

auch die simulierte Transferkennlinie des Referenzbauelements mit SiO_2 als alleinigem Oxidmaterial eingezeichnet. Aus dieser Abbildung kann man deutlich erkennen, dass ein signifikanter Einfluss auf die Schalteigenschaften resultiert: So verschiebt sich einmal die Einsatzspannung zu niedrigeren Werten der angelegten Gatespannung. Dieser Effekt ist besonders ausgeprägt, wenn das Gateoxid teilweise oder ganz aus HfO_2 besteht. Zudem steigen die Ströme im Durchlassbereich um bis zu zwei Größenordnungen an und bei den vier high-k Varianten nimmt die Steigung der Transferkennlinien im Bereich zwischen Sperr- und Durchlassbereich zu. Im *spacer*-Fall ist die Verschiebung der Einsatzspannung wesentlich geringer und der Strom im Sperrbereich bleibt nahezu unverändert. Der Grund hierfür ist, dass nur das Seitenoxid auf der Sourceseite aus HfO_2 besteht. Da dennoch der Strom im Durchlassbereich deutlich ansteigt, ist die Verbesserung der Schalteigenschaften bei diesem Bauelement sehr ausgeprägt. Aus den abgebildeten Transferkennlinien kann man herauslesen, dass je nach Wahl des Oxids und der eingesetzten Bauelementvariante entweder ein passendes Gatekontaktmaterial gewählt oder, wie bereits in Abschnitt 4.2.2 beschrieben, durch eine Anpassung der Kanaldotierung die Einsatzspannung verschoben werden muss. In den Fällen mit einem SiO_2 -Gateoxid ist die Wahl von Aluminium als Gatekontaktmaterial sinnvoll (siehe Abbildung 4.33). Auch im Hinblick auf HfO_2 macht diese Wahl noch Sinn, da die Sperrströme bei einer Gatespannung von Null Volt immer noch hinreichend klein bleiben. Werden allerdings Oxidmaterialien mit höheren relativen Dielektrizitätszahlen technologisch verfügbar, so kann es wiederum sinnvoll sein, auf TiN als Gatekontaktmaterial zurückzugreifen. Wegen der beschriebenen Möglichkeit durch die Prozesstechnologie die Austrittsarbeit in einem weiten Bereich zu variieren, ist dies besonders interessant.

Wird zu Vergleichszwecken als high-k Material beispielsweise Si_3N_4 mit $\varepsilon = 7.5$ oder Al_2O_3 mit $\varepsilon = 9.6$ verwandt, so resultieren die auf der linken Seite von Abbildung 4.34 aufgetragenen Ströme zwischen Drain und Source bei einer angelegten Gatespannung von 1.5 V. Dabei ist Si_3N_4 ein in der Herstellung von MOS-Bauelementen gebräuchliches Material und Al_2O_3 wegen des Einsatzes von Aluminium als Gatekontaktmaterial von Interesse. Hierbei fällt auf, dass in den Fällen *mixgate* und *full* die erwarteten Zunahmen des Drainstromes auftreten. Diese sind wegen der um mehr als Faktor zwei kleineren Dielektrizitätszahlen naturgemäß deutlich geringer als im Fall von HfO_2 . Allerdings sollte nicht vergessen werden, dass auch die Verschiebung der Einsatzspannung einen nicht zu vernachlässigenden Anteil an der Zunahme des Stromes im Durchlassbetrieb liefert. Eine Möglichkeit diese zwei Effekte zu trennen, ist in Abbildung 4.35 links illustriert. Dabei werden die einzelnen simulierten Transferkennlinien so verschoben, dass diese sich bei einer Gatespannung von 0.5 V und einem Drainstrom von 10^{-14} A schneiden. Bei diesem Strom handelt es sich um einen hinreichend kleinen Sperrstrom und daher stellt dieser Punkt einen geeigneten Bezugspunkt dar. Daran anschließend werden die Ströme bei einer Gatespannung von 1.5 V ausgewertet. Auf diese Weise wird untersucht, inwieweit sich der Drainstrom bei einer Variation der Gatespannung von 1 V verändert.

Es resultieren dann die auf der rechten Seite von Abbildung 4.35 aufgetragenen Ströme im Durchlassbereich. Im Vergleich mit Abbildung 4.34 fällt auf, dass die Unterschiede zwischen den einzelnen Varianten nicht mehr so ausgeprägt sind. Wie erwartet, zeigt sich bei der *spacer*-Variante im Fall von HfO_2 ein sehr gutes Schaltverhalten. So ist der Strom

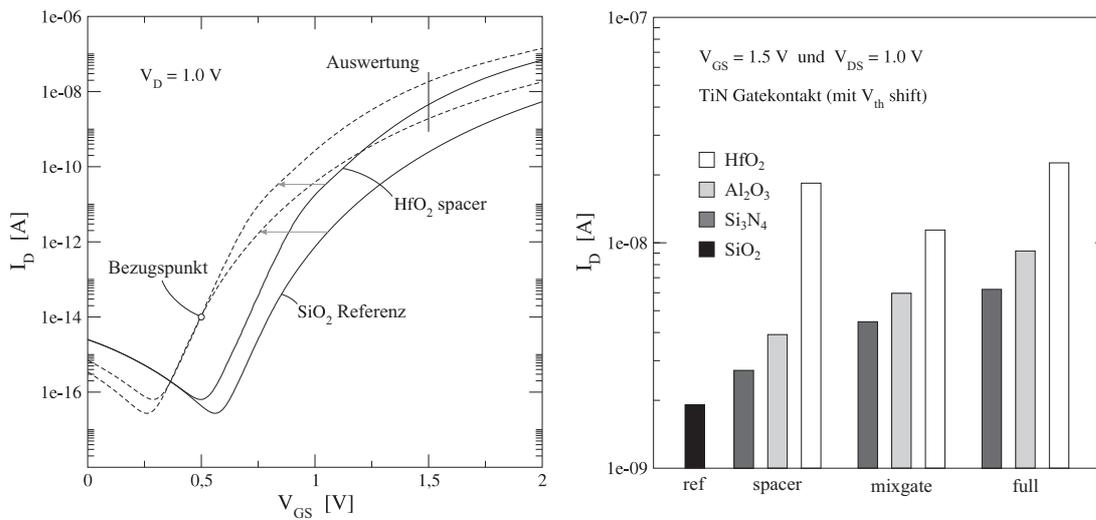


Abb. 4.35: Methode zu Kompensation der Verschiebung der Einsatzspannung (links) und daraus resultierende Durchlass-Ströme der CNW-TFETs mit verschiedenen Oxidmaterialien in unterschiedlichen Realisierungen nach Abb. 4.33 (rechts).

nahezu gleich hoch, wie im Fall der vollständigen HfO₂-Variante (Fall A). Allerdings ist zu berücksichtigen, dass bei den *spacer*-Varianten keine SiO₂-Schicht zwischen Silizium und dem high-k Oxid vorliegt. Daher werden die Ströme wegen dieser Zwischenschicht unter Umständen ein wenig geringer ausfallen. Zudem ist bei diesen Bauelementen die Abhängigkeit der Ströme im Durchlassbereich von der relativen Dielektrizitätszahl des Seitenoxids besonders ausgeprägt. Damit ist der Vorteil der Tunneltransistoren mit einem Si₃N₄ oder Al₂O₃ Seitenoxid im Vergleich zum SiO₂-Referenzbauelement nicht mehr wirklich vorhanden. Insgesamt sind auch die Verschiebungen der Einsatzspannungen der Tunneltransistoren im Verhältnis zum Referenzbauelement von Interesse: Beispielsweise wird die Einsatzspannung bei einer vollständigen Ersetzung aller Oxide durch HfO₂ um 0.38 V verschoben. Je geringer die Dielektrizitätszahl ist, desto weniger ausgeprägt ist die Verschiebung. Zum Vergleich wird bei der *spacer*-Variante mit HfO₂ die Einsatzspannung nur um 90 mV geringer.

Abschließend werden die elektrischen Feldstärken im Tunneltransistor im Hinblick auf die einzelnen Bauelementvarianten untersucht. So sind in Abbildung 4.36 die elektrischen Feldstärken und die Band-zu-Band Tunnelraten bei unterschiedlichen HfO₂-Varianten des CNW-TFETs entlang einer Schnittlinie 2 nm unter dem Gateoxid dargestellt. Durch den Einbau von HfO₂ steigt hierbei die elektrische Feldstärke deutlich an. Zudem verschiebt sich im Fall der Varianten mit einem HfO₂-Seitenoxid die Position des Maximums. Der Gatekontakt über dem Siliziumkanal befindet sich zwischen 100 und 140 nm. Damit findet das Band-zu-Band Tunneln bei diesen Bauelementen auch vermehrt unter dem Seitenoxid statt. Wird die Tunnelrate betrachtet, so bestätigt sich diese Aussage. Allerdings fällt eine Besonderheit auf: So sind die Maxima der Tunnelraten bei der *gate*- und *full*-Variante gleich hoch, obwohl die elektrischen Feldstärken unterschiedlich sind. Wegen der hohen Dotierung der Sourceregion nimmt in diesem Bereich auch die Energiebandlücke ab und

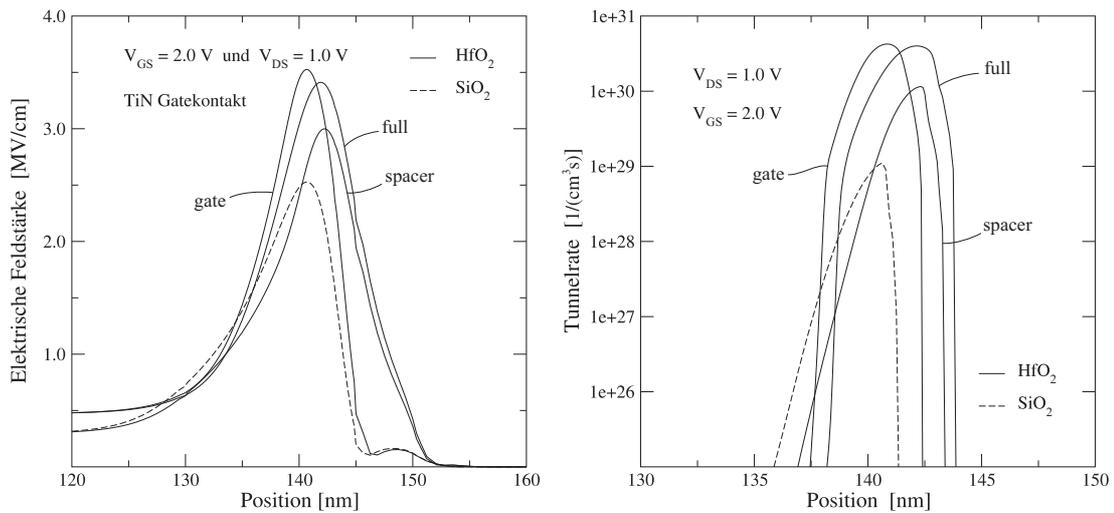


Abb. 4.36: Verlauf der elektrischen Feldstärke (links) und der Band-zu-Band Tunnelrate (rechts) bei unterschiedlichen HfO₂-Varianten des CNW-TFETs entlang einer Schnittlinie 2 nm unter dem Gateoxid.

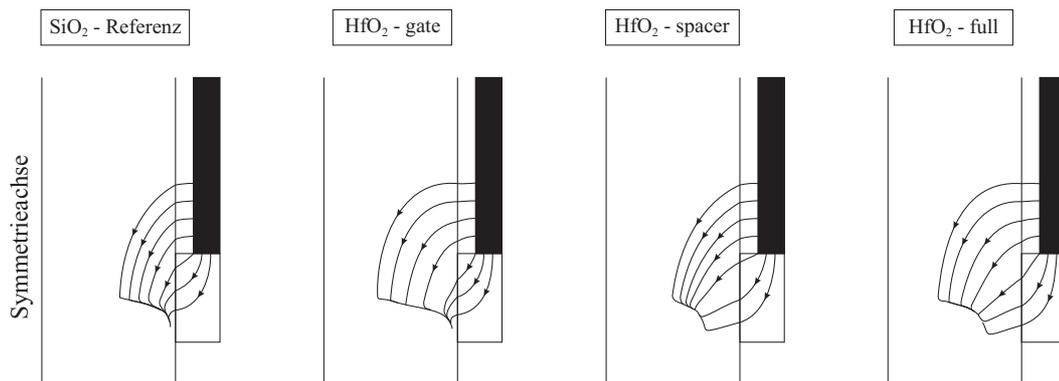


Abb. 4.37: Verlauf der elektrischen Feldlinien unter dem Gate bei $V_{DS}=1V$ und $V_{GS}=2V$ im Fall dreier verschiedener HfO₂-Varianten des CNW-TFETs im Vergleich zur reinen SiO₂-Ausführung.

damit steigt letztlich die Tunnelrate. Weil das Tunneln in dem Übergangsbereich zwischen schwach dotiertem Kanal und der hochdotierten Sourcereion stattfindet, folgt ein starker Einfluss der Position auf die Tunnelrate. Aus Abbildung 4.37 können die Verläufe der elektrischen Feldlinien der einzelnen Bauelementvarianten abgelesen werden. So reichen die elektrischen Feldlinien bei der Verwendung von high-k Oxidmaterialien tiefer in den Siliziumkanal hinein und damit steigt auch die Tunnelrate in größerer Entfernung zum Gatestack an. Außerdem zeigt sich, dass bei high-k Seitenoxiden nun auch Tunnelpfade in Richtung dieser Seitenoxide vorliegen.

In jedem Fall sind high-k Oxide im Gatestack ein wichtiges Optimierungswerkzeug, da auch nach der Eliminierung der Auswirkungen der Verschiebung der Einsatzspannung die

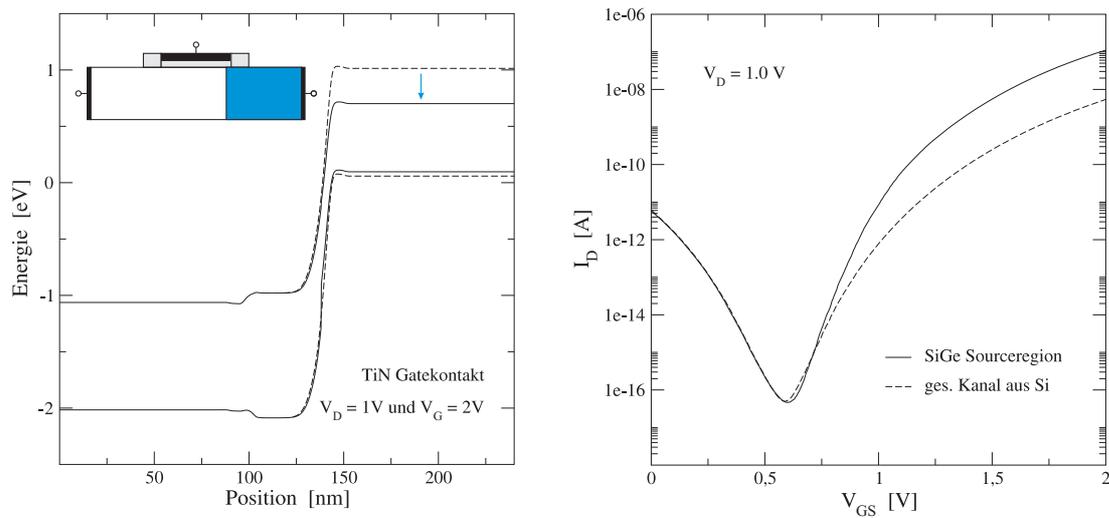


Abb. 4.38: Schematische Darstellung und Banddiagramm eines CNW-TFETs mit einer SiGe-Sourcereion (links) und der n-Betriebs-Transferkennlinie (rechts) im Vergleich zu dem Bauelement mit einem reinen Si-Kanal.

Ströme um bis zu eine Größenordnung zunehmen. Bei einer technologiereifen Einführung von high-k Materialien mit höheren Dielektrizitätszahlen resultiert zudem eine weitere Verbesserung der Bauelementeigenschaften von Tunnelfeldeffekttransistoren. Außerdem stellt die untersuchte *spacer*-Variante eine interessante Möglichkeit zur Fertigung von extremen Kurzkanaltransistoren dar.

4.4.2 Einfluss verschiedener Materialien im Kanalgebiet

Im letzten Abschnitt wurde bereits die Abhängigkeit der Tunnelrate von der Bandlücke beobachtet. Hierbei war die Abhängigkeit der Bandlücke von der Dotierkonzentration der auslösende Faktor. Durch den gezielten Einbau von Halbleitermaterialien mit einer kleinen Bandlücke in der Kanalregion wird nun untersucht, inwieweit die Eigenschaften von Tunneltransistoren weiter verbessert werden können. Ein attraktives Material ist der Verbundhalbleiter aus Silizium und Germanium. Bei Raumtemperatur ($T = 300$ K) hat Silizium eine Bandlücke von 1.12 eV und der ebenfalls indirekte Halbleiter Germanium weist eine Bandlücke von 0.67 eV auf [16]. Der Verbundhalbleiter SiGe hat, abhängig von der Zusammensetzung, eine Energiebandlücke zwischen diesen beiden Werten. In diesem Abschnitt werden verschiedene Zusammensetzungen und auch der Einfluss der Position der SiGe-Region auf das Schaltverhalten von Tunneltransistoren sowohl im n- wie auch im p-Betrieb untersucht.

Zuerst wird die Sourcereion des Referenztunneltransistors durch SiGe ersetzt. Dabei wird eine 50%-ige Germaniumkonzentration verwendet und zwischen SiGe-Region und dem Gateapfel gibt es einen Überlapp von 2 nm. Somit resultiert die in Abbildung 4.38 links dargestellte Bandstruktur im Durchlassbereich des Tunneltransistors. Aus Gründen der Vergleichbarkeit ist zudem der entsprechende Bandverlauf des Referenzbauelements

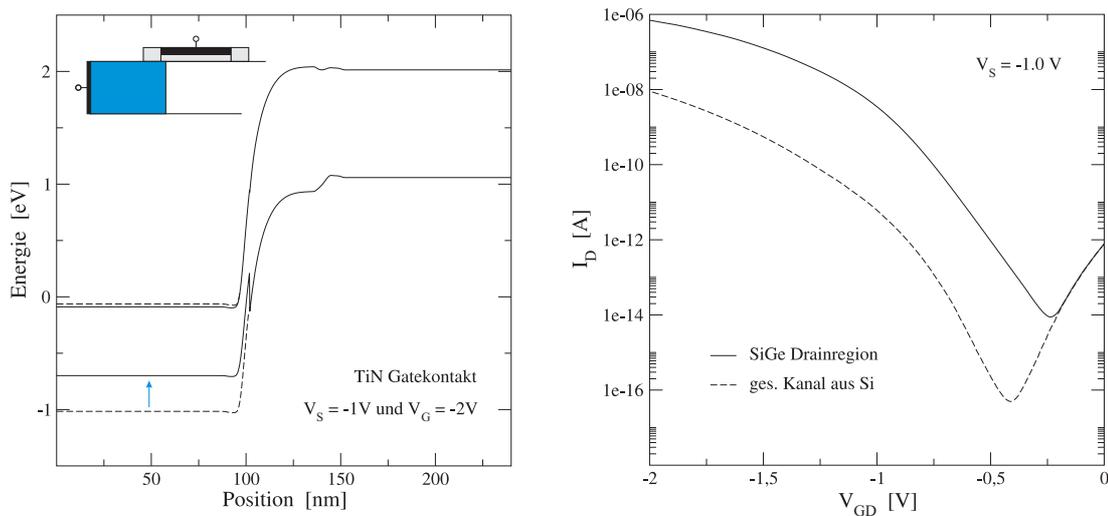


Abb. 4.39: Schematische Darstellung und Banddiagramm eines CNW-TFETs mit einer SiGe-Drainregion (links) und der p-Betriebs-Transferkennlinie (rechts) im Vergleich zu dem Bauelement mit einem reinen Si-Kanal.

eingezeichnet. Die dazugehörigen simulierten Transferkennlinien sind auf der rechten Seite dieser Abbildung zu sehen. Hierbei fällt auf, dass bei niedrigen Gatespannungen der Strom zwischen Source- und Drainkontakt unverändert ist und damit keine negativen Auswirkungen auf das Sperrverhalten zu beobachten sind. Im Durchlassbereich zeigt sich allerdings der erwartete Effekt: So nimmt der durch den Tunneltransistor fließende Strom um mehr als eine Größenordnung zu. Auf diese Weise bewirkt der Einbau von SiGe auf der Source-Seite die erwartete Verbesserung der Bauelementeigenschaften [57].

Wird die Drainseite des Tunneltransistors auf äquivalente Weise durch SiGe ersetzt, so folgt die in Abbildung 4.39 dargestellte Situation. Hier wird nun die Valenzbandkante um etwa $0.34 V$ nach oben verschoben und auf diese Weise die Tunnelbarriere signifikant verringert. Da sich die Position der Leitungs- beziehungsweise Valenzbandkante an dem Fermi-Niveau orientiert, ist wegen der n-dotierten Drainseite vornehmlich das Valenzband von der Bandlückenänderung betroffen. So folgt die auf der rechten Seite der Abbildung dargestellte Transferkennlinie. Im Vergleich mit dem Referenzbauelement zeigt sich eine etwas andere Situation wie im n-Betrieb des Tunneltransistors: Bei diesem Bauelement verschiebt sich die Einsatzspannung in erheblichem Maß, wodurch der minimale Strom um circa zwei Größenordnungen ansteigt. Insgesamt ist dieser Anstieg jedoch nicht weiter problematisch, da einmal der Sperrstrom noch immer hinreichend niedrig ist und zudem der Strom im Durchlassbereich bei einer angelegten Gatespannung von $-2 V$ um nahezu zwei Größenordnungen zunimmt. Somit ist mit diesen Simulationen bestätigt, dass die Verwendung von SiGe in den jeweiligen hochdotierten Bereichen sowohl im n- wie auch im p-Betrieb die Schalteigenschaften verbessert.

Es ist noch zu klären, auf welche Weise der Germaniumanteil in dem Verbundhalbleiter eine Rolle spielt. Zu diesem Zweck werden unterschiedliche Zusammensetzungen von $Si_{1-x}Ge_x$ analysiert. Hierbei folgt eine Beschränkung auf den n-Betriebszustand und der

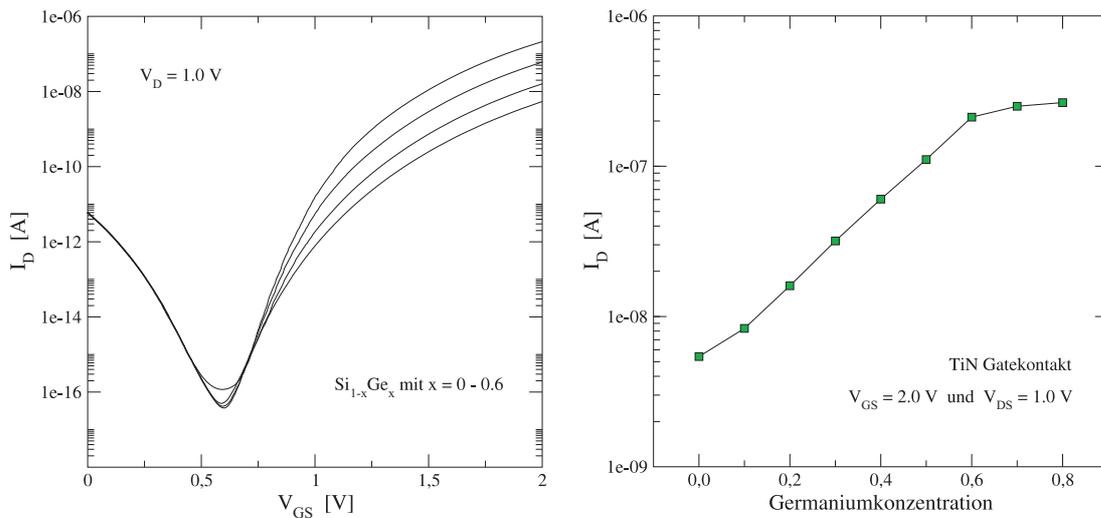


Abb. 4.40: Transferkennlinien eines CNW-TFETs mit einer SiGe-Sourcereion (links) und Durchlass-Ströme in Abhängigkeit von der Ge-Konzentration (rechts).

Überlapp zwischen der SiGe-Sourcereion und dem Gatestapel beträgt auch weiterhin 2 nm. Es resultieren dann die in Abbildung 4.40 dargestellten Transferkennlinien. Bei diesen ist der Germaniumkonzentrationsfaktor x in Schritten von 0.2 zwischen 0 (reines Silizium) und 0.6 variiert. Im Sperrbereich des Tunneltransistors, d.h. bei dem minimalen Stromfluss zwischen Source- und Drainkontakt, ist nur eine geringfügige Veränderung des Drainstroms zu beobachten. Demgegenüber nimmt der Strom im Durchlassbereich mit einer Erhöhung des Germaniumanteils um nahezu zwei Größenordnungen zu. Wird nun der Strom bei einer angelegten Gatespannung von 2 V über der Germaniumkonzentration aufgetragen, so folgt die auf der rechten Seite der Abbildung gezeigte Situation. Hierbei bestätigt sich zuerst einmal der signifikante Anstieg der Ströme. Allerdings erfolgt für hohe Germaniumanteile dann eine stark ausgeprägte Sättigung. Eine Begründung hierfür liefert die Abhängigkeit der Bandlücke von der Germaniumkonzentration: So weist die Energiebandlücke in linearer Skalierung qualitativ denselben Verlauf wie die dargestellten Ströme in logarithmischer Skalierung auf.

Abschließend bleibt noch zu klären, welchen Einfluss die genaue Lage der SiGe-Region auf das Verhalten des Tunneltransistors hat. Weil eine Germaniumkonzentration über 60% keine weitere Verbesserung liefert, ist es für die Analyse ausreichend, wie zu Beginn des Abschnittes, einen Tunneltransistor im n-Betrieb mit einem 50%-igen Germaniumanteil in der Sourcereion zu betrachten. Die aus den durchgeführten Simulationen folgenden Ströme im Durchlassbereich bei einer Gatespannung von 2 V sind auf der linken Seite von Abbildung 4.41 in Abhängigkeit von dem Überlapp zwischen SiGe-Sourcereion und Gatestapel dargestellt. Wird dieser Überlapp größer als 2 nm, so sättigt der Strom und zusätzlich wird bei einer weiteren Erhöhung der Sperrstrom zunehmen und damit werden sich die Schalteigenschaften des Tunneltransistors letztlich verschlechtern. Dieser Effekt spielt aber erst bei extremen Kurzkanalbauelementen eine signifikante Rolle, da dann die SiGe-Region auch das Gebiet betrifft, in dem das drainseitige Band-zu-Band Tunneln

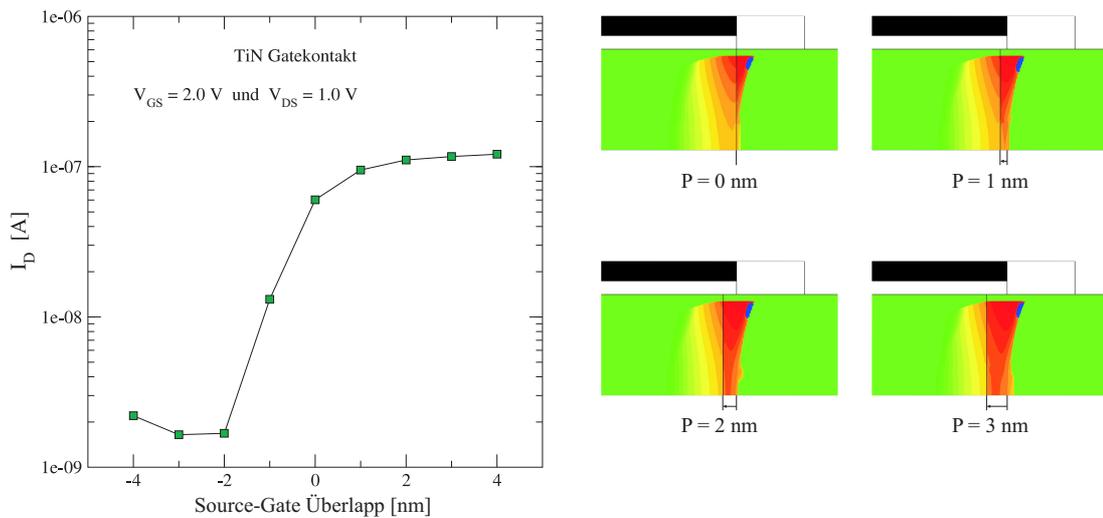


Abb. 4.41: Durchlass-Ströme des CNW-TFETs in Abhängigkeit von dem Überlapp der SiGe-Sourcereion mit dem Gate (links) und die jeweiligen Band-zu-Band Tunnelraten bei $V_{DS}=1$ V und $V_{GS}=2$ V (rechts).

stattfindet. Allerdings ist es die Absicht einen Tunneltransistor mit geringer Kanallänge und dennoch einer guten Transfercharakteristik zu entwerfen. Im Bereich zwischen einem Abstand von 2 nm und einem Überlapp von 1 nm ist eine ausgeprägte Veränderung des Stroms zu beobachten. Auf der rechten Seite der Abbildung 4.41 sind Konturbilder der Tunnelrate und die jeweilige Position der SiGe-Region für vier unterschiedliche Varianten dargestellt. Hieraus ist zu erkennen, dass die Tunnelregion bei Bauelementen bis zu einem Überlapp von 2 nm nahezu vollständig in dem SiGe-Gebiet liegt. Bei einer Verringerung des Überlapps befindet sich ein immer geringerer Anteil der Tunnelregion in SiGe und der Strom nimmt ab. Sobald der Abstand zwischen Gatestack und SiGe-Region mehr als 2 nm beträgt, findet das Band-zu-Band Tunneln nur noch in Silizium statt und damit hat der Strom im Durchlassbereich bei einem Abstand von 3 nm ein Minimum. Die Ursache ist die in diesem Bereich auftretende Verbreiterung der Tunnelbarriere durch den Einbau von SiGe. Die Zunahme des Stromes bei einer weiteren Abstandsvergrößerung ist ab 6 nm beendet und das Stromniveau des reinen Siliziumbauelements ist erreicht.

4.5 Optimierungstrategien bei zylindrischen Nanodraht-Tunneltransistoren

In den vorangehenden Abschnitten wurden mit Hilfe von Bauelementsimulationen die einzelnen Auswirkungen unterschiedlicher Geometrie- und Dotierungsparameter auf das Schaltverhalten von zylindrischen Nanodraht-Tunneltransistoren untersucht. Mit Hilfe der daraus abgeleiteten Erkenntnisse wird nun ein optimiertes Bauelement entworfen. Die einzelnen Resultate der Analyse von zylindrischen Nanodraht-Tunneltransistoren sind in Abbildung 4.42 graphisch aufbereitet. Hierbei ist der Einfluss auf die Kennlinien, der

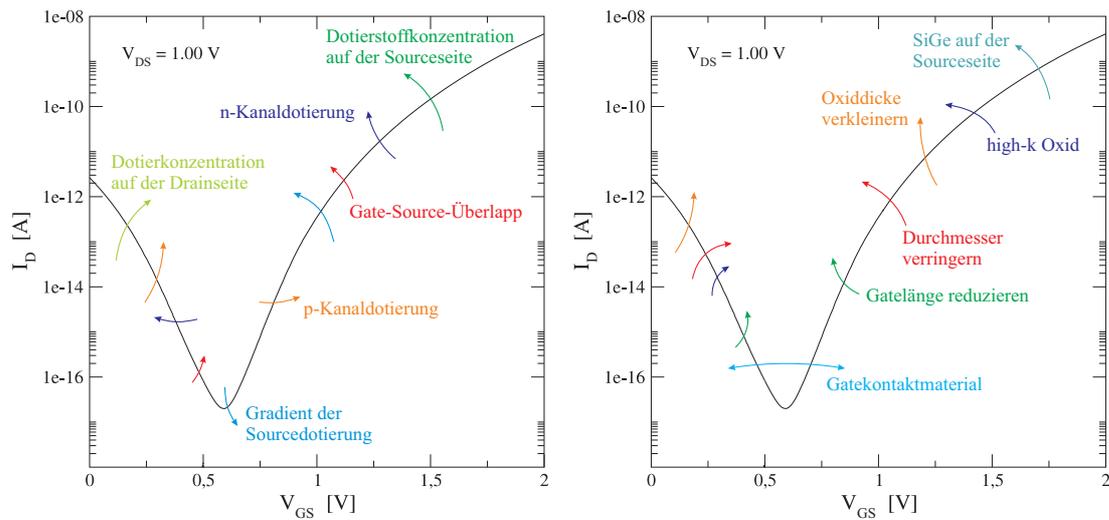


Abb. 4.42: Zusammenfassung der einzelnen Optimierungsstrategien von CNW-TFETs bei einer Variation der Dotierungen (links) und der gewählten Materialien bzw. der Skalierungseigenschaften (rechts).

durch die Dotierprofile verursacht wird, im linken Teil zusammengefasst. So sollte die Dotierstoffkonzentration auf der Drainseite möglichst gering sein und gleichzeitig ist eine hohe Dotierung der Sourcereion anzustreben. Zudem ist ein großer Gradient der Sourcedotierung und ein Überlapp zwischen Sourcereion und Gatestapel sinnvoll. Der Überlapp sollte allerdings gerade bei Bauelementen mit sehr kleinen Gatelängen nicht zu groß ausfallen. Mit Hilfe der Dotierstoffkonzentration der Kanalregion kann letztlich eine Feinjustierung der Einsatzspannung vorgenommen werden. Hier sind besonders niedrige Bordotierungen empfehlenswert. Die in der Abbildung eingezeichneten Pfeile deuten in der Regel den Einfluss einer Erhöhung der entsprechenden Parameterwerte an. Werden die Geometrieparameter oder die eingesetzten Materialien variiert, so ergeben sich die in Abbildung 4.42 rechts schematisch dargestellten Änderungen der Transferkennlinien des zylindrischen Tunneltransistors. Dabei bleibt festzuhalten, dass die Reduzierung der Abmessungen unter Umständen sogar eine Verbesserung der Bauelementeigenschaften bewirkt. Durch die Implementierung von Materialien mit niedrigen Energiebandlücken in der Sourcereion werden die Schalteigenschaften von Tunneltransistoren im n-Betrieb signifikant verbessert. Ein Einbau in der Drainregion führt zu besseren Bauelementen im p-Betriebszustand. Werden sogenannte high-k Oxidmaterialien eingesetzt, so können die Ströme im Durchlassbereich erhöht werden, ohne unerwünschte Leckströme zwischen Kanal und Gatekontakt zu erhöhen. Die Einsatzspannung kann durch maßgeschneiderte Gatekontaktmaterialien eingestellt werden.

Basierend auf diesen Ergebnissen wird nun ein optimierter zylindrischer Tunneltransistor zusammengestellt und mit Hilfe von Simulationen analysiert. Das Gate ist 5 nm lang und der Durchmesser der Kanalregion beträgt 30 nm. Hierbei besteht das 2 nm dicke Gateoxid aus HfO_2 und als Kontaktmaterial wird TiN im p-Betrieb und Aluminium im n-Betrieb eingesetzt. Im Fall des n-TFETs ist die Sourceseite mit einer Konzentration von 10^{20} cm^{-3}

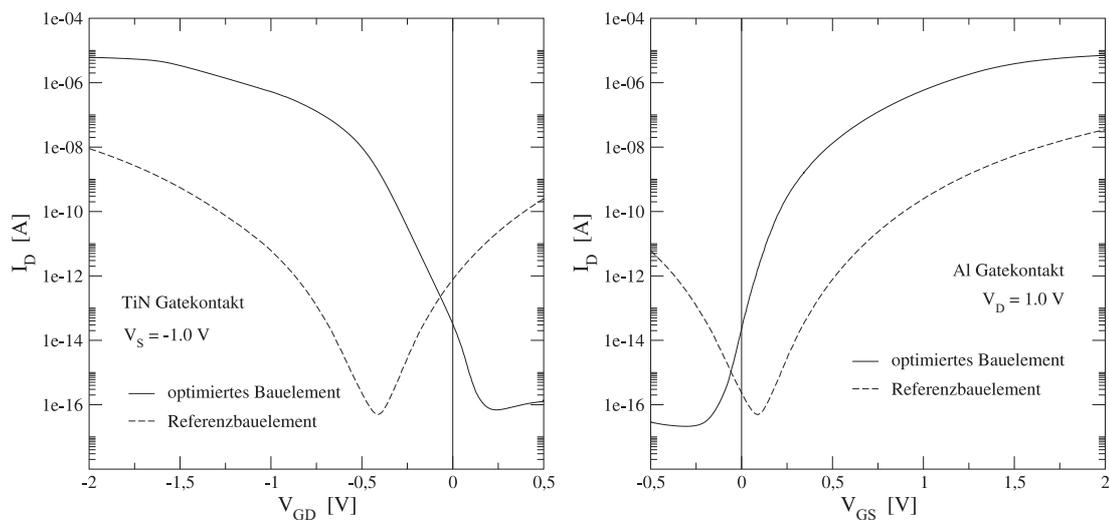


Abb. 4.43: Transfercharakteristiken von optimierten CNW-TFETs mit 5 nm Gatelänge im p- (links) und n-Betrieb (rechts) im Vergleich zu den Referenzbauelementen aus diesem Kapitel.

bor- und die Drainseite mit einer Konzentration von 10^{18}cm^{-3} arsen-dotiert. Der Kanal ist mit 10^{16}cm^{-3} schwach bordotiert und der Gradient der Dotierprofile beträgt auf der Source- und der Drainseite jeweils 2.5 nm/dec und folgt einem Gaußprofil. Der Abstand der Dotierprofilmaxima zum Gatestapel ist 5 nm , die Sourcereion besteht aus SiGe mit einem 50%-igen Germaniumanteil und hat einen 2 nm Überlapp mit dem Gatestapel. Im Fall des p-TFETs wird die Source- mit der Draindotierstoffkonzentration vertauscht und die SiGe-Region befindet sich auf der Drainseite. Alle weiteren Bauelementkenngrößen bleiben jedoch identisch. Aus der Simulation folgen die in Abbildung 4.43 dargestellten Transferkennlinien. Zwischen dem Sperr- und dem Durchlassbereich steigt der Strom um nahezu neun Größenordnungen an. Zudem ist die Steigung der Kennlinie im n-Betrieb 50 mV/dec , und die Ströme im Durchlassbereich sind mit etwa 10^{-5} A relativ hoch. Aus Gründen der Konvergenz sind Tunneltransistoren mit höheren Dotierstoffkonzentrationen nicht untersucht worden. Hiervon ist allerdings eine zusätzliche Steigerung der Ströme im Durchlassbereich zu erwarten.

4.6 Analyse von dreidimensional strukturierten Tunneltransistoren

Bisher lag in diesem Kapitel der Schwerpunkt auf der Untersuchung von zylindrischen Nanodraht-Tunnelfeldeffekttransistoren. Allerdings sind diese nicht besonders leicht zu fertigen. Daher werden in diesem Abschnitt einige einfacher herzustellende Bauelemente betrachtet: Der Fokus liegt hierbei zunächst auf Doppelgatestrukturen und anschließend werden sogenannte Quadgatebauelemente analysiert. Prinzipiell können alle Ergebnisse aus den letzten Abschnitten auch bei dem Entwurf von Multigatestrukturen herangezogen

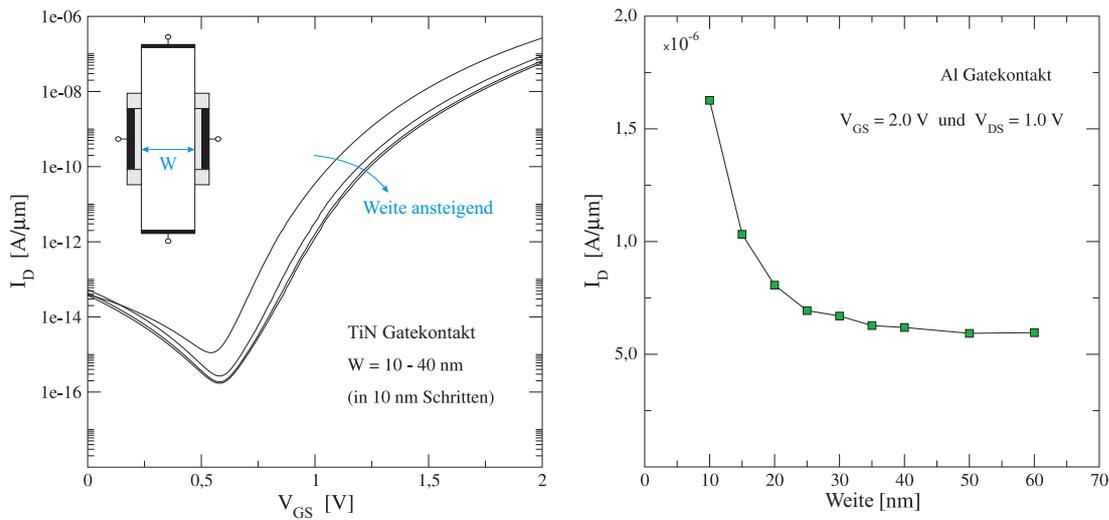


Abb. 4.44: Simulierte Transferkennlinien eines DG-TFETs (links) und Durchlass-Ströme in Abhängigkeit von der Kanalweite (rechts).

werden. Die einzige Ausnahme stellt die Abhängigkeit der Bauelementeigenschaften von dem Durchmesser dar. Anstelle des Durchmessers treten im Fall quaderförmiger Kanäle dann die Breite und Höhe. Im letzten Teilabschnitt werden die verschiedenen Varianten der Tunneltransistoren miteinander verglichen.

4.6.1 Doppelgate als Vorstufe zu Nanodraht-TFETs

Tunnelfeldeffekttransistoren in einer Doppelgateausführung sind in planarer Technologie schichtweise aufgebaut. Allerdings ist die Kontaktierung der unteren Gateplatte und die Positionierung der beiden Gatekontakte zueinander nicht einfach [58]. Der nachfolgend untersuchte DG-TFET hat dieselben Dotierprofile und bis auf den Durchmesser dieselben Geometriedaten wie der zylindrische Tunneltransistor in Abschnitt 4.3.1 (Analyse der Auswirkungen der Gatelänge auf die Schalteigenschaften). Zudem ist abweichend vom Referenzbauelement die Draindotierstoffkonzentration $5 \cdot 10^{18} \text{ cm}^{-3}$ und die Spacerweite auf der Sourceseite beträgt null Nanometer. In Abbildung 4.44 ist links der schematische Aufbau des DG-TFETs und die aus den Simulationen resultierenden Transferkennlinien für unterschiedliche Weiten des Siliziumkanals dargestellt. Aus diesen Simulationen ist abzulesen, dass ab einer Weite von etwa 30 nm keine signifikanten Unterschiede zwischen den einzelnen Bauelementen erkennbar sind. Auf der rechten Seite der Abbildung sind die Ströme zwischen Source- und Drainkontakt im Durchlassbetrieb in Abhängigkeit von der Kanalweite zusammengefasst. Dabei fällt auf, dass zwischen 10 und 20 nm die Ströme noch stark variieren, aber letztlich ab 50 nm sättigen. Im Gegensatz zu den zylindrischen Tunneltransistoren steigen die Ströme nicht wieder an, da bei dieser Bauelementvariante die Gatefläche im Fall einer Erhöhung der Kanalweite konstant ist. Daher liegt der einzige Einfluss der Kanalweite in der wechselseitigen Verstärkung der elektrischen Felder und des daraus folgenden Anstiegs der Tunnelraten.

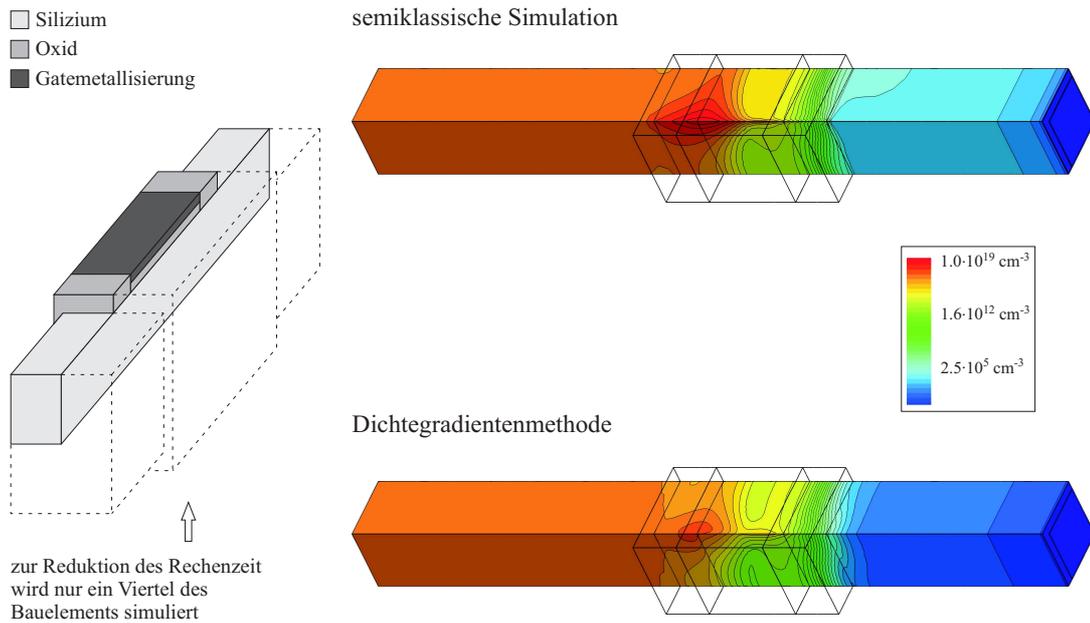


Abb. 4.45: Schematischer Aufbau mit grau hervorgehobenem Simulationsgebiet eines Quadgate-Tunneltransistors (links) und Elektronendichteverteilung im Kanal dieses Bauelements bei semiklassischer und quantenmechanisch korrigierter Simulation (rechts).

4.6.2 Eckeneffekte bei Quadgate-Nanodraht-TFETs

Werden dreidimensional strukturierte Tunneltransistoren betrachtet, so erhöht sich der erforderliche Rechenaufwand signifikant. Um dennoch vertretbare Simulationszeiten zu erhalten, können Symmetrieeigenschaften ausgenutzt werden: Im Fall eines sogenannten Quadgate-Transistors ist es daher ausreichend nur ein Viertel der Bauelementregion zu simulieren, wenn entsprechende Spiegelungsrandbedingungen angenommen werden. In Abbildung 4.45 ist auf der linken Seite dieser Sachverhalt schematisch illustriert. Dieses Vorgehen kann auch direkt auf Trigate-Tunneltransistoren übertragen werden. Wegen der guten Vergleichbarkeit zu zylindrischen Nanodraht-Tunneltransistoren und des kleineren Rechenaufwands werden im weiteren Verlauf dieses Abschnitts allerdings nur rechteckige Quadgate-Bauelemente betrachtet.

Den Ausgangspunkt bildet ein Tunnelfeldeffekttransistor mit einer Gatelänge von 10 nm und einer Kanalweite und -höhe von 20 nm. Der Gatestapel besteht hierbei aus einem 2 nm dicken Gateoxid aus SiO_2 und einem Gatekontakt aus Aluminium. Daran anschließend sind auf der Source- und Drainseite 10 nm lange Seitenoxide aus SiO_2 angefügt. Letztlich ist der Gatestapel auf allen vier Seiten des rechteckigen Kanals identisch. Das Substrat ist auf der Sourceseite $5 \cdot 10^{19} \text{cm}^{-3}$ bor- und auf der Drainseite 10^{18}cm^{-3} arsen-dotiert. Der Kanal ist mit 10^{16}cm^{-3} nur schwach bordotiert und der Gradient der Dotierprofile beträgt auf der Source- und der Drainseite jeweils 2 nm/dec und folgt einem Gaußprofil. Zudem beträgt der Abstand des Dotiermaximums auf der Sourceseite zum Gatestapel 5 nm und

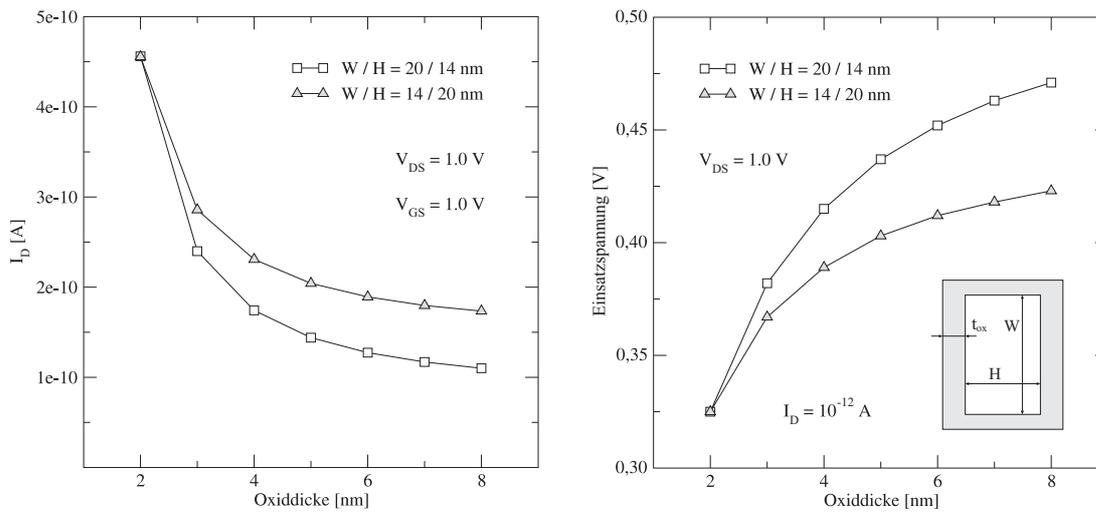


Abb. 4.46: Simulierte Ströme im Durchlassbereich eines Quadgate-Tunneltransistors bei der Variation der Oxiddicke auf zwei Seiten des Siliziumkanals (links) und dazugehörige Einsatzspannung (rechts) für zwei verschiedene Kanalprofile.

auf der Drainseite 15 nm. Auf diese Weise werden die Sperrströme gering gehalten. Die vergleichsweise niedrige Dotierstoffkonzentration auf der Sourceseite hat seine Ursache in den teilweise erheblich schlechteren Konvergenzeigenschaften der Simulationen im Fall hoher Dotierungen. Durch das Zusammenspiel mit dem dreidimensionalen Aufbau der Bauelemente verstärkt sich dieser Effekt. Ein Beheben dieses Problems ist ohne einen grundlegenden Eingriff in das Simulationswerkzeug nicht möglich. Bei einer angelegten Gatespannung von 1.5 V und einer Drainspannung von 1 V resultieren dann die auf der rechten Seite von Abbildung 4.45 dargestellten Elektronendichteverteilungen [59]. Dabei sind die Resultate einmal mittels einer semiklassischen Simulation und ein zweites Mal unter Verwendung der Dichtegradientenmethode erzielt worden. Ebenso wie im Fall des zylindrischen Nanodraht-Tunnelfeldeffekttransistors zeigt sich eine erhebliche Änderung der Dichteverteilungen und es folgt auch bei einer rechteckigen Tunneltransistorgeometrie ein deutlicher Einfluss auf die Bauelementcharakteristiken. Auf eine quantenmechanische Korrektur darf daher nicht verzichtet werden.

Eine erste Analyse von Quadgate-Tunneltransistoren betrifft deren Abhängigkeit von der Oxiddicke. Hierzu wird die Dicke des Gateoxids auf zwei gegenüberliegenden Seiten des Siliziumkanals variiert. Der Grund für diese Vorgehensweise liegt in der Absicht, auch Aussagen hinsichtlich von Trigate- beziehungsweise Fin-TFET-Strukturen zu treffen. So ist bei diesen Bauelementen aus fertigungstechnischen Gründen meist die Oxiddicke auf der Oberseite des Kanals größer als die der vertikalen Gateoxide an den Seitenflächen des Siliziumkanals. In Abbildung 4.46 ist der Einfluss auf die Ströme im Durchlassbereich des Tunneltransistors dargestellt. Hierbei werden die Resultate zweier unterschiedlicher Kanalprofile miteinander verglichen. Erwartungsgemäß nimmt in beiden Fällen der Strom bei einer Vergrößerung der Oxiddicke ab. Die Gründe für dieses Verhalten wurden bereits in Abschnitt 4.4.1 ausführlich erläutert. Es ist letztlich auch wenig überraschend, dass das

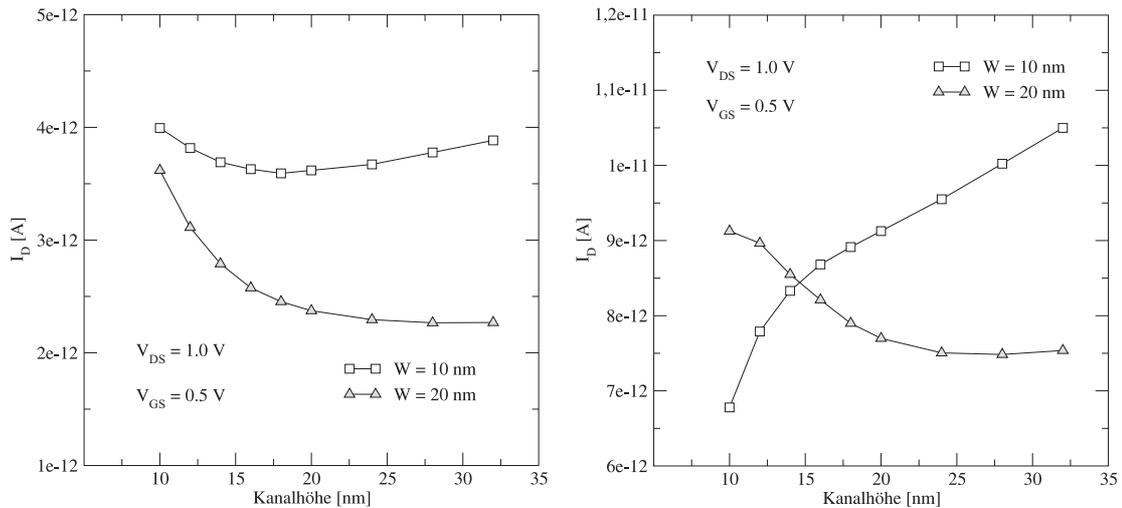


Abb. 4.47: Simulierte Ströme bei einer Gatespannung von 0.5 V in Abhängigkeit von der Kanalhöhe bei zwei unterschiedlichen Weiten des Kanals. Der Abstand des Dotierprofilmaximums auf der Sourceseite zum Gatestapel ist hierbei einmal 5 nm (links) und bei einer zweiten Simulationsreihe 0 nm (rechts).

Bauelement mit der größeren Höhe eine geringere Abhängigkeit von der Veränderung der Oxiddicke aufweist: Hinsichtlich der gesamten Gatefläche ist schließlich ein geringerer Anteil von der Änderung der Oxiddicke betroffen. Wird die Einsatzspannung des TFETs betrachtet, ergibt sich ein grundsätzlich ähnliches Bild: Die Einsatzspannung verschiebt sich mit steigender Dicke des Gateoxids zu höheren Werten und bei dem Tunneltransistor mit der größeren Kanalhöhe ist eine etwas geringere Abhängigkeit von der Veränderung der Oxiddicke zu beobachten. Als Einsatzspannung wird dabei der Gatespannungswert in der Transferkennlinie bezeichnet, bei dem der Strom durch das Bauelement einen Wert von 10^{-12} A erreicht. Für das Design von Trigate-Tunneltransistoren folgt letztlich, dass diese nach Möglichkeit höher als weit sein sollten.

In einer zweiten Simulationsreihe werden die Auswirkungen verschiedener Kanalprofile auf das Bauelementverhalten analysiert. Dazu werden die Kanalhöhe und -weite variiert und die Ströme bei unterschiedlichen Gatespannungen miteinander verglichen. Bei allen durchgeführten Simulationen beträgt die Drainspannung 1 V. Wird der Übergangsbereich zwischen Sperr- und Durchlassbetrieb bei einer Gatespannung von 0.5 V betrachtet, so folgen die in Abbildung 4.47 dargestellten Ergebnisse. Für Bauelemente deren Abstand des Dotierprofilmaximums auf der Sourceseite zum Gatestapel 5 nm beträgt, ergibt sich ein ähnliches Bild wie bei zylindrischen Nanodraht-Tunneltransistoren: So fällt die Kurve bei sehr kleinen Kanalhöhen wegen der starken wechselseitigen Beeinflussung der jeweils gegenüberliegenden Gatekontakte deutlich ab, hat dann in einem Übergangsbereich ein Minimum und steigt bei einer weiteren Vergrößerung der Höhe nahezu linear an. Der zweite Anstieg hat hierbei seine Ursache in der linearen Erhöhung der Gatefläche. Es fällt auf, dass bei den Tunneltransistoren mit einer kleineren Kanalweite dieser Effekt weniger ausgeprägt ist. Im Fall rechteckiger Nanodraht-Tunneltransistoren dominiert prinzipiell

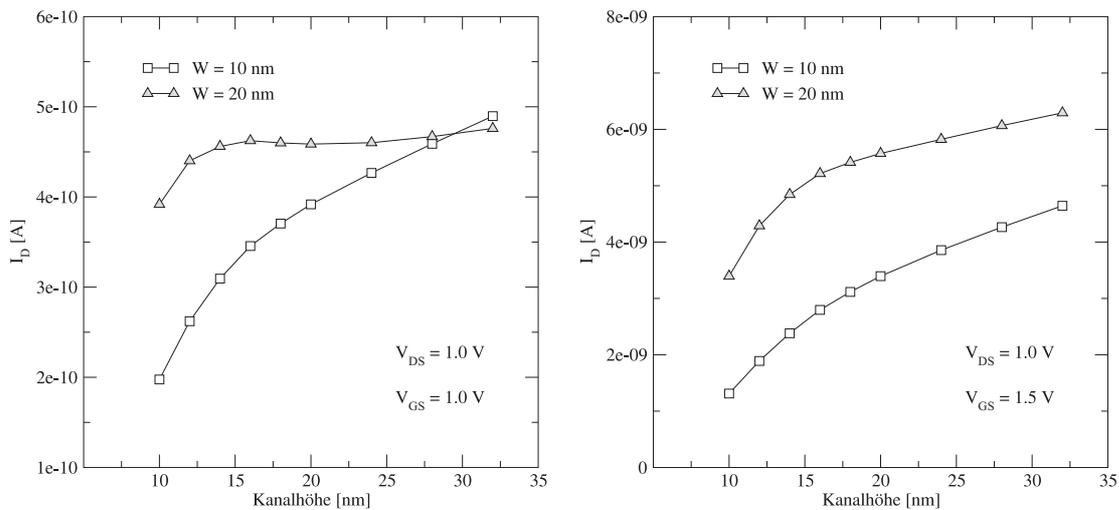


Abb. 4.48: Simulierte Ströme in Abhängigkeit von der Kanalhöhe bei zwei verschiedenen Kanalweiten und unterschiedlichen angelegten Gatespannungen. Der Abstand des Dotierprofilmaximums auf der Sourceseite zum Gatestapel beträgt in allen Fällen 0 nm.

die kleinere Kantenlänge der Kanalregion. Wenn der Abstand des Dotierprofilmaximums auf der Sourceseite zum Gatestapel 0 nm beträgt, folgt eine interessante Abweichung von dem vorangehend beschriebenem Verhalten der Bauelemente. Für eine genaue Analyse sind in Abbildung 4.48 zudem die Drainströme in Abhängigkeit von der Kanalhöhe bei Gatespannungen von 1 V beziehungsweise 1.5 V aufgetragen. Werden die Bauelemente mit einer Weite von 20 nm untersucht, so zeigt sich, dass bei niedrigen Spannungen die voranstehend beschriebene Abhängigkeit des Drainstroms von der Kanalhöhe besteht. Im Fall einer Erhöhung der Gatespannung verändert sich die Charakteristik: So nimmt bei einer Spannung von 1.5 V (siehe Abbildung 4.48 rechts) der Strom bei einer Vergrößerung der Kanalhöhe monoton zu. Die Ursache für dieses Verhalten liegt darin begründet, dass im Fall eines Anstiegs der Gatespannung das Tunneln auch für größere Kanalhöhen in der gesamten Querschnittsfläche stattfindet. Zusätzlich erfolgt bei Bauelementen mit einer kleinen Kanalfläche eine Sättigung der Transferkennlinien und damit des Drainstromes bei hohen Gatespannungen.

Abschließend bleibt zu klären, welchen Einfluss der Abstand des Dotierprofilmaximums auf der Sourceseite zum Gatestapel hat. Im Fall eines geringeren Abstands verschiebt sich die Einsatzspannung zu höheren Werten, da die Ladungsträgerdichten in der Tunnelregion anfangs zu niedrig sind, um einen relevanten Tunnelstrom zu ermöglichen. Zudem stellt sich die Frage, weshalb bei den zylindrischen Nanodraht-Tunneltransistoren eine solche Änderung der Abhängigkeiten von dem Durchmesser auch bei hohen Gatespannungen nicht zu beobachten war. Zum Einen wurde bei diesen Bauelementen ein Spacerabstand von 5 nm gewählt. Zusätzlich war die Sourceseite mit einer Konzentration von 10^{20} cm^{-3} bordotiert. In Abbildung 4.49 sind auf der linken Seite die simulierten Drainströme in Abhängigkeit von der Kanalhöhe für einen rechteckigen Nanodraht-Tunneltransistor bei

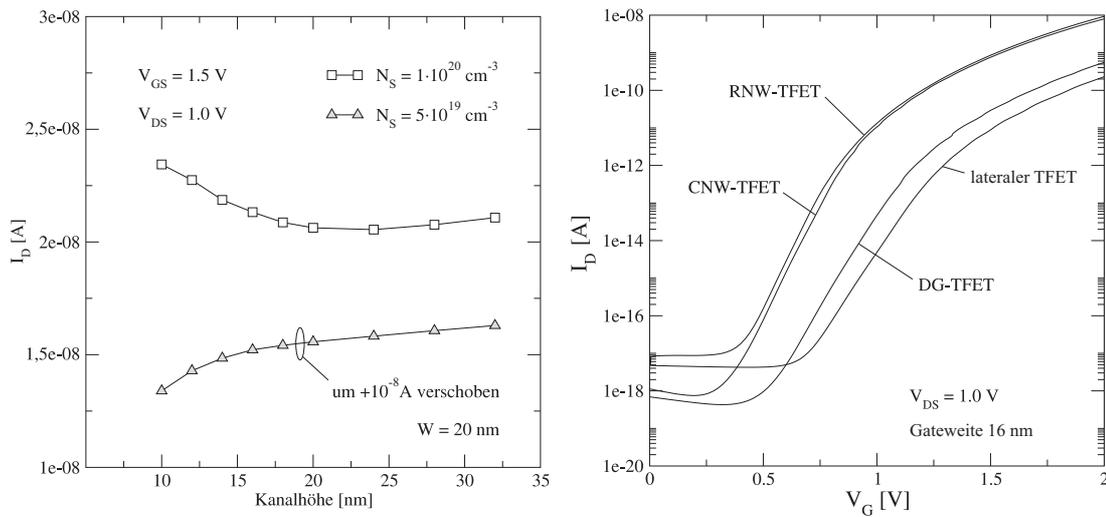


Abb. 4.49: Vergleich der Abhängigkeit der Ströme von der Dotierstoffkonzentration auf der Sourceseite bei unterschiedlichen rechteckigen Kanalprofilen (links) und Transferkennlinien bei verschiedenen Tunneltransistorvarianten (rechts).

einer Gatespannung von 1.5 V und zwei unterschiedlichen Dotierstoffkonzentrationen in der Source region dargestellt. Hierbei zeigt sich, dass das bei zylindrischen Bauelementen beschriebene Verhalten auch für rechteckige Tunneltransistoren bei hinreichend hohen Sourcedotierstoffkonzentrationen zu beobachten ist. Im Fall hoher Dotierungen nimmt der Tunnelstrom zu und die Sättigung in der Transfercharakteristik tritt erst bei größeren Gatespannungswerten auf. Daher resultiert dann der Anstieg des Drainstromes bei kleinen Kanalquerschnittsflächen.

4.6.3 Vergleich unterschiedlicher Querschnittstypen

Für die endgültige Bewertung der Optimierungsmöglichkeiten der Schalteigenschaften von Tunneltransistoren fehlt nur noch der Vergleich der verschiedenen Bauformen. Zu diesem Zweck werden vier Tunneltransistorvarianten mit einer Gatelänge von 10 nm und einer Kanalweite von 16 nm untersucht. Die Dotierprofile und der Aufbau der Gatestapel sind hierbei äquivalent zu denen aus dem zweiten Absatz des letzten Abschnitts. Um in den Transferkennlinien auch den Sperrbereich besser zu erkennen, wird allerdings TiN als Gatekontaktmaterial eingesetzt. Die Ergebnisse dieser Simulationsreihe sind auf der rechten Seite von Abbildung 4.49 zusammengestellt. Der Doppelgate-Tunneltransistor ist im Hinblick auf die Schalteigenschaften ein wenig besser als die laterale Variante. So ist der Sperrstrom niedriger und der Strom im Durchlassbereich ist höher. Grundsätzlich fällt auf, dass die Nanodraht-Tunneltransistoren noch höhere Ströme im Durchlassbereich aufweisen. Zudem nimmt die Steigung der Kennlinien bei diesen Bauformen zu. Wegen dieser Eigenschaften sind Multigate-Tunneltransistoren zu bevorzugen. Die nur geringen Unterschiede zwischen den zylindrischen und den rechteckigen Tunneltransistoren legen nahe, dass die erheblich einfacher zu fertigenden Bauelemente mit einem rechteckigen

Kanalquerschnitt die bessere Wahl darstellen. Abschließend sei darauf hingewiesen, dass alle Aussagen zur Optimierung von zylindrischen Nanodraht-Tunneltransistoren auch auf die rechteckigen Bauformen anwendbar sind.

5. Verbesserte Modellierung von Band-zu-Band Tunnelprozessen

Die Ziele dieses Kapitels sind die kritische Analyse des Band-zu-Band Tunnelmodells aus Kapitel 2.4.2 und die Möglichkeiten dieses zu verbessern. Den Ausgangspunkt bildet dabei die Feststellung, dass das ursprüngliche Modell nach Schenk [6] einen numerischen Approximationsfehler enthält und daher einer Korrektur bedarf. Zu diesem Zweck wird im ersten Abschnitt zunächst der Fehler identifiziert. Hierbei ist die fehlerhafte Näherung des Airy-Integrals die Ursache der, für die Bauelementesimulation entwickelten, falschen Tunnelrate. Danach wird eine geeignete Näherung für Airy-Integrale vorgestellt und ein korrigiertes Modell für das Band-zu-Band Tunneln abgeleitet. Dieses wird im Rahmen des *physical model interface* von Sentaurus Device [38] implementiert und anschließend werden die hieraus folgenden Auswirkungen auf die Simulationen von Tunneltransistoren diskutiert. Der zweite Abschnitt behandelt die Frage, inwieweit das Modell innerhalb der Bauelementesimulation vereinfacht werden kann, ohne dabei die Ergebnisse nennenswert zu beeinträchtigen. Außerdem wird in einem kleinen Exkurs auf die Möglichkeiten der Beschreibung von Tunnelbauelementen auf Kompaktmodellebene eingegangen.

5.1 Korrektur der Band-zu-Band Tunnelmodelle

Bevor in diesem Abschnitt auf Fehler in der Rate des indirekten phononenunterstützten Band-zu-Band Tunnelns eingegangen wird, ist es sinnvoll die wichtigsten Gleichungen nochmals in einer geeigneten Form bereitzustellen. Den Ausgangspunkt der Analyse stellt die Generationsrate des Modells nach Schenk aus Gleichung (2.59) dar:

$$R_t^{ind} = A^* \cdot D^* \cdot F^2 \cdot \sum_{\alpha} \sqrt{m_{\perp}^{\alpha} m_{\parallel}^{\alpha}} \cdot \left[f_B H(x_{\alpha}^{\mp}) + (f_B + 1) H(x_{\alpha}^{\pm}) \right] \quad (5.1)$$

In diesem richtungsabhängigen Modell ist f_B die Bose-Einstein-Verteilungsfunktion der mittleren Phononenbesetzungszahl, F der Betrag der elektrischen Feldstärke und D^* der von den Ladungsdichten abhängige Vorfaktor aus Gleichung (2.66). Wird nun mit Hilfe der Annahme (2.67) ein richtungsunabhängiges Modell aufgestellt, so folgt

$$R_t^{ind} = C \cdot D^* \cdot F^2 \cdot \left[f_B H(x^{\mp}) + (f_B + 1) H(x^{\pm}) \right] \quad (5.2)$$

dabei ist C aus A^* analog zu B aus A in (2.69) abgeleitet. Diese Modellgleichung für das Band-zu-Band Tunneln wird in den folgenden Abschnitten analysiert. Hierbei gilt die

Abkürzung

$$H(x) = \frac{\text{Ai}(x)}{x^2} + \frac{\text{Ai}'(x)}{x} + \text{Ai}_1(x) \quad (5.3)$$

mit der Airy-Funktion $\text{Ai}(x)$, der Ableitung $\text{Ai}'(x)$ der Airy-Funktion nach x und dem folgendermaßen definierten Airy-Integral

$$\text{Ai}_1(x) = \int_x^\infty \text{Ai}(y) dy \quad (5.4)$$

Die Abhängigkeit der Variablen x^\pm in der Funktion H von der kritischen Feldstärke F_c^\pm aus Gleichung (2.63) und der elektrischen Feldstärke F wird beschrieben durch

$$x^\pm = \left(\frac{3 F_c^\pm}{2 F} \right)^{2/3} \quad (5.5)$$

Diese Formeln zur Beschreibung der Band-zu-Band Tunnelrate sind im Rahmen der von Schenk getroffenen physikalischen Annahmen soweit richtig. Grundsätzlich ist dieses Modell in der vorliegenden Form für die Bauelementesimulation einsetzbar. Allerdings besteht das Problem, dass die Funktion H nur numerisch gelöst werden kann und, allein wegen der Komplexität der Lösung des Airy-Integrals, der rechentechnische Aufwand bei jeder einzelnen Auswertung der Tunnelrate inakzeptabel hoch wird.

5.1.1 Fehleranalyse beim bislang verwendeten Tunnelmodell

Um den rechentechnischen Aufwand zu verringern, verwendet Schenk [6] die folgende Approximation

$$H(x) \simeq \frac{x^{-9/4}}{2\sqrt{\pi}} \cdot \exp \left[-\frac{2}{3} x^{3/2} \right] \quad (5.6)$$

und somit resultiert letztlich auch eine Näherungslösung für die Rekombinationsrate

$$R_t^{\text{ind}} \simeq C \cdot \left(\frac{2}{3} \right)^{3/2} \cdot \frac{1}{2\sqrt{\pi}} \cdot D^* \cdot F^{7/2} \cdot \left\{ f_B (F_c^\mp)^{-3/2} \exp \left[-\frac{F_c^\mp}{F} \right] + \right. \\ \left. + (f_B + 1) (F_c^\pm)^{-3/2} \exp \left[-\frac{F_c^\pm}{F} \right] \right\} \quad (5.7)$$

Dieses Modell ist beispielsweise im Bauelementesimulator Sentaurus Device [38] von Synopsys implementiert. Bedauerlicherweise ist das Modell in dieser Form aber leider nicht mit Gleichung (5.2) identisch und somit falsch. Wird die Funktion $H(x)$ mit hoher Genauigkeit numerisch ausgewertet und mit der Näherung (5.6) verglichen, so resultiert die in Abbildung 5.1 dargestellte Situation. Die Abbildung ist im semilogarithmischen Maßstab skaliert und für die Bauelementesimulation ist nur der hervorgehobene Bereich von Bedeutung. Die obere Grenze ergibt sich aus der Beobachtung, dass die elektrische Feldstärke in Tunnelbauelementen wenige MV/cm nicht überschreitet. Zum anderen ist

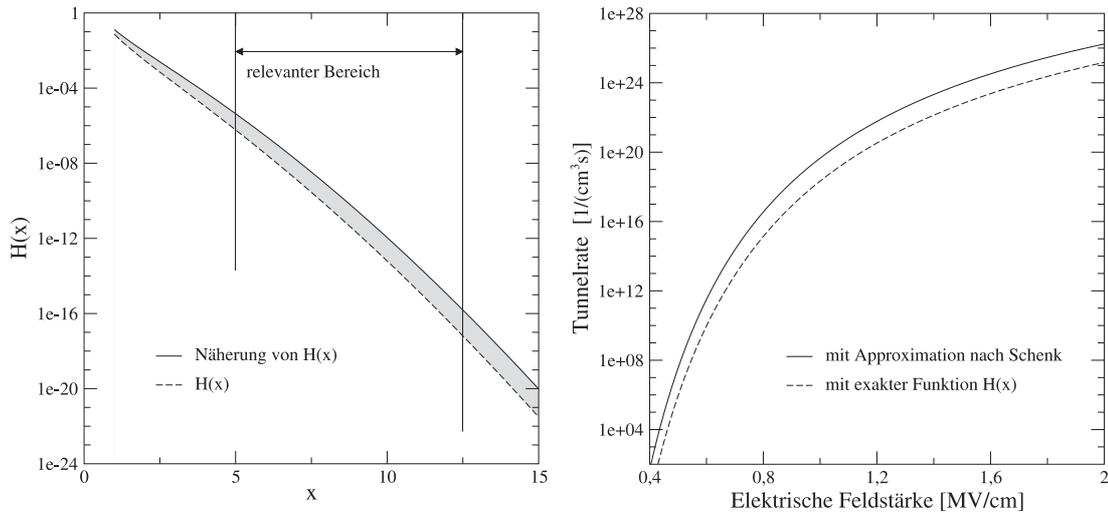


Abb. 5.1: Darstellung der exakten analytischen Funktion $H(x)$ und der Näherungsformel nach Schenk (links) und Auswirkungen auf die feldabhängige Band-zu-Band Tunnelgenerationsrate (rechts).

im Fall kleiner Feldstärken die Generationsrate so gering, dass dann dem Tunnelstrom keine physikalische Relevanz mehr beigemessen werden kann. Der relative Fehler der Approximation in diesem Bereich liegt zwischen einer und zwei Größenordnungen und ist damit nicht akzeptabel. Auf der rechten Seite der Abbildung sind die dazugehörigen Tunnelraten in Abhängigkeit von der elektrischen Feldstärke aufgetragen.

Bevor eine Lösung dieses Problems möglich ist, bleibt zu klären, auf welche Weise eine solche signifikante Abweichung zustandekommt. In Schenk [6] wird die Airy-Funktion durch die Reihenentwicklung [60] beschrieben

$$\text{Ai}(x) \simeq \left(1 - \frac{5}{48}x^{-3/2}\right) \cdot \frac{x^{-1/4}}{2\sqrt{\pi}} \cdot \exp\left[-\frac{2}{3}x^{3/2}\right] \quad \text{für } x \rightarrow \infty \quad (5.8)$$

Gleichermaßen gilt für die Reihenentwicklung der Ableitung der Airy-Funktion [60]

$$\text{Ai}'(x) \simeq -\left(1 + \frac{7}{48}x^{-3/2}\right) \cdot \frac{x^{1/4}}{2\sqrt{\pi}} \cdot \exp\left[-\frac{2}{3}x^{3/2}\right] \quad \text{für } x \rightarrow \infty \quad (5.9)$$

Diese beiden Reihen gelten streng genommen nur für große Werte von x , allerdings sind die Fehler im hier betrachteten Definitionsbereich ausreichend gering. Als Näherung für das Airy-Integral wird dann

$$\text{Ai}_1(x) \simeq \left(1 + \frac{7}{48}x^{-3/2}\right) \cdot \frac{x^{-3/4}}{2\sqrt{\pi}} \cdot \exp\left[-\frac{2}{3}x^{3/2}\right] \quad \text{für } x \rightarrow \infty \quad (5.10)$$

eingesetzt. Diese Approximation ist jedoch falsch. Wie dieser Fehler genau entstanden ist kann leider nicht nachvollzogen werden. Die Auswirkungen auf die Approximation der

Funktion $H(x)$ sind aber fatal, da sich der Term mit der Ableitung der Airy-Funktion und der mit dem Airy-Integral nun gegenseitig exakt aufheben. Damit wird die Funktion $H(x)$ nur durch die Reihenentwicklung der Airy-Funktion beschrieben. Bei der Herleitung der Tunnelstromdichte war die Beziehung $\text{Ai}'(x) + x\text{Ai}_1(x)$ explizit erforderlich. Unter der Annahme, dass die Näherungsformel (5.10) für das Airy-Integral stimmt, ist dann der Tunnelstrom immer Null.

5.1.2 Näherungsmethoden für Airy-Integrale

Nachdem mit der fehlerhaften Approximationsformel für das Airy-Integral die Ursache der Abweichungen in der Tunnelrate gefunden ist, muss eine korrekte Beschreibung für dieses bestimmt werden. In einem ersten Schritt werden die vollständigen asymptotischen Reihenentwicklungen der Airy-Funktion und ihrer Ableitung angegeben. Wird die Reihe für die Airy-Funktion nach n Elementen abgebrochen, so folgt die Näherungsformel

$$\text{Ai}^n(x) = \frac{x^{-1/4}}{2\sqrt{\pi}} \cdot \exp\left[-\frac{2}{3}x^{3/2}\right] \cdot \sum_{k=0}^n c_k \cdot \left(-\frac{3}{2} \cdot x^{-3/2}\right)^k \quad (5.11)$$

mit den Koeffizienten

$$c_k = \frac{(6k-1)!!}{6^{3k} \cdot k! \cdot (2k-1)!!} \quad (5.12)$$

und der sogenannten Doppelfakultät

$$k!! = \begin{cases} k \cdot (k-2) \cdot \dots \cdot 1 & \text{für } k > 0, \text{ ungerade} \\ k \cdot (k-2) \cdot \dots \cdot 2 & \text{für } k > 0, \text{ gerade} \\ 1 & \text{für } k = -1, 0 \end{cases} \quad (5.13)$$

Für die ersten vier Koeffizienten gilt dann

$$c_0 = 1, \quad c_1 = \frac{5}{72}, \quad c_2 = \frac{385}{10368}, \quad c_3 = \frac{85085}{2239488}, \quad \dots \quad (5.14)$$

wobei auffällt, dass die höheren Koeffizienten relativ klein sind. Es ergibt sich dieselbe Reihe wie in Gleichung (5.8). Für die asymptotische Reihenentwicklung der Ableitung der Airy-Funktion kann ein äquivalenter Ausdruck wie für die Reihe der Airy-Funktion gefunden werden. Wenn nun diese Reihe nach n Elementen abgebrochen wird, so folgt die Näherungsformel

$$\text{Ai}'^n(x) = -\frac{x^{1/4}}{2\sqrt{\pi}} \cdot \exp\left[-\frac{2}{3}x^{3/2}\right] \cdot \sum_{k=0}^n d_k \cdot \left(-\frac{3}{2} \cdot x^{-3/2}\right)^k \quad (5.15)$$

mit den Koeffizienten

$$d_k = -c_k \cdot \frac{6k+1}{6k-1} \quad (5.16)$$

Das negative Vorzeichen der Reihe und das der einzelnen Koeffizienten heben sich zwar auf, jedoch wird zur besseren Vergleichbarkeit mit Gleichung (5.8) diese Schreibweise bevorzugt. Damit folgen als Zahlenwerte für die ersten vier Koeffizienten

$$d_0 = 1, \quad d_1 = -\frac{7}{72}, \quad d_2 = -\frac{455}{10368}, \quad d_3 = -\frac{95095}{2239488}, \quad \dots \quad (5.17)$$

Somit verbleibt die Aufgabe eine asymptotische Reihenentwicklung für das Airy-Integral zu finden. Nach Nikishov und Ritus [61] gilt

$$\text{Ai}_1^n(x) = \frac{x^{-3/4}}{2\sqrt{\pi}} \cdot \exp\left[-\frac{2}{3}x^{3/2}\right] \cdot \sum_{k=0}^n a_k \cdot \left(-\frac{3}{2} \cdot x^{-3/2}\right)^k \quad (5.18)$$

mit den Koeffizienten

$$a_k = c_k + \left(k - \frac{1}{2}\right) \cdot a_{k-1} \quad \text{mit} \quad a_0 = 1 \quad (5.19)$$

Diese sind zur besseren Lesbarkeit in Abhängigkeit von den entsprechenden Koeffizienten der Reihenentwicklung der Airy-Funktion c_k in rekursiver Form dargestellt. Für die Werte der ersten vier Koeffizienten gilt dann

$$a_0 = 1, \quad a_1 = \frac{41}{72}, \quad a_2 = \frac{9241}{10368}, \quad a_3 = \frac{5075225}{2239488}, \quad \dots \quad (5.20)$$

Werden die drei Formeln für die asymptotischen Reihenentwicklungen (5.11), (5.15) und (5.18) miteinander verglichen, so fällt auf, dass die Form dieser identisch ist. Die einzigen Unterschiede sind die Koeffizienten c_k , d_k beziehungsweise a_k und die vorangestellten Potenzterme. So steht zu erwarten, dass in der resultierenden Reihe der Funktion $H(x)$ ebenso eine kompakte Darstellung gefunden werden kann und sich eventuell bestimmte Elemente herauskürzen. In Abbildung 5.2 sind auf der linken Seite die relativen Fehler der Reihenentwicklung der Airy-Integrale in Abhängigkeit von der Zahl der verwandten Elemente der Reihenentwicklung dargestellt. Mit steigender Anzahl an Reihenelementen nimmt der relative Fehler in einem weiten Bereich signifikant ab. Der Fehler wird aber für Werte $x < 2$ sehr groß. Allerdings ist dieser Bereich für das Band-zu-Band Tunnelmodell nicht relevant.

Werden die drei asymptotischen Reihenentwicklungen in Gleichung (5.3) eingesetzt, so resultiert die Reihe

$$H_{se}^n(x) = \frac{x^{-3/4}}{\sqrt{\pi}} \cdot \exp\left[-\frac{2}{3}x^{3/2}\right] \cdot \sum_{k=0}^n b_k \cdot (-x^{-3/2})^k \quad (5.21)$$

wobei der Index n die Zahl der berücksichtigten Elemente der prinzipiell unendlichen Reihe angibt. Man beachte die etwas andere Form dieser Reihenentwicklung. Für die Koeffizienten gilt

$$b_k = \frac{1}{2} \cdot \left(a_k - d_k - \frac{2}{3} \cdot c_{k-1}\right) \cdot \left(\frac{3}{2}\right)^k \quad \text{mit} \quad b_0 = 0 \quad (5.22)$$

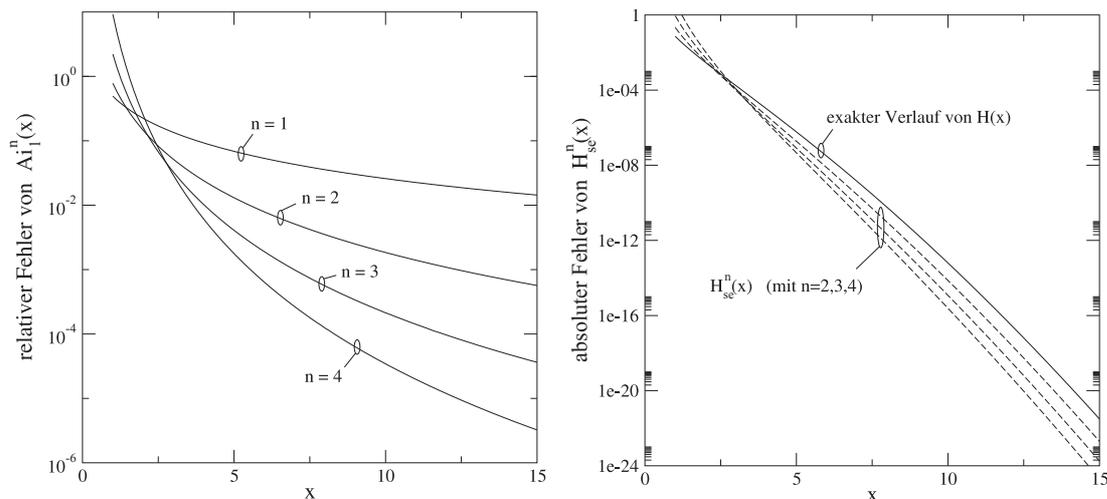


Abb. 5.2: Relativer Fehler der Reihenentwicklung des Airy-Integrals nach Nikishov und Ritus [61] (links) und die resultierenden absoluten Fehler der Funktion $H_{sc}^n(x)$ in Abhängigkeit von der Zahl der berücksichtigten Reihenelemente (rechts).

Auch diese werden mittels einer rekursiven Formel berechnet, da jede andere Darstellung wesentlich unübersichtlicher ist. Für die Zahlenwerte der ersten vier Koeffizienten gilt dann

$$b_0 = 0, \quad b_1 = 0, \quad b_2 = 1, \quad b_3 = \frac{185}{48}, \quad \dots \quad (5.23)$$

Auf der rechten Seite der Abbildung 5.2 sind die absoluten Fehler der Reihenentwicklung in Abhängigkeit von der Zahl der berücksichtigten Reihenelemente im Vergleich zum exakten Verlauf der Funktion $H(x)$ gezeigt. Hierbei zeigt sich allerdings, dass eine hohe Zahl an Elementen erforderlich ist, um eine gute Approximation zu erzielen.

5.1.3 Ableitung eines korrigierten Band-zu-Band Tunnelmodells

Die im letzten Teilabschnitt berechnete Näherungsformel (5.21) für die Funktion $H(x)$ ist erheblich besser als die Ausgangsnäherung (5.6) nach Schenk [6]. Allerdings wird bei dieser Approximation der Fehler bei hohen elektrischen Feldstärken, das heißt bei kleinen Werten von x , verhältnismäßig groß und deshalb wird nachfolgend versucht, eine für die vorliegende Aufgabenstellung besser geeignete Reihenentwicklung zu finden. Nach dem Artikel von Nikishov und Ritus [61] gilt

$$\int_x^\infty y^{-n} \text{Ai}(y) dy = -x^{-n-1} \cdot \text{Ai}'(x) - (n+1) \cdot x^{-n-2} \cdot \text{Ai}(x) + (n+1) \cdot (n+2) \cdot \int_x^\infty y^{-n-3} \text{Ai}(y) dy \quad (5.24)$$

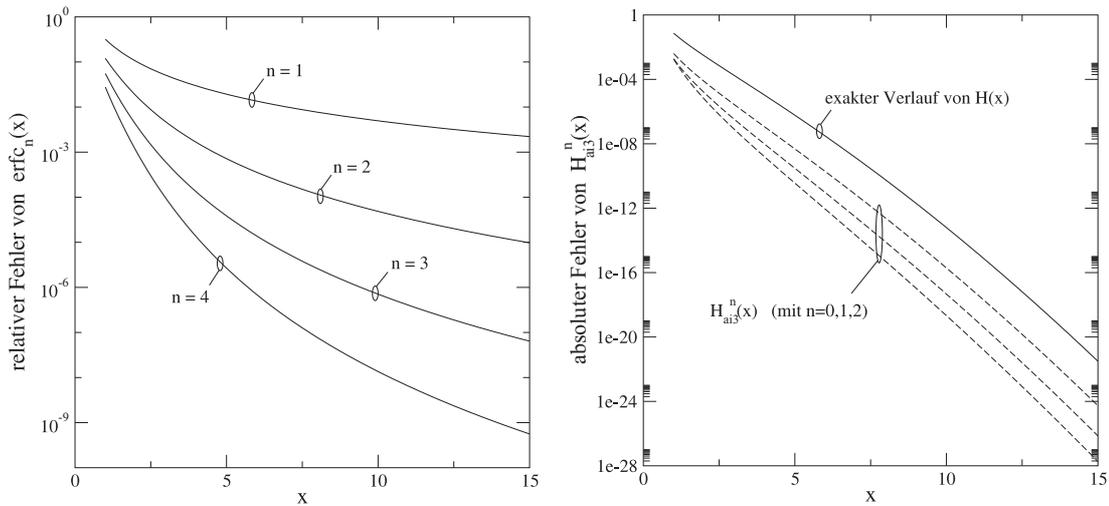


Abb. 5.3: Relativer Fehler der Kettenbruchnäherung der Fehlerfunktion (links) und der daraus resultierende absolute Fehler der Funktion $H_{ai3}^n(x)$ in Abhängigkeit von der Zahl der berücksichtigten Reihenelemente (rechts).

und es folgt im Fall $n = 0$ die Beziehung

$$H(x) = 2 \cdot \int_x^\infty \frac{\text{Ai}(y)}{y^3} dy \quad (5.25)$$

In diese Gleichung wird dann die Reihenentwicklung der Airy-Funktion eingesetzt und anschließend komponentenweise integriert. Auf diese Weise resultieren die Näherungen $H_{ai3}^n(x)$ in Abhängigkeit von der Ordnung der eingesetzten Airy-Reihenentwicklung. So folgt durch Einsetzen der Beziehung $\text{Ai}^0(x)$ in diese Gleichung beispielsweise

$$H_{ai3}^0(x) = \frac{1}{\sqrt{\pi}} \cdot \int_x^\infty y^{-13/4} \cdot \exp\left[-\frac{2}{3}y^{3/2}\right] dy \quad (5.26)$$

Mit der Abkürzung $z = 2/3 \cdot x^{3/2}$ ergibt sich dann der Ausdruck

$$H_{ai3}^0(x) = \sqrt{6} \cdot \frac{16}{81} \cdot \text{erfc}(\sqrt{z}) + \frac{1}{\sqrt{\pi}} \cdot \left(\frac{2}{3}\right)^{7/2} \cdot \exp(-z) \cdot \left[z^{-3/2} - 2z^{-1/2}\right] \quad (5.27)$$

Die hieraus resultierenden absoluten Fehler der Funktionen $H_{ai3}^n(x)$ in Abhängigkeit von der Zahl der berücksichtigten Reihenelemente sind in Abbildung 5.3 auf der rechten Seite im Vergleich zur exakten Funktion $H(x)$ dargestellt. Im Gegensatz zu den Näherungen aus der Reihenentwicklung nach Gleichung (5.21) ist im relevanten Bereich ein erheblich geringerer Fehler zu beobachten. Zudem nimmt der Fehler bei kleinen Werten von x nur unwesentlich zu. Selbst im Fall der einfachsten Näherung nach Gleichung (5.27) bleibt der relative Fehler geringer als ein Prozent.

Allerdings sind diese Näherungslösungen nicht direkt für die Bauelementesimulation zu verwenden, da in jedem Fall komplementäre Fehlerfunktionen enthalten sind. Und diese sind nur numerisch auswertbar. Daher ist bisher noch kein ernsthafter Fortschritt erzielt und es muss nach einer geeigneten Beschreibungsmöglichkeit der Fehlerfunktion gesucht werden. Mit der Kettenbruchapproximation gilt

$$\operatorname{erfc}(x) \approx \operatorname{erfc}_n(x) = \frac{\exp(-x^2)}{\sqrt{\pi}} \cdot \frac{1}{x + \frac{1}{2x + \frac{1}{x + \frac{2}{2x + \dots}}}} \quad (5.28)$$

Somit folgt für die ersten drei Näherungslösungen

$$\operatorname{erfc}(x) \approx \operatorname{erfc}_0(x) = \frac{\exp(-x^2)}{\sqrt{\pi}} \cdot \frac{1}{x} \quad (5.29)$$

$$\operatorname{erfc}(x) \approx \operatorname{erfc}_1(x) = \frac{\exp(-x^2)}{\sqrt{\pi}} \cdot \frac{1}{x + \frac{1}{2x}} \quad (5.30)$$

$$\operatorname{erfc}(x) \approx \operatorname{erfc}_2(x) = \frac{\exp(-x^2)}{\sqrt{\pi}} \cdot \frac{1}{x + \frac{1}{2x + 2/x}} \quad (5.31)$$

Der relative Fehler der Kettenbruchapproximation der komplementären Fehlerfunktion ist in Abbildung 5.3 (links) für unterschiedliche Grade der Näherung dargestellt. So fallen die Fehler sehr schnell bei einer Erhöhung der Zahl der Kettenbruchelemente. Im Fall der zweiten Näherung der Fehlerfunktion $\operatorname{erfc}_2(x)$ sind hierbei die relativen Fehler über den gesamten relevanten Bereich von x kleiner als ein Promille. Wird nun $\operatorname{erfc}_2(x)$ in $H_{a;3}^0(x)$ aus Gleichung (5.27) eingesetzt, so folgt

$$H_2^0(x) = \frac{1}{\sqrt{\pi}} \cdot \frac{x^{-9/4}}{x^{3/2} + 9/4} \cdot \exp\left(-\frac{2}{3}x^{3/2}\right) \quad (5.32)$$

Mit dieser Näherung ist eine erste sinnvolle Lösung der Problemstellung gefunden. Es kann schließlich eine weitere Reduktion der Fehler durch Terme höherer Ordnung erzielt werden, doch führt dies zu wesentlich komplexeren Band-zu-Band Tunnelraten und wird dann wegen der größeren Zahl an notwendigen Rechenoperationen bei jeder Auswertung der Tunnelraten insgesamt weniger attraktiv. Abschließend wird versucht, mit Hilfe einer numerischen Parametervariation, eine optimierte Lösung zu finden. Zu diesem Zweck wird in die Gleichung (5.32) der Parameter a eingeführt und es gilt

$$H_2^0(x, a) = \frac{1}{\sqrt{\pi}} \cdot \frac{x^{-9/4}}{x^{3/2} + a} \cdot \exp\left(-\frac{2}{3}x^{3/2}\right) \quad (5.33)$$

In Abbildung 5.4 sind auf der linken Seite der relative Fehler der Approximation $H(x, a)$ aus dieser Gleichung und das Integral über den relativen Fehler in Abhängigkeit vom

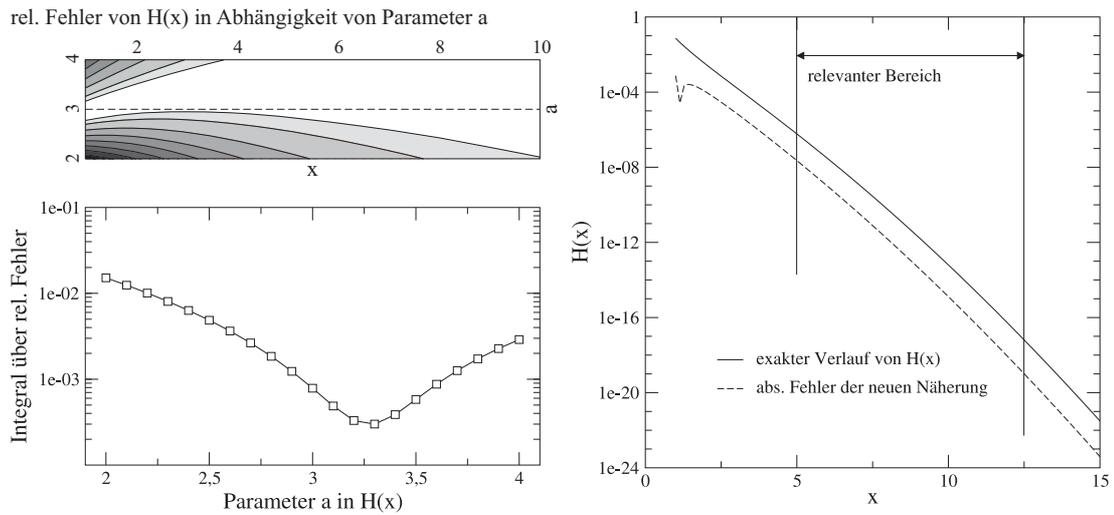


Abb. 5.4: Relativer Fehler der Approximation $H(x, a)$ aus Gleichung (5.33), Integral über den relativen Fehler in Abhängigkeit vom Parameter a und Darstellung des absoluten Fehlers der letztlich ausgewählten Approximation von $H(x)$ aus Gleichung (5.34) im Vergleich zu der exakten Funktion.

Parameter a gezeigt. Das obere Konturbild zeigt den relativen Fehler in Abhängigkeit von unterschiedlichen Parameterwerten. Dunkle Flächen weisen auf einen hohen und helle auf einen niedrigen Fehler hin. Wird zusätzlich der aufintegrierte Fehler in der unteren Abbildung betrachtet, so scheint ein Wert $a = 13/4$ angebracht. Da aber für $a > 3$ der Fehler bei kleinen Werten von x deutlich ansteigt, fällt die Wahl auf $a = 3$ und es gilt

$$H(x) \approx \frac{1}{\sqrt{\pi}} \cdot \frac{x^{-9/4}}{x^{3/2} + 3} \cdot \exp\left(-\frac{2}{3}x^{3/2}\right) \quad (5.34)$$

Der absolute Fehler dieser Approximation von $H(x)$ ist im Vergleich zu dem Verlauf der exakten Funktion auf der rechten Seite der Abbildung 5.4 dargestellt. So ist im Gegensatz zur Ausgangssituation ein geeignetes Beschreibungsmodell gefunden. Der Vergleich mit den Ergebnissen aus Abbildung 5.3 legt nahe, dass eine weitere Verbesserung durch mehr Reihenelemente in Gleichung (5.25) erzielt werden kann. Das ist aber nicht der Fall, da einer guten Näherung der Fehlerfunktion wesentlich mehr Gewicht zukommt. Wird nun eine höhere Genauigkeit benötigt, so kann mit Hilfe der Gleichungen (5.11), (5.25) und (5.28) schnell eine genauere Näherung bestimmt werden. Einsetzen von Gleichung (5.34) in Gleichung (5.2) liefert das neue Modell für die Band-zu-Band Tunnelrate [62]

$$R_t^{\text{ind}} \simeq C \cdot \left(\frac{2}{3}\right)^{5/2} \cdot \frac{1}{\sqrt{\pi}} \cdot D^* \cdot F^{9/2} \cdot \left\{ \frac{f_B}{(F_c^\mp)^{3/2} \cdot (F_c^\mp + 2F)} \cdot \exp\left[-\frac{F_c^\mp}{F}\right] + \frac{(f_B + 1)}{(F_c^\pm)^{3/2} \cdot (F_c^\pm + 2F)} \cdot \exp\left[-\frac{F_c^\pm}{F}\right] \right\} \quad (5.35)$$

Hierbei ist zu erkennen, dass im Vergleich zu dem ursprünglichen Modell aus Gleichung (5.7) die Komplexität ein wenig zugenommen hat. Allerdings ist das neue Modell ebenso

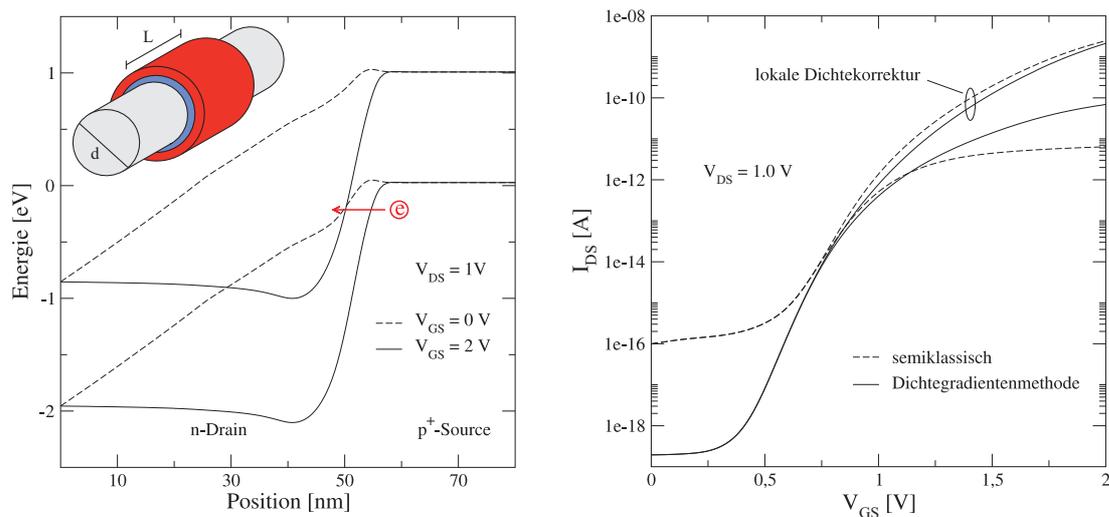


Abb. 5.5: Banddiagramm des untersuchten zylindrischen Tunneltransistors (links) und Transferkennlinien bei verschiedenen Modellierungsvarianten (rechts).

ohne Probleme in einen Bauelementsimulator implementierbar und führt auch zu keiner wesentlichen Erhöhung des Rechenaufwands. Dieser Schritt wird nachfolgend mit dem sogenannten *physical model interface* von Sentaurus Device vollzogen. Zu diesem Zweck wird ein zylindrischer Tunneltransistor mit 15 nm Durchmesser und einer Gatelänge von 10 nm betrachtet. Das Siliziumsubstrat ist auf der Source-Seite $5 \cdot 10^{19} \text{cm}^{-3}$ bordotiert und der Kanal, wie auch die Drainregion, sind konstant mit einer Konzentration von 10^{17}cm^{-3} arsen-dotiert. Der Gradient der Dotierstoffkonzentration der Source-Seite beträgt 2 nm/dec und folgt dabei einem Gaußprofil. Zudem beträgt der Abstand des Dotierstoffmaximums der Source-Region zum Gate-stapel 5 nm. Das Gateoxid ist 2 nm dick und besteht wie die Seitenoxide aus Siliziumdioxid. Als Gatekontaktmaterial wird TiN eingesetzt.

In einem ersten Schritt wird dieses Bauelement unter Verwendung verschiedener, bereits in Sentaurus Device implementierten, Modelle simuliert. Dabei werden die Simulationen sowohl mit quantenmechanischer Dichtekorrektur als auch semiklassisch durchgeführt und zusätzlich wird die lokale Dichtekorrektur ein- beziehungsweise ausgeschaltet. Das Banddiagramm des zylindrischen Tunneltransistors und die jeweiligen Transferkennlinien der Modellvarianten sind in Abbildung 5.5 dargestellt. Auf die physikalischen Ursachen wurde bereits in den letzten Kapiteln ausführlich eingegangen. Um die Möglichkeiten des *physical model interface* zu testen, werden in einem zweiten Schritt die ursprünglichen Band-zu-Band Tunnelmodelle aus Gleichung (5.7) erneut in den Bauelementsimulator eingebunden. Somit kann die korrekte Funktionsweise des PMI überprüft werden. Auf der linken Seite von Abbildung 5.6 werden die auf diese Weise erhaltenen Transferkennlinien mit den Referenzsimulationen verglichen. Damit die Abbildung nicht zu unübersichtlich wird, ist eine Beschränkung auf zwei Fälle vorgenommen worden. Insgesamt zeigt sich aber eine hervorragende Übereinstimmung der jeweiligen Kennlinien. Damit ist das PMI sehr gut geeignet, die verschiedenen Auswirkungen bei der Verwendung neuer Modelle auf Bauelementebene zu untersuchen. In dem hier vorliegenden Fall wurde anschließend

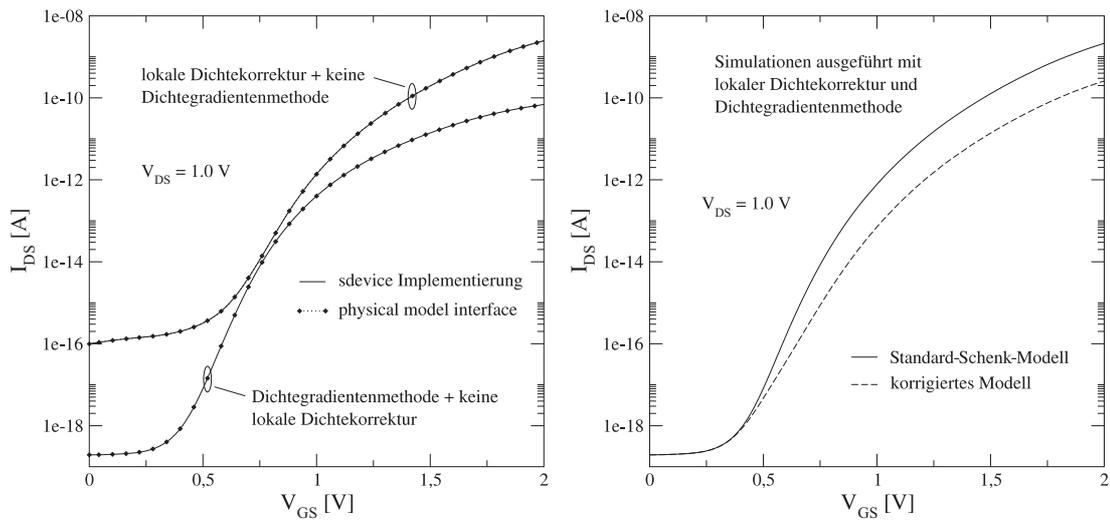


Abb. 5.6: Vergleich zweier Transferkennlinien bei Einsatz des in Sentaurus Device [38] eingebundenen Tunnelmodells nach Schenk [6] und der mittels des *physical model interface* implementierten Modelle (links). Auswirkung des korrigierten Tunnelmodells aus Gleichung (5.35) auf die simulierte Transferkennlinie des Referenz-Tunneltransistors (rechts).

das korrigierte Band-zu-Band Tunnelmodell nach Gleichung (5.35) implementiert. Auf der rechten Seite von Abbildung 5.6 sind die resultierenden Transferkennlinien der beiden Tunnelmodelle dargestellt. Hierbei bleibt festzuhalten, dass auch auf makroskopischer Ebene signifikante Unterschiede zwischen den beiden Modellen auftreten. Grundsätzlich sollte bei allen zukünftigen Simulationen, bei denen Band-zu-Band Tunnelprozesse das Verhalten von Halbleiterbauelementen beeinflussen, das hier vorgeschlagene korrigierte Modell eingesetzt werden. Deshalb ist aus praktischen Gesichtspunkten ein vollständiger Einbau in den Bauelementesimulator Sentaurus Device wünschenswert.

5.2 Alternative Modellkonzepte zur Beschreibung des Band-zu-Band Tunnels

Nachdem im letzten Abschnitt ein korrigiertes Beschreibungsmodell für Band-zu-Band Tunnelprozesse im Rahmen eines Bauelementesimulators aufgestellt wurde, ist es Ziel dieses Abschnitts verschiedene Möglichkeiten zur Verringerung des rechentechnischen Aufwands zu untersuchen.

5.2.1 Vereinfachte Beschreibung des Band-zu-Band Tunnels

Um letztendlich eine Vereinfachung der Tunnelrate aus Gleichung (5.35) zu erzielen, gibt es verschiedene Möglichkeiten. Eine erste Option ist die Vernachlässigung der effektiven

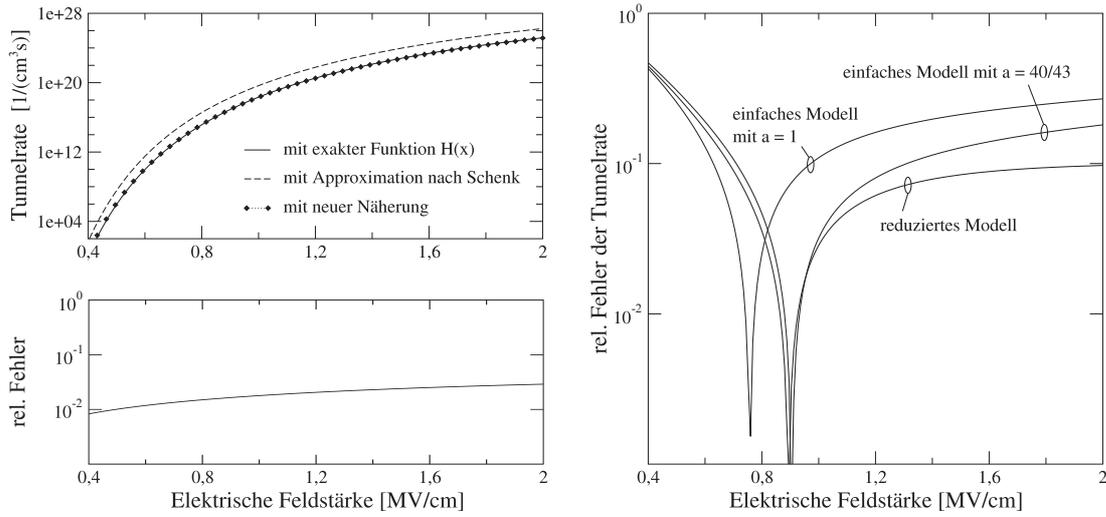


Abb. 5.7: Feldstärkeabhängigkeit der Band-zu-Band Tunnelrate und relativer Fehler der korrigierten Näherungslösung der Tunnelrate aus Gleichung (5.35) und relative Fehler des reduzierten Modells aus Gleichung (5.36) und zweier Varianten des einfachen Modells aus Gleichung (5.38).

Phononenenergie in der kritischen Feldstärke aus Gleichung (2.63). Dies führt zu einer modifizierten Rekombinationsrate der Form

$$R_t^{\text{ind}} \simeq C \cdot \left(\frac{2}{3}\right)^{5/2} \cdot \frac{1}{\sqrt{\pi}} \cdot D^* \cdot F^{9/2} \cdot \frac{2f_B + 1}{F_c^{3/2} \cdot (F_c + 2F)} \cdot \exp\left[-\frac{F_c}{F}\right] \quad (5.36)$$

Eine genauere Betrachtung der Gleichung (5.34) zeigt, dass die Funktion $H(x)$ einerseits exponentiell von x abhängt und zudem proportional zu $x^{-15/4}$ ist. Diese Proportionalität ist streng genommen nur für große Werte von x gültig. Im für die Tunnelrate relevanten Bereich von x ist allerdings der Fehler nicht besonders groß. Somit gilt näherungsweise

$$H_s(x) \approx \frac{1}{\sqrt{\pi}} \cdot x^{-15/4} \cdot a \cdot \exp\left(-\frac{2}{3}x^{3/2}\right) \quad (5.37)$$

mit dem Optimierungsparameter a , um eine möglicherweise auftretende Abweichung in der Feldstärkeabhängigkeit der Tunnelrate kompensieren zu können. Dann folgt für die Tunnelrate

$$R_t^{\text{ind}} \simeq C \cdot \left(\frac{2}{3}\right)^{5/2} \cdot \frac{1}{\sqrt{\pi}} \cdot D^* \cdot a \cdot F^{9/2} \cdot \frac{2f_B + 1}{F_c^{5/2}} \cdot \exp\left[-\frac{F_c}{F}\right] \quad (5.38)$$

Die Bewertung der vorgestellten alternativen Modellierungskonzepte ist in Abbildung 5.7 zusammengefasst. In dem Teilbild oben links ist die Abhängigkeit der Tunnelraten von der elektrischen Feldstärke abgebildet. Hierbei sind drei Fälle unterschieden: Einmal wird die Tunnelrate nach dem ursprünglichen Modell von Schenk aus Gleichung (5.7), dann mit dem in dieser Arbeit verbesserten Modell aus Gleichung (5.35) und schließlich nach

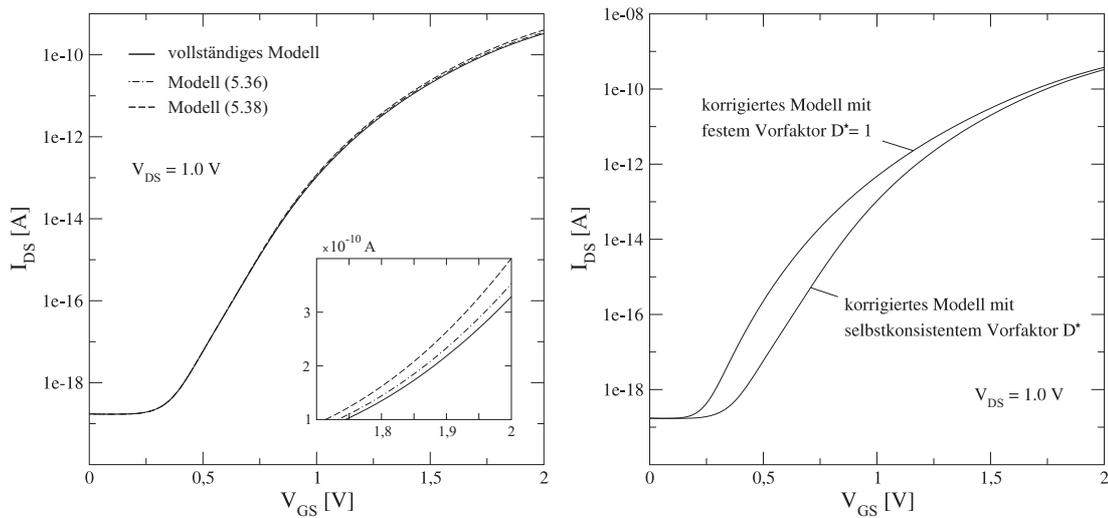


Abb. 5.8: Mit Hilfe des PMI (*physical model interface*) simulierte Transferkennlinien bei Vernachlässigung der Dichteabhängigkeit der Tunnelrate im Vergleich zum vollständigen Modell (links) und Transferkennlinien bei Verwendung der zwei vereinfachten Tunnelmodelle (rechts).

dem vollständigen Modell auf Basis der exakten Lösung der Funktion $H(x)$ berechnet. In dieser Darstellung ist kein Unterschied zwischen dem vollständigen und dem korrigierten Modell zu erkennen. Die Abbildung 5.7 unten links zeigt den relativen Fehler der Rate im Fall des korrigierten Modells aus Gleichung (5.35) und bestätigt die hohe Qualität der Näherung. Werden nun die relativen Fehler der vereinfachten Modelle betrachtet, so zeigt sich bei niedrigen elektrischen Feldstärken ein großer Fehler. Allerdings sind in diesem Bereich die Raten so niedrig, dass dem Band-zu-Band Tunneln im makroskopischen Sinn keine Bedeutung zukommt. Das reduzierte Modell nach Gleichung (5.36) stellt aber bei höheren Feldstärken eine halbwegs befriedigende Näherung dar. Im Fall des Modells aus Gleichung (5.38) erweist sich ein Parameterwert $a = 40/43$ noch als beste Wahl. Aber bei diesem Modell sind die Fehler insgesamt sehr hoch. Daher steht zu erwarten, dass auch die jeweiligen simulierten Ströme deutlich abweichen werden. Um diesen Sachverhalt zu untersuchen, wird erneut der Tunneltransistor aus dem letzten Abschnitt simuliert. Die auf diese Weise erhaltenen Transferkennlinien sind in Abbildung 5.8 auf der linken Seite dargestellt. So fällt auf den ersten Blick ein nur marginaler Unterschied zwischen den Kennlinien auf. Jedoch muss berücksichtigt werden, dass die Ströme semilogarithmisch abgebildet sind und selbst deutliche quantitative Abweichungen nicht sofort erkennbar sind. Zu diesem Zweck ist zusätzlich der Durchlassbereich in einem linearen Maßstab gezeigt. Hieraus folgen bei einer angelegten Gatespannung von 2 V im Strom zwischen Drain- und Sourcekontakt relative Fehler von 10 beziehungsweise 30%. Der qualitative Verlauf der Ströme wird aber in jedem Fall sehr gut wiedergegeben.

Im Hinblick auf die Entwicklung von Kompaktmodellen ist eine andere Vereinfachung der Tunnelraten von Relevanz. Wird der von den Ladungsdichten abhängige Vorfaktor D^* konstant auf eins gesetzt, reduziert sich das Modell auf ein nur von der elektrischen

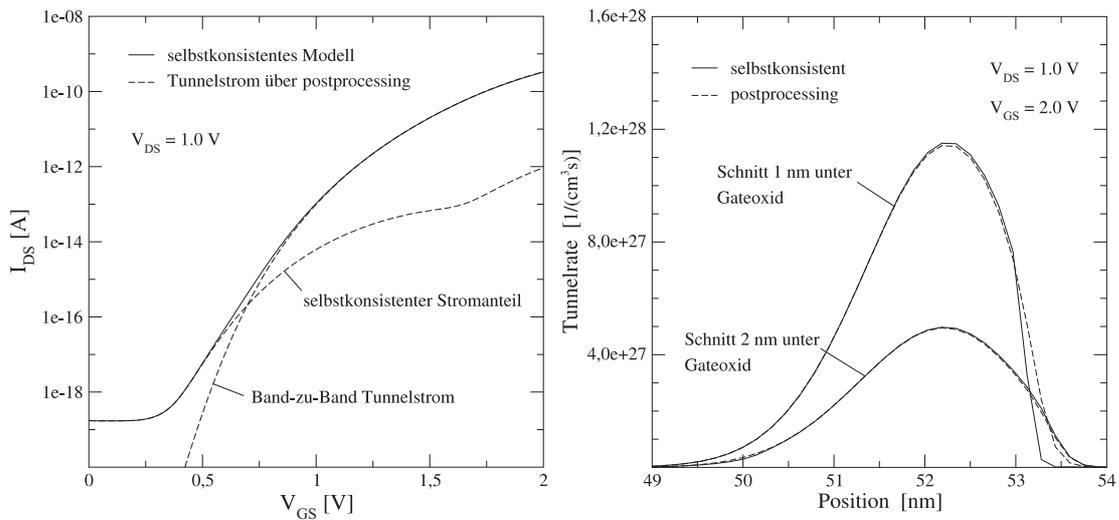


Abb. 5.9: Simulierte Stromkomponenten in der Transferkennlinie bei Anwendung einer *postprocessing*-Strategie im Vergleich zur selbstkonsistenten Lösung (links) und die zugehörigen Band-zu-Band Tunnelraten in zwei Schnittebenen unter dem Gateoxid (rechts).

Feldstärke und der Energiebandlücke abhängiges Modell. Für die Kompaktmodellierung ist dieser Ansatz deshalb so interessant, da dann nur noch eine sinnvolle Abschätzung des elektrischen Potentialverlaufs erforderlich ist. Aus diesem folgt auf einfache Weise die Feldstärke und damit auch die orts aufgelöste Tunnelrate. Das Problem einer solchen Vorgehensweise wird auf der rechten Seite von Abbildung 5.8 anschaulich illustriert. So ist im Übergangsbereich zwischen An- und Aus-Zustand eine extrem große Abweichung bei der Vernachlässigung der Ladungsträgerdichten in der Tunnelrate festzustellen. Der Grund für dieses Verhalten ist die Tatsache, dass der Vorfaktor die Verfügbarkeit von Elektronen und Löchern in der Tunnelregion beschreibt. Auf diese Weise wird bei nicht ausreichender Zahl an Ladungsträgern auch die Tunnelrate beschränkt und der durch das Bauelement fließende Strom ist in der Folge geringer.

Um den Rechenaufwand bei der Simulation von Tunneltransistoren zu verringern, kann auch eine sogenannte *postprocessing* Strategie eingesetzt werden. Hierbei werden die Grundgleichungen des Drift-Diffusions-Modells mit den Quantenkorrekturgleichungen selbstkonsistent gelöst, jedoch wird die Band-zu-Band Tunnelrate in einem ersten Schritt vernachlässigt. Ausgehend von der Kontinuitätsgleichung der Elektronen (2.16) gilt im stationären Fall unter Verwendung des Gaußschen Satzes für den Tunnelstrom

$$\nabla \mathbf{j}_n = q \cdot (R - G) \quad \Rightarrow \quad I_{DS}^{\text{tun}} = q \cdot \int_V R_t^{\text{ind}} dV \quad (5.39)$$

Hierbei wird angenommen, dass keine Leckströme durch das Gateoxid fließen. Aus der Addition des selbstkonsistenten Sperrstromes zwischen Source- und Drainkontakt mit dem aus der Formel (5.39) berechneten Tunnelstrom folgt dann der Gesamtstrom durch den Tunneltransistor. Eine Reduktion des Rechenaufwandes resultiert aus der erheblich

besseren Konvergenz des Gleichungssystems. In Abbildung 5.9 sind für den in diesem Kapitel betrachteten Tunneltransistor die einzelnen Stromkomponenten im Vergleich zu der vollständigen selbstkonsistenten Lösung dargestellt. Dabei fällt die außergewöhnlich gute Übereinstimmung zwischen diesen beiden Methoden auf. Auf der rechten Seite der Abbildung sind zusätzlich die jeweiligen Tunnelraten in zwei Schnittebenen unter dem Gateoxid gezeigt. Auch hier sind die Abweichungen nur marginal. Diese Vorgehensweise stellt somit eine interessante Alternative zur vollständigen, selbstkonsistenten Berechnung der Ströme dar. Allerdings ist darauf hinzuweisen, dass bei hohen Dotierkonzentrationen ein etwas größerer Fehler zu beobachten ist.

5.2.2 Simulation von Tunneltransistoren auf Kompaktmodellebene

Zur Beschreibung von Tunneltransistoren mittels Kompaktmodellen wird die Arbeit von Vandenberghe [63] herangezogen und die dort vorgestellten Resultate werden zudem für das in diesem Kapitel erarbeitete, korrigierte Band-zu-Band Tunnelmodell angepasst. Den Startpunkt bilden einige vereinfachende Annahmen: So wird das elektrische Potential in der Sourceregion als konstant angenommen und der Einfluss der freien Ladungsträger im Bauelement wird vernachlässigt. Der Strom berechnet sich hierbei über Gleichung (5.39) und die elektrische Feldstärke wird gemäß Gleichung (2.44) durch die Tunnelwegstrecke ersetzt. Damit ein homogenes Material vorliegt wird die effektive Oxiddicke

$$t'_{\text{ox}} = t_{\text{ox}} \cdot \frac{\varepsilon_{\text{Si}}}{\varepsilon_{\text{ox}}} \quad (5.40)$$

eingeführt. Mit der Näherung, dass die Feldlinien kreisförmig verlaufen, kann dann das elektrische Feld berechnet werden und es gilt

$$F = \frac{E_g}{qt} = \frac{E_g}{qr\vartheta_0} \quad (5.41)$$

Für die Berechnung werden zylindrische Koordinaten verwandt. Der Winkel zwischen zwei Äquipotentiallinien mit einem Potentialunterschied in der Größe der Bandlücke ist dann

$$\vartheta_0 = \frac{\pi \cdot E_g}{2 \cdot q \cdot \psi_g} = \frac{\pi \cdot E_g}{2 \cdot q \cdot (V_G - V_{\text{FB}})} \quad (5.42)$$

mit der Flachbandspannung V_{FB} und der angelegten Gatespannung V_G . Für den zwischen Source- und Drainkontakt fließenden Tunnelstrom gilt dann

$$I_{\text{DS}}^{\text{tun}} = q \cdot W \cdot \int_{r_0}^{\infty} (\arccos(t'_{\text{ox}}/r) - \vartheta_0) R_t^{\text{ind}}(r) r dr \quad (5.43)$$

Der Radius r_0 bezeichnet den Punkt ab dem sowohl der Start-, wie auch der Endpunkt der Tunnelwegstrecke im Siliziumsubstrat liegen. Aus diesem Grund müssen kleinere Radien bei der Integration nicht berücksichtigt werden. Die Größe W bezeichnet die Weite des Gatestapels. Da die Tunnelrate nur vom Radius abhängt, bleibt von dem ursprünglichen

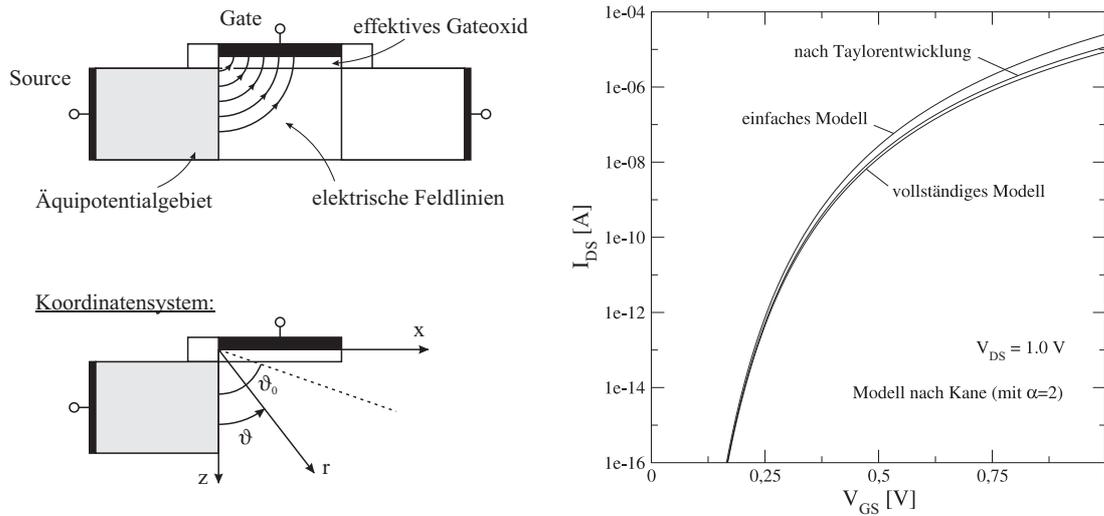


Abb. 5.10: Schematische Darstellung der von Vandenberghe [63] getroffenen Annahmen zur Ableitung eines Kompaktmodells für TFETs (links) und die resultierenden Transferkennlinien bei drei verschiedenen Modellvarianten unter Verwendung des Band-zu-Band Tunnelmodells nach Kane [38] (rechts).

Volumenintegral nur das Integral über den Radius übrig. In Abbildung 5.10 sind auf der linken Seite einige der getroffenen Annahmen schematisch dargestellt und zudem wird das für die Berechnung verwandte Koordinatensystem gezeigt. Vandenberghe verwendet die auf den Arbeiten von Kane [64] basierende Band-zu-Band Tunnelrate

$$R_t^{\text{Kane}} = A_{\text{Kane}} \cdot F^\alpha \cdot \exp \left[-\frac{B_{\text{Kane}}}{F} \right] \quad (5.44)$$

Nach einer Taylorreihenentwicklung des Integranden ohne die Exponentialfunktion um den Punkt $r = r_0$ und einer Beschränkung auf den linearen Anteil resultiert dann für den Tunnelstrom

$$I_{\text{DS}}^{\text{tun}} \approx \frac{W \cdot A_{\text{Kane}} \cdot E_g^{\alpha+2} \cdot t_{\text{ox}}'}{q^{\alpha+1} \cdot B_{\text{Kane}}^2 \cdot \vartheta_0^{\alpha+2} \cdot r_0^{\alpha+1}} \cdot \exp \left[-q B_{\text{Kane}} r_0 \vartheta_0 / E_g \right] \quad (5.45)$$

Dieser analytische Ausdruck für den Tunnelstrom kann nun als Kompaktmodell für einen Tunneltransistor eingesetzt werden. Eine weitere Vereinfachung ist möglich, wenn der Kosinusterm in $r_0 = t_{\text{ox}}' / \cos(\vartheta_0)$ in eine Sinusfunktion transformiert und anschließend durch die Kleinwinkelnäherung $\sin(x) \approx x$ ersetzt wird. Diese Näherung ist allerdings sehr ungenau, da die Annahme kleiner Winkel in dem untersuchten Fall nicht wirklich gültig ist. Auf der rechten Seite der Abbildung 5.10 sind die aus diesen Kompaktmodellen folgenden Transferkennlinien dargestellt: Als vollständiges Modell wird die numerische Lösung des Integrals aus Gleichung (5.43) mit der Tunnelrate aus (5.44) bezeichnet. Das zweite Modell ist durch Formel (5.45) definiert und das einfache Modell ergibt sich durch die beschriebene Kleinwinkelnäherung. Insgesamt zeigt sich, dass die Taylorentwicklung

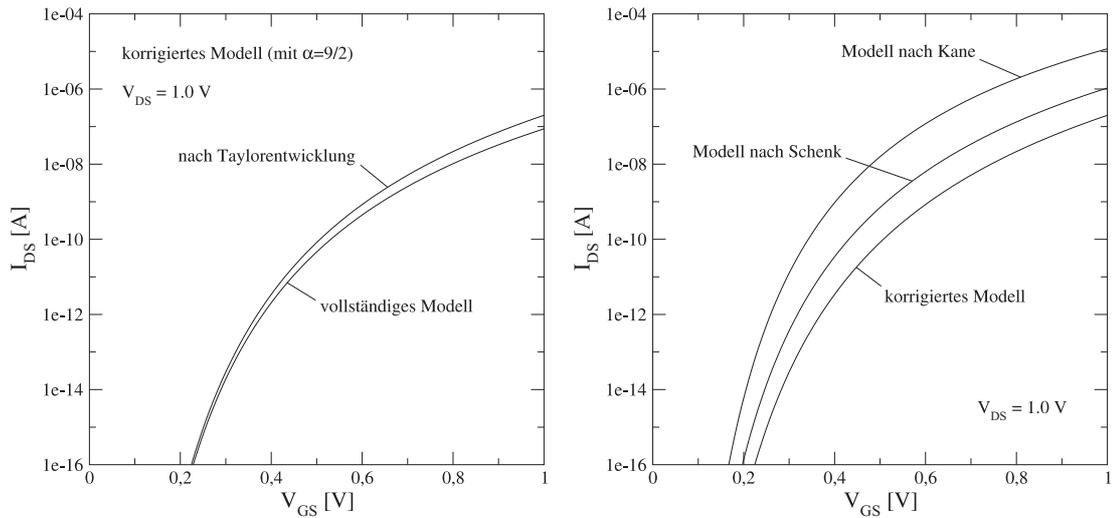


Abb. 5.11: Aus den Kompaktmodellgleichungen (5.46) und (5.47) analytisch berechnete Transferkennlinien eines Tunnelfeldeffekttransistors (links) und Vergleich der Auswirkungen verschiedener Band-zu-Band Tunnelmodelle (rechts).

eine annehmbare Näherung darstellt. Hingegen führt die von Vandenberghe eingesetzte Kleinwinkelnäherung erwartungsgemäß zu deutlichen Abweichungen und kann nicht als empfehlenswert eingestuft werden. Der betrachtete laterale Tunneltransistor hat dabei ein 2 nm dickes Gateoxid aus HfO_2 und die Gatelänge wird als so groß angenommen, dass die Bauelementeigenschaften nicht beeinflusst werden.

Wird anstelle der Tunnelrate nach Kane das in Abschnitt 5.2.1 vorgestellte Band-zu-Band Tunnelmodell aus Gleichung (5.36) verwandt und in Gleichung (5.43) eingesetzt, so folgt für den Tunnelstrom

$$I_{\text{DS}}^{\text{tun}} \approx W \cdot \frac{E_g^2 \cdot t'_{\text{ox}}}{q \cdot F_c^2 \cdot \vartheta_0^2 \cdot r_0} \cdot R_t^{\text{ind}} \Big|_{F=F_0} \quad (5.46)$$

mit

$$F_0 = \frac{E_g}{qr_0\vartheta_0} \quad (5.47)$$

Diese Formel ist eine etwas allgemeinere Form der Lösung von Vandenberghe [63]. Die Anwendung auf das vollständige Modell aus Gleichung (5.35) ist hierbei nicht weiter schwierig. Die Tunnelrate muss einfach in zwei von den jeweiligen kritischen Feldstärken abhängige Terme zerlegt und dann in die oben gegebene Formel eingesetzt werden. Auf der linken Seite von Abbildung 5.11 ist die Transferkennlinie dieses Kompaktmodells im Vergleich zur numerischen Lösung des Integrals aus Gleichung (5.43) mit der Tunnelrate aus Gleichung (5.36) gezeigt. Zudem ist rechts die Kennlinie des ursprünglichen Modells nach Schenk aus Gleichung (5.7) dargestellt. Hierbei zeigt sich erneut der signifikante Unterschied der Tunnelströme zwischen dem Ausgangsmodell für die Tunnelrate nach Schenk und der in dieser Arbeit vorgeschlagenen korrigierten Fassung.

Abschließend sind noch einige Anmerkungen erforderlich: So ist die Annahme, dass die Tunnelwegstrecke genau entlang der elektrischen Feldlinien verläuft, mit einem hohen

Fehler behaftet. Die Stromdichte verläuft unter dem Gatestapel nahezu parallel zu der Grenzfläche zwischen Gateoxid und Siliziumsubstrat. Daher müsste auch die Integration entsprechend angepasst werden. Auch in [65] findet sich dieser Hinweis, jedoch wird diese Idee dort nicht weiterverfolgt. Dieser Ansatz sollte noch einmal genauer untersucht werden. Darüberhinaus ist zu beachten, dass die Auswirkungen der freien Ladungsträger und der Dotierstoffkonzentrationen nicht zu vernachlässigen sind. Insgesamt stellt dieses Kompaktmodell aber eine interessante Basis für weitere Untersuchungen dar.

6. Zusammenfassung und Ausblick

Das Ziel dieser Arbeit ist die Bewertung der Eignung von Tunnelfeldeffekttransistoren für einen zukünftigen Einsatz in der CMOS-Technologie. Zum gegenwärtigen Zeitpunkt ist der MOSFET noch immer das zentrale Bauelement für Logikschaltungen. Wenn auch seit über dreißig Jahren wiederholt das Ende der Skalierbarkeit von MOSFETs vorhergesagt wird [66, 67], so ist dieses dennoch nicht erreicht. Allerdings treten durch die Reduktion der physikalischen Abmessungen vermehrt Kurzkanaleffekte auf. Auch die Problematik der diskreten Dotierstoffkonzentrationen in der Kanalregion und die damit verbundenen Schwankungen in dem Bauelementverhalten führt zusätzlich zu einer immer größeren Herausforderung bei Entwicklung von kleineren MOSFETs. Viele Entwicklungen, wie der Einsatz von high-k Materialien im Gatestapel [68], mechanischen Verspannungen im Kanal zur Erhöhung der Beweglichkeit [69] oder Multigate-Strukturen zur Erhöhung der Steuerwirkung des Gatekontaktes, stellen voraussichtlich die nächsten Jahre den Einsatz von MOSFETs sicher. Allerdings ist es in jedem Fall erforderlich, über diese Zeitspanne hinauszuplanen und alternative Bauelementkonzepte zu entwickeln. Gerade wegen der erwartungsgemäß immensen Kosten bei der Umstellung auf alternative Bauelemente und Materialsysteme sind hierbei siliziumbasierte Transistoren zu bevorzugen.

Ein mögliches Bauelement ist der sogenannte Tunnelfeldeffekttransistor. Dieser kann im Rahmen der bestehenden Prozesstechnologien hergestellt werden und es sind dafür auch keine in der Siliziumtechnologie neuen Materialien erforderlich. Mit dem gategesteuerten Band-zu-Band Tunneln wird einer der physikalischen Mechanismen die bei MOSFETs zu Kurzkanaleffekten führen (*gate-induced drain leakage*) als Funktionsprinzip genutzt. Mit Ausnahme des direkten Tunnelns durch das Gateoxid treten bei dem TFET keine weiteren Kurzkanaleffekte auf. Ein besonderer Vorteil liegt hierbei in den zu erwartenden geringen Sperrströmen des Tunneltransistors und der steileren Transferkennlinien. Auf diese Weise ist ein geringerer Gatespannungsbereich erforderlich, um zwischen Ein- und Aus-Zustand zu unterscheiden. Darüber hinaus findet das Band-zu-Band Tunneln nur am pn-Übergang statt und ist damit räumlich sehr stark begrenzt. Daher steht zu erwarten, dass mit dem TFET auch Bauelemente mit extrem kurzen Gatelängen realisiert werden können.

In dieser Arbeit wurde in einem ersten Schritt die Funktionsweise von dreidimensional strukturierten Nanodraht-Tunneltransistoren untersucht. Im Wesentlichen handelt es sich dabei um die Anwendung des Multigatekonzepts auf Tunnelfeldeffekttransistoren. Daran anschließend wurden unterschiedliche Optimierungsstrategien entwickelt und mit Hilfe dieser ein Ultrakurzkanal-TFET mit 5 nm Gatelänge entworfen. Der Schwerpunkt des letzten Teils lag auf der Verbesserung des eingesetzten Band-zu-Band Tunnelmodells und den verschiedenen Möglichkeiten vereinfachte Modellierungsmethoden zur Verringerung der Simulationszeiten anzuwenden.

Die im Rahmen der vorliegenden Arbeit erhaltenen Ergebnisse und die dazugehörigen Schlussfolgerungen können folgendermaßen zusammengefasst werden:

- Mit dem auf der Arbeit von Schenk [17] basierenden Modell zur Beschreibung von phononenunterstützten Band-zu-Band Tunnelprozessen durch eine Generationsrate ist es möglich, physikalisch sinnvolle Resultate bei der Simulation unterschiedlicher Tunnelbauelemente zu erzielen. Bei nahezu allen vorangehenden Analysen wurde zur Simulation von Band-zu-Band Tunnelbauelementen zumeist nur das Modell von Kane benutzt [70, 71]. Allerdings ist dabei zu bemerken, dass die Bezeichnung etwas irreführend ist, da es sich bei dem in Sdevice implementierten Modell um eine signifikant vereinfachte Variante des Modells nach Kane [64] handelt. Besonders problematisch ist bei diesen Generationsraten die vollständige Vernachlässigung der Ladungsträgerdichten in der Tunnelregion. Zudem wird auch die Abhängigkeit von der elektrischen Feldstärke nur unvollständig wiedergegeben. Unabhängig von diesen Beobachtungen ist das Modell nach Schenk die beste Wahl. Somit konnte erstmals im Rahmen einer semiklassischen Simulation die vollständige Kennlinie einer Tunneldiode simuliert werden. Hierbei war es zudem möglich, den durch das defektunterstützte Tunneln hervorgerufenen Stromanteil zu bestimmen. Auch die Kennlinien des Tunneltransistors wurden so mit den Band-zu-Band Tunnelmodellen nach Schenk am Rechner nachgebildet.
- Eine wichtige Rolle in dem Verhalten von nanostrukturierten MOS-Bauelementen spielen Quantisierungseffekte. Bei einer quantenmechanischen Herangehensweise folgt unter dem Gatestapel im Vergleich zu einer rein klassischen Simulation eine deutliche Abweichung der Ladungsverteilungen. Da es mit dem in dieser Arbeit eingesetzten Band-zu-Band Tunnelmodell möglich ist, auch Ladungsträgereffekte zu berücksichtigen, wird die gesamte Charakteristik von Tunneltransistoren von dieser veränderten Situation erheblich beeinflusst. Allerdings ist eine vollständige quantenmechanische Behandlung bei zwei- und erst recht bei dreidimensionalen Problemen nur mit einem extrem hohen rechentechnischen Aufwand möglich. Die Näherung über die Dichtegradientenmethode weist eine gute Übereinstimmung und eine hohe Praktikabilität auf. So müssen zwei zusätzliche Differentialgleichungen selbstkonsistent mit dem bestehenden System gelöst werden, wodurch letztlich die Simulationszeiten deutlich ansteigen. Dennoch folgt im Vergleich zur vollständigen quantenmechanischen Lösung eine erhebliche Beschleunigung. Mit der erstmaligen Einbindung von Quantisierungseffekten bei der Simulation von Tunneltransistoren war es möglich aufzuzeigen, dass ein solcher Mehraufwand für eine physikalisch richtige Beschreibung von Tunneltransistoren erforderlich ist.
- Um die Simulationszeit möglichst gering zu halten, wurden in einem ersten Schritt Nanodraht-Tunnelfeldeffekttransistoren mit einem zylindrischen Kanalquerschnitt betrachtet. Anhand dieser Grundgeometrie konnten die jeweiligen Auswirkungen einer Variation der Dotierprofile, der Bauelementabmessungen und des Materials untersucht werden. Bei diesen Simulationen wurden jeweils Quantisierungseffekte unter dem Gatestapel mittels der Dichtegradientenmethode berücksichtigt. Es hat

sich im Rahmen der Untersuchungen gezeigt, dass zum Erreichen eines möglichst hohen Stroms im Durchlassbereich besonders die Form des Sourcedotierprofils von Bedeutung ist. So sollte dieses möglichst hoch p-dotiert sein und gleichzeitig auch einen großen Gradienten aufweisen. Für geringe Ströme im Sperrbereich sollte die Draindotierung nicht zu hohe Werte erreichen und über die Kanaldotierung kann die Einsatzspannung eingestellt werden. Mittels eines Spaceroxids auf der Sourceseite kann die Position des Maximums der Sourcedotierung angepasst werden. Dieses sollte hinsichtlich geringer Gate-längen in Transportrichtung weder einen Überlapp noch einen Abstand zum Gatestapel haben. Wurde schließlich der Durchmesser des Querschnitts verringert, so konnte unter bestimmten Umständen eine Verbesserung der Bauelementeigenschaften beobachtet werden. Insgesamt konnten die sehr guten Skalierungseigenschaften des TFETs aufgezeigt werden. Die Einbindung von SiGe in der Source- beziehungsweise Drainregion des Tunneltransistors führt zu einer weiteren Verbesserung der Transfercharakteristiken des Bauelements. Durch den Einsatz von high-k Oxiden resultierten signifikant bessere Schalteigenschaften. Von der Einführung von Materialien mit noch höheren Dielektrizitätszahlen kann eine weitere Verbesserung erwartet werden. Hinsichtlich des Designs von TFETs mit sehr kurzen Gate-längen erwiesen sich auch Bauelemente, bei denen die Steuerung des Kanals über das Spaceroxid auf der Sourceseite des Gatestapels erfolgt, als sehr vielversprechend (siehe auch [72]). Beispielhaft wurden die erzielten Resultate zum Entwurf eines Tunneltransistors mit 5 nm Gate-länge herangezogen. Für dieses Bauelement ergaben sich dann sehr gute Schalteigenschaften und es konnte durch geringfügige Änderungen auch ein p-Typ Tunneltransistor erhalten werden. Weitere Ergebnisse der Untersuchungen zur Optimierung von zylindrischen TFETs sind in Abschnitt 4.5 zusammengefasst.

- Da der technologische Aufwand zur Fertigung von Nanodraht-Tunneltransistoren mit einem rechteckigen Kanalquerschnitt wesentlich geringer als bei Bauelementen mit zylindrischem Querschnitt ist, wurden auch ausgiebige Simulationsreihen von Quad- und Trigate-Tunneltransistoren durchgeführt. Gerade bei diesen TFETs ist die Berücksichtigung von Quantisierungseffekten unter dem Gate notwendig, da die Ladungsdichten in den Ecken des leitenden Kanals in der Realität deutlich von den mit einer rein semiklassischen Simulation erhaltenen Werten abweichen. Wegen der fertigungstechnisch begründeten dickeren Gateoxide auf der Oberseite des Kanals bei Trigate-TFETs, sollte der Kanal bei diesen höher als weit sein. Bei rechteckigen Bauelementen haben die Band-zu-Band Tunnelraten in den Ecken des Kanals ein Maximum. Allerdings sind bei gleichem Durchmesser des Kanals die Ströme nicht größer als bei zylindrischen Tunneltransistoren, da bei rechteckigen Bauelementen die Raten außerhalb der Ecken niedriger ausfallen. Im Fall hoher Dotierungen der Source-region folgt, wie bei zylindrischen Tunneltransistoren, ein Optimum für den Durchmesser des Kanals. Hierbei ist zu beachten, dass der Kanal deshalb nicht zu dünn werden darf, da ab einem gewissen Durchmesser das Band-zu-Band Tunneln im gesamten Querschnitt stattfindet und damit der Drainstrom bei einer Reduktion der Fläche abnimmt. Insgesamt stellen rechteckige Nanodraht-Tunneltransistoren bei kurzen Kanallängen eine vielversprechende Alternative zu MOSFETs dar.

- Bei der detaillierten Untersuchung des Band-zu-Band Tunnelmodells nach Schenk zeigte sich, dass dieses fehlerhaft ist. So wurde bei der Berechnung der Tunnelrate eine falsche Näherung des Airy-Integrals eingesetzt. In dieser Arbeit konnte eine korrigierte Näherung eingeführt und somit die korrekte Formel für die Tunnelrate abgeleitet werden. Der Test dieses verbesserten Modells erfolgte dann mit Hilfe des *physical model interface* von Sdevice. Die Auswirkungen auf das makroskopische Verhalten von Tunneltransistoren sind erheblich: Mit dem Originalmodell resultiert ein um mehr als eine Größenordnung abweichender Drainstrom. Das neue Modell weist diesen Fehler nicht mehr auf. Zudem wird hierdurch der rechentechnische Aufwand nicht signifikant erhöht. Eine Verringerung der Rechenzeit ist durch eine sogenannte *postprocessing*-Strategie möglich. Bei dieser werden die Band-zu-Band Tunnelströme in der Selbstkonsistenzschleife weggelassen und erst im Nachhinein aus den Ladungsdichten und der Feldstärke berechnet. Dieses Verfahren liefert sehr gute Ergebnisse, allerdings nimmt mit steigender Dopingkonzentration auch der Fehler zu. Abschließend wurde das einfache Kompaktmodell für Tunneltransistoren nach Vandenberghe [63] für die in dieser Arbeit verwendeten Tunnelraten angepasst und analysiert. Wegen der Nichtberücksichtigung der Ladungsverteilungen, kann dieses nur als Baustein eines geeigneten Kompaktmodells interpretiert werden.

Mit Hilfe der durchgeführten Simulationen von Nanodraht-Tunnelfeldeffekttransistoren mit einer rechteckigen Querschnittsfläche konnten die vielversprechenden Eigenschaften dieser Bauelemente festgestellt werden. So ist es, mit den in dieser Arbeit beschriebenen Optimierungsstrategien, möglich, Tunneltransistoren zu entwickeln, die sowohl ein gutes Sperrverhalten als auch hohe Ströme im Durchlassbereich aufweisen. In einem nächsten Schritt ist es gerade im Fall rechteckiger Nanodraht-Tunneltransistoren erforderlich auch komplexere, und damit stärker an der Herstellung orientierte, Dotierstoffverteilungen zu untersuchen. Da aus fertigungstechnischen Gründen die Dotierkonzentrationen im Kanal mit zunehmender Tiefe abnehmen, wird die Kanalhöhe von Nanodraht-Tunneltransistoren vermutlich einen optimalen Wert haben. Des Weiteren ist bei diesen Bauelementen der Einfluss der Position der Source- und Drainkontakte zu berücksichtigen. Im Hinblick auf eine fortgesetzte Untersuchung von Tunnelbauelementen sind einige Anforderungen an die hierfür eingesetzten Bauelementensimulatoren zu stellen: Das im Rahmen dieser Arbeit entwickelte, korrigierte Tunnelmodell sollte in jedem Fall in die Simulatoren eingebaut werden. Darüberhinaus ist eine Analyse des extrem schlechten Konvergenzverhaltens bei Simulationen von Tunnelbauelementen mit Dotierstoffkonzentrationen über 10^{20} cm^{-3} unbedingt erforderlich. Letztlich ist voraussichtlich eine besser an die Problemstellung angepasste numerische Behandlung der Band-zu-Band Tunnelrate notwendig. Außerdem sollte auf Basis von physikalisch fundierten theoretischen Modellen die Nichtlokalität der Stoßionisation und des Band-zu-Band Tunnelns berücksichtigt werden.

Symbolverzeichnis

Da sich nicht in allen Fällen Doppelbelegungen von Symbolzeichen vermeiden lassen, sei darauf hingewiesen, dass die Bedeutung der einzelnen Symbole immer aus dem Kontext eindeutig abzulesen ist. Zur besseren Übersicht sind die jeweiligen Symbolbezeichnungen kapitelweise aufgelistet.

Kapitel 2

Symbol	Bedeutung	Einheit
A	Vorfaktor der richtungsabhängigen Tunnelrate	$(\text{mV}^2\text{s})^{-1}$
A^*	Vorfaktor der Band-zu-Band Tunnelstromdichte	$(\text{mV}^2\text{s})^{-1}$
$\text{Ai}(x)$	Airy-Funktion	1
$\text{Ai}'(x)$	Ableitung der Airy-Funktion	1
$\text{Ai}_1(x)$	Airy-Integral	1
B	Vorfaktor der richtungsunabhängigen Tunnelrate	$(\text{mV}^2\text{s})^{-1}$
c_s	Schallgeschwindigkeit	m/s
d_{ij}	Kantenlänge der Zelle zwischen den Gitterpunkten i und j	m
$D_{ac,n}$	Deformationspotential der akustischen Elektronenstreuung	eV
D_n	Diffusionskoeffizient der Elektronen	m^2s^{-1}
D_p	Diffusionskoeffizient der Löcher	m^2s^{-1}
D^*	Differenz der Fermi-Verteilungsfunktionen	1
E_0	Energieniveau des optimalen horizontalen Übergangspfades	eV
E_c	Leitungsbandenergie	eV
E_g	Energiebandlücke	eV
E_g^d	direkte Energiebandlücke	eV
E_g^i	minimale indirekte Bandlücke	eV
E_{tn}	modifizierte Störstellenenergie der Elektronen	eV

Symbol	Bedeutung	Einheit
E_{tp}	modifizierte Störstellenenergie der Löcher	eV
E_v	Valenzbandenergie	eV
f	Verteilungsfunktion der Boltzmann-Gleichung	1
f_B	Phononenbesetzungszahl (Bose-Einstein-Verteilung)	1
f_c	Fermi-Verteilungsfunktion im Leitungsband	1
f_v	Fermi-Verteilungsfunktion im Valenzband	1
F	Betrag der elektrischen Feldstärke	V/m
\mathbf{F}	elektrisches Feld	V/m
F_c^\pm	richtungsunabhängige kritische elektrische Feldstärke	V/m
$F_c^{\alpha\pm}$	richtungsabhängige kritische elektrische Feldstärke	V/m
F^{ii}	treibende Kraft der Stoßionisation	V/m
F_n	Quasi-Fermi-Niveau der Elektronen	eV
F_p	Quasi-Fermi-Niveau der Löcher	eV
g_n	Feldabhängigkeit der Elektronenlebensdauer	1
g_p	Feldabhängigkeit der Löcherlebensdauer	1
G	Generationsrate	$\text{m}^{-3}\text{s}^{-1}$
G^{ii}	Generationsrate der Stoßionisation	$\text{m}^{-3}\text{s}^{-1}$
h	Plancksches Wirkungsquantum	Js
\hbar	reduziertes Plancksches Wirkungsquantum	Js
$\hbar\omega_0$	effektive Phononenenergie	eV
$\hbar\omega_{TA}$	Energie des akustischen Phonons	eV
\mathbf{j}	Stromdichtetensor	Am^{-2}
\mathbf{j}_n	Elektronenstromdichte	Am^{-2}
\mathbf{j}_p	Löcherstromdichte	Am^{-2}
j_t	Stromdichte des phononenuntersützten Tunnelns	Am^{-2}
j_t^{ind}	Stromdichte für indirekte Band-zu-Band Übergänge	Am^{-2}
j_t^{dir}	Stromdichte für direkte Band-zu-Band Übergänge	Am^{-2}
k	Boltzmann-Konstante	J/K
\mathbf{k}	Wellenvektor	m^{-1}
\mathbf{k}_e	Extremum des Wellenvektors	m^{-1}
l_{ij}	Abstand der Gitterpunkte i und j	m

Symbol	Bedeutung	Einheit
l_t	Tunnelweglänge	m
m_e	effektive Masse der Elektronen	kg
m_h	effektive Masse der Löcher	kg
m_l	effektive longitudinale Elektronenmasse	kg
m_t	effektive transversale Elektronenmasse	kg
m_{ν, \mathbf{k}_e}^*	effektiver Massentensor	kg
m_{\perp}^{α}	effektive Elektronenmasse orthogonal zur Feldrichtung	kg
m_{\parallel}^{α}	effektive Elektronenmasse in Feldrichtung	kg
$M_{\mathbf{k}'\mathbf{k}}$	Matrizelement des Streupotentials	$\text{m}^{-3/2}$
n	Elektronendichte	m^{-3}
n_i	intrinsische Ladungsträgerdichte	m^{-3}
N	Zahl der Elektronen	1
N_0	Referenzkonzentration der Scharfetter-Relation	m^{-3}
N_A	Akzeptorenkonzentration	m^{-3}
N_A^-	aktivierte Akzeptorenkonzentration	m^{-3}
N_c	Zustandsdichte des Leitungsbandes	m^{-3}
N_D	Donatorenkonzentration	m^{-3}
N_D^+	aktivierte Donatorenkonzentration	m^{-3}
N_v	Zustandsdichte des Valenzbandes	m^{-3}
p	Löcherdichte	m^{-3}
p^{cv}	Impulsmatrizelement des Band-zu-Band Tunnelübergangs	kgms^{-1}
q	Elementarladung	C
R	Rekombinationsrate	$\text{m}^{-3}\text{s}^{-1}$
R^{SRH}	Shockley-Read-Hall-Rekombinationsrate	$\text{m}^{-3}\text{s}^{-1}$
R_t^{ind}	Rekombinationsrate für indirektes Band-zu-Band Tunneln	$\text{m}^{-3}\text{s}^{-1}$
S	Huang-Rhys-Faktor der Elektron-Phonon-Kopplung	1
$S(f)$	Stoßintegral für klassische Teilchen	s^{-1}
t	Zeit	s
T	Temperatur	K
T_0	Raumtemperatur (300 K)	K
v	Gruppengeschwindigkeit	m/s

Symbol	Bedeutung	Einheit
v_n	Driftgeschwindigkeit der Elektronen	m/s
v_p	Driftgeschwindigkeit der Löcher	m/s
W	Streurrate	$\text{m}^{-3}\text{s}^{-1}$
x_c	Umkehrpunkt des Tunnelweges im Leitungsband	m
x_v	Umkehrpunkt des Tunnelweges im Valenzband	m
α_n	Ionisationskoeffizient der Elektronen	m^{-1}
α_p	Ionisationskoeffizient der Löcher	m^{-1}
γ	relativer Subbandbesetzungsfaktor der Quantenkorrektur	1
γ_{cv}	Oszillatorstärke	1
γ^d	Elektronenruhemasse	kg
γ^{ii}	Temperaturabhängigkeit der Ionisationskoeffizienten	1
$\Delta\Phi_{\text{WF}}$	Differenz der Arbeitsfunktion zwischen Metall und Referenz	V
ε	Energie	eV
ε_R	Gitterrelaxationsenergie	eV
$\varepsilon^{\text{trans}}$	Transferenergie der Streuung	eV
ε_ν	ν -tes Energieband	eV
$\tilde{\varepsilon}_\nu$	parabolische Näherung des ν -ten Energiebandes	eV
Θ_n	elektro-optische Frequenz der Elektronen	s^{-1}
Θ_p	elektro-optische Frequenz der Löcher	s^{-1}
$\Theta_{r\parallel}^\alpha$	elektro-optische Frequenz des indirekten Tunnelübergangs	s^{-1}
Θ_Γ	elektro-optischen Frequenz des Tunnelübergangs	s^{-1}
Λ_n	Quantenkorrekturpotential der Elektronen	V
μ_n	Elektronenbeweglichkeit	$\text{As}^2\text{kg}^{-1}$
μ_p	Löcherbeweglichkeit	$\text{As}^2\text{kg}^{-1}$
μ_Γ	reduzierte effektive Masse im Gammapunkt	kg
ϱ	Ladungsträgerdichte	m^{-3}
ρ	Massendichte	kgm^{-3}
$\rho(t)$	Dichtetensor des großkanonischen Ensembles	m^{-3}
σ	Leitfähigkeitstensor des phononenunterstützten Tunnelns	$\text{AV}^{-1}\text{m}^{-1}$
τ	mikroskopische Relaxationszeit	s
τ_d	Dotierstoffabhängigkeit der Lebensdauer	s

Symbol	Bedeutung	Einheit
τ_n	Elektronenlebensdauer	s
$\tau_{\max,n}$	maximale Elektronenlebensdauer der Scharfetter-Relation	s
$\tau_{\max,p}$	maximale Löcherlebensdauer der Scharfetter-Relation	s
$\tau_{\min,n}$	minimale Elektronenlebensdauer der Scharfetter-Relation	s
$\tau_{\min,p}$	minimale Löcherlebensdauer der Scharfetter-Relation	s
τ_p	Löcherlebensdauer	s
Φ	elektrostatisches Potential	V
Φ_F	Fermi-Potential	V
Φ_g	elektrostatisches Potential an einem Gatekontakt	V
Φ_{ref}	Referenzpotential	V
χ	Elektronenaffinität	eV
ω	Frequenz	s ⁻¹
Ω_i	Volumen der i -ten Zelle bei Boxdiskretisierung	m ²

Kapitel 3

Symbol	Bedeutung	Einheit
$E_{\text{ion},c}$	Ionisationsenergie der Elektronen bzw. Löcher	eV
F_c	richtungsunabhängige kritische elektrische Feldstärke	V/m
I	Strom	A
V	Spannung	V
ΔPot	Potentialdifferenz in Richtung der elektrischen Feldstärke (ist im Abstand $\Delta\text{Pot}/F$ die Potentialdifferenz kleiner als $\Delta\text{Pot}/2$ wird die Tunnelrate auf Null gesetzt)	V

Kapitel 4

Symbol	Bedeutung	Einheit
D	Durchmesser	nm

Symbol	Bedeutung	Einheit
L_{gate}	Gatelänge	nm
N_C	Kanaldotierstoffkonzentration	cm^{-3}
N_D	Dotierstoffkonzentration der Drainregion	cm^{-3}
N_S	Dotierstoffkonzentration der Sourcereion	cm^{-3}
S	Abstand zwischen Maximum der Dotierung und Gatestapel	nm
t_{ox}	Oxididicke	nm
W	Gateweite	nm
ε	relative Dielektrizitätskonstante	1
δ	Gradient der Dotierprofile	nm/dec

Kapitel 5

Symbol	Bedeutung	Einheit
C	Vorfaktor der richtungsunabhängigen Tunnelrate	$(\text{mV}^2\text{s})^{-1}$
R_t^{Kane}	Rekombinationsrate für Band-zu-Band Tunneln nach Kane	$\text{m}^{-3}\text{s}^{-1}$
t'_{ox}	effektive Oxididicke	nm

Naturkonstanten und Modellparameter

In dieser Aufstellung werden die für die vorliegende Arbeit relevanten Naturkonstanten bereitgestellt. Zudem sind die Parameterwerte der Generations-/Rekombinationsmodelle und der Dichtegradientenmethode angegeben. Alle nicht hier aufgelisteten Parameter der Bauelementesimulation aus den Kapiteln zwei bis fünf haben die in Sentaurus Device voreingestellten Standardwerte.

Naturkonstanten

Symbol	Bedeutung	Größe	Einheit
ε_0	Dielektrizitätskonstante des Vakuums	$8.854187817 \cdot 10^{-12}$	C ² /(Jm)
k	Boltzmann-Konstante	$1.380658 \cdot 10^{-23}$	J/K
\hbar	Plancksches Wirkungsquantum	$1.05457266 \cdot 10^{-34}$	Js
q	Elementarladung	$1.60217733 \cdot 10^{-19}$	C
m_0	Elektronenruhemasse	$9.1093897 \cdot 10^{-31}$	kg

Material und Bandstruktur in Silizium

Symbol	Bedeutung	Größe	Einheit
ϵ_{Si}	relative Dielektrizitätskonstante in Silizium	11.7	1
ϵ_{SiO_2}	relative Dielektrizitätskonstante in Siliziumdioxid	3.9	1
m_l/m_0	relative laterale Elektronenmasse	0.9163	1
m_t/m_0	relative transversale Elektronenmasse	0.1905	1
m_h/m_0	relative effektive Löchermasse	0.16	1
E_g	minimale Bandlücke bei 300 K	1.12	eV
χ_0	Elektronenaffinität bei 0 K	4.05	eV
α	erster Temperaturkoeffizient der Bandstruktur	4.73e-4	eV/K
β	zweiter Temperaturkoeffizient der Bandstruktur	636	K
$E_g(0)$	Bandlücke bei 0 K	1.15365	eV
E_{bgn}	Koeffizient der Bandlückenvariation	9.0e-3	eV
N_{ref}	Referenzdotierung	1.0e17	cm ⁻³

Shockley-Read-Hall-Rekombination

Symbol	Bedeutung	Größe	Einheit
$\tau_{\text{min},n}$	minimale Elektronenlebensdauer	0.0	s
$\tau_{\text{max},n}$	maximale Elektronenlebensdauer	$1 \cdot 10^{-5}$	s
$\tau_{\text{min},p}$	minimale Löcherlebensdauer	0.0	s
$\tau_{\text{max},p}$	maximale Löcherlebensdauer	$3 \cdot 10^{-6}$	s
N_0	Referenzdichte bei störstellenabh. Lebensdauer	$1 \cdot 10^{16}$	cm ⁻³
γ	Exponent bei störstellenabhängiger Lebensdauer	1.0	1
S	Huang-Rhys-Faktor	3.5	1
$\hbar\omega_0$	effektive Phononenenergie	0.068	eV
m_n/m_0	effektive Elektronentunnelmasse in Feldrichtung	0.258	1
m_p/m_0	effektive Löchertunnelmasse in Feldrichtung	0.24	1

Interbandtunnelmodelle

Symbol	Bedeutung	Größe	Einheit
E_g^i	minimale Bandlücke in Silizium	1.12	eV
$E_g^d(\Gamma)$	direkte Bandlücke im Gammapunkt	3.42	eV
c_s	Schallgeschwindigkeit in Silizium	9200	ms^{-1}
ρ	Massendichte	2.32831	gcm^{-3}
$\hbar\omega_{TA}$	Energie des akustischen Phonons	18.6	meV
$D_{ac,n}$	Konstante des Deformationspotentials	9	eV
A_{Schenk}	Vorfaktor der Band-zu-Band Tunnelrate	$8.977 \cdot 10^{20}$	$(\text{cmV}^2\text{s})^{-1}$
B_{Schenk}	Vorfaktor der kritischen Feldstärke	$2.147 \cdot 10^7$	V/cm
A_{Kane}	Vorfaktor der Band-zu-Band Tunnelrate	$3.5 \cdot 10^{21}$	$(\text{cmV}^2\text{s})^{-1}$
B_{Kane}	Vorfaktor der kritischen Feldstärke	$2.25 \cdot 10^7$	V/cm

Anmerkung:

Der Koeffizient A_{Schenk} entspricht dem Faktor B aus Gleichung (2.68) und B_{Schenk} ist der bandlückenunabhängige Vorfaktor aus Gleichung (2.63). Die hierbei abweichenden Bezeichnungen wurden wegen der besseren Vergleichbarkeit zu der Namensgebung in Sentaurus Device gewählt.

Dichtegradientenmethode

Symbol	Bedeutung	Größe	Einheit
γ_n	Besetzung der Subbänder der Elektronen	3.6	1
γ_p	Besetzung der Subbänder der Löcher	5.6	1

Literaturverzeichnis

- [1] M. Fulde, A. Heigl, G. Wachutka, G. Knoblinger, and D. Schmitt-Landsiedel. Complementary multi-gate tunneling FETs: fabrication, optimisation and application aspects. *International Journal on Nanotechnology*, 6(7/8):628–639, 2009.
- [2] Ch.R. Müller, L. Worschech, S. Höfling, and A. Forchel. Characterization of three-terminal junctions operated as in-plane gated field-effect transistors. *IEEE Transactions on Electron Devices*, 56(2):306–311, 2009.
- [3] C. Dekker. Carbon nanotubes as molecular quantum wires. *Physics Today*, 52(5):22–28, 1999.
- [4] L. Esaki. Long journey into tunneling. *Review on Modern Physics*, 46(2):237–244, 1974.
- [5] W. Hansch, C. Fink, J. Schulze, and I. Eisele. A vertical MOS-gated Esaki tunneling transistor in silicon. *Thin Solid Films*, 369:387–389, 2000.
- [6] A. Schenk. Rigorous theory and simplified model of the band-to-band tunneling in silicon. *Solid State Electronics*, 36(1):19–34, 1993.
- [7] W. Hänsch. *The drift diffusion equation and its application in MOSFET modeling*. Springer-Verlag, Wien, New York, 1991.
- [8] Ch. Jungemann and B. Meinerzhagen. *Hierarchical device simulation - The Monte-Carlo perspective*. Springer-Verlag, Wien, New York, 2003.
- [9] E. Wigner. On the quantum correction for thermodynamic equilibrium. *Physical Review*, 40:749–759, 1932.
- [10] S. Datta. *Electronic Transport in Mesoscopic Systems*. Cambridge University Press, Cambridge, 1995.
- [11] M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas. Generalized many-channel conductance formula with applications to small rings. *Physical Review B*, 31(10):6207–6215, 1986.
- [12] M. Büttiker. Symmetry of electrical conduction. *IBM Journal of Research and Development*, 32(3):317–334, 1988.

- [13] S.E. Laux, A. Kumar, and M.V. Fischetti. Analysis of quantum ballistic electron transport in ultrasmall silicon devices including space-charge and geometric effects. *Journal of Applied Physics*, 95(10):5545–5582, 2004.
- [14] G. Klimeck, R. Lake, R.C. Bowen, W.R. Frensley, and T.S. Moise. Quantum device simulation with a generalized tunneling formula. *Applied Physics Letters*, 67(12):2539–2541, 1995.
- [15] S. Datta. Nanoscale device modeling: The Greens function method. *Superlattices and Microstructure*, 28(4):253–278, 2000.
- [16] Ch. Kittel. *Einführung in die Festkörperphysik*. Oldenbourg Verlag, München, 2006.
- [17] A. Schenk. *Advanced physical models for silicon device simulation*. Springer-Verlag, Wien, New York, 1998.
- [18] E. Forsberg and J.O.J. Weström. Self-consistent simulations of mesoscopic devices operating under a finite bias. *Solid State Electronics*, 48:1147–1154, 2004.
- [19] E. Polizzi, N. Ben Abdallah, O. Vanbesien, and D. Lippens. Space lateral transfer and negative differential conductance regimes in quantum waveguide junctions. *Journal of Applied Physics*, 87(12):8700–8706, 2000.
- [20] R.C. Bowen, G. Klimeck, and R.K. Lake. Quantitative simulation of a resonant tunneling diode. *Journal of Applied Physics*, 81(7):3207–3213, 1997.
- [21] R. Lake, G. Klimeck, R.C. Bowen, and D. Jovanovic. Single and multiband modeling of quantum electron transport through layered semiconductor devices. *Journal of Applied Physics*, 81(12):7845–7869, 1997.
- [22] G. Klimeck. Quantum and Semi-Classical Transport in NEMO 1D. *Journal of Computational Electronics*, 2:177–182, 2003.
- [23] D. Mamaluy, D. Vasileska, M. Sabathil, T. Zibold, and P. Vogl. Contact block reduction method for ballistic transport and carrier densities of open nanostructures. *Physical Review B*, 71(245321), 2005.
- [24] M. Nedjalkov, H. Kosina, S. Selberherr, C. Ringhofer, and D.K. Ferry. Unified particle approach to Wigner-Boltzmann transport in small semiconductor devices. *Physical Review B*, 70(115319), 2004.
- [25] G. Masetti, M. Severi, and S. Solmi. Modeling of carrier mobility against carrier concentration in arsenic-, phosphorous- and boron-doped silicon. *IEEE Transactions on Electron Devices*, 30(7):764–769, 1983.
- [26] D.M. Caughey and R.E. Thomas. Carrier Mobilities in silicon empirically related to doping and field. *Proceedings of the IEEE*, 55(12):2192–2193, 1967.

- [27] S. Reggiani, M. Valdinoci, L. Colalongo, M. Rudan, G. Baccarani, A.D. Stricker, F. Illien, W. Fichtner, and L. Zullino. Electron and hole mobility in silicon at large operating temperatures - Part I: Bulk mobility. *IEEE Transactions on Electron Devices*, 49(3):490–499, 2002.
- [28] D.B.M. Klaassen. A unified mobility model for device simulation - I. Model equations and concentration dependence. *Solid State Electronics*, 35(7):953–959, 1992.
- [29] D.L. Scharfetter and H.K. Gummel. Large-signal analysis of a silicon read diode oscillator. *IEEE Transactions on Electron Devices*, 16:64–77, 1969.
- [30] W. Shockley and W.T. Read. Statistics of the recombinations of holes and electrons. *Physical Review*, 87(5):835–842, 1952.
- [31] D.J. Roulston, N.D. Arora, and S.G. Chamberlain. Modeling and measurement of minority-carrier lifetime versus doping in diffused layers of n⁺p silicon diodes. *IEEE Transactions on Electron Devices*, 29(2):284–291, 1982.
- [32] M.S. Tyagi and R. van Overstraeten. Minority carrier recombination in heavily-doped silicon. *Solid-State Electronics*, 26(6):577–597, 1983.
- [33] A. Schenk. A model for the field and temperature dependence of Shockley-Read-Hall lifetimes in silicon. *Solid State Electronics*, 35(11):1585–1596, 1992.
- [34] G.A.M. Hurkx, D.B.M. Klaassen, and M.P.G. Knuvers. A new recombination model for device simulation including tunneling. *IEEE Transactions on Electron Devices*, 39(2):331–338, 1992.
- [35] A.G. Chynoweth. Ionization rates for electrons and holes in silicon. *Physical Review*, 109(5):1537–1540, 1958.
- [36] R. van Overstraeten and H. de Man. Measurement of the ionization rates in diffused silicon p-n junctions. *Solid-State Electronics*, 13(1):583–608, 1970.
- [37] R. Kubo. Statistical-mechanical theory of irreversible processes. I. General theory and simple applications of magnetic and conduction problems. *Journal of the Physical Society of Japan*, 12(6):3917–34, 1980.
- [38] *Sentaurus device user guide*. Synopsys Inc., Mountain View, CA, 2010.
- [39] D. Bohm. A suggested interpretation of the quantum theory in terms of hidden variables I. *Physical Review*, 85(2):166–179, 1952.
- [40] D. Bohm. A suggested interpretation of the quantum theory in terms of hidden variables II. *Physical Review*, 85(2):180–193, 1952.
- [41] M.G. Ancona and H.F. Tiersten. Macroscopic physics of silicon inversion layer. *Physical Review B*, 35(15):7959–7965, 1987.

- [42] J.J. Quinn, G. Kawamoto, and B.D. McCombe. Subband spectroscopy by surface channel tunneling. *Surface Science*, 73:190–196, 1978.
- [43] S. Sedlmaier, J. Schulze, T. Sulima, C. Fink, C. Tolksdorf, A. Bayerstadler, I. Eisele, P.F. Wang, K. Hilsenbeck, and W. Hansch. Phonon assisted tunneling in gated pin diodes. *Materials Science and Engineering*, B89:116–119, 2002.
- [44] J. Koga and A. Toriumi. Negative differential conductance in three-terminal silicon tunneling device. *Applied Physics Letters*, 69(10):1435–1437, 1996.
- [45] A. Heigl and G. Wachutka. Simulation of Advanced Tunneling Devices. In *Proceedings of the International Conference on Advanced Semiconductor Devices and Microsystems (ASDAM)*, pages 121–124, Smolenice, Slovakia, 26.–29. Oktober 2006.
- [46] A. Heigl and G. Wachutka. Simulation of Advanced Tunneling Devices. In *Technical Digest of the IEEE EDS Workshop on Advanced Electron Devices (AED)*, pages 1–2, Duisburg, Germany, 26.–29. Mai 2006.
- [47] J.L. Moll and R. van Overstraeten. Charge multiplication in silicon p-n junctions. *Solid-State Electronics*, 6(2):147–157, 1963.
- [48] C.L. Anderson and C.R. Crowell. Threshold energies for electron-hole pair production by impact ionization in semiconductors. *Physical Review B*, 5(6):2267–2272, 1972.
- [49] A. Heigl and G. Wachutka. Simulation of Silicon Nanowire Tunneling Field-Effect Transistors Including Quantum Effects. In *Proceedings of the International Semiconductor Device Research Symposium (ISDRS)*, pages 1–2, College Park, MD, USA, 26.–29. Dezember 2007.
- [50] A. Heigl and G. Wachutka. Optimization of Vertical Tunneling Field-Effect Transistors. In *Proceedings of the 8th International Conference on the Ultimate Integration on Silicon (ULIS)*, pages 133–136, Leuven, Belgium, 26.–29. Mai 2007.
- [51] A. Heigl and G. Wachutka. Study on the Optimized Design of Nanowire Tunneling Transistors Including Quantum Effects. In *Proceedings of the International Conference on the Simulation of Semiconductor Processes and Devices (SISPAD)*, pages 225–228, Hakone, Japan, 26.–29. September 2008.
- [52] G.D. Wilk, R.M. Wallace, and J.M. Anthony. High- κ gate dielectrics: Current status and materials properties considerations. *Journal of Applied Physics*, 89(10):5243–5275, 2001.
- [53] R. Chau, S. Datta, M. Doczy, B. Doyle, J. Kavalieros, and M. Metz. High- κ /metal-gate stack and its MOSFET characteristics. *IEEE Electron Device Letters*, 25(6):408–410, 2004.

- [54] L.R.C. Fonseca and A.A. Knizhnik. First-principles calculation of the TiN effective work function on SiO₂ and on HfO₂. *Physical Review B*, 74:195304, 2006.
- [55] J. Westlinder, G. Sjöblom, and J. Olsson. Variable work function in MOS capacitors utilizing nitrogen-controlled TiNx gate electrodes. *Microelectronic Engineering*, 75(4):389–396, 2004.
- [56] D.R. Linde. *CRC Handbook of chemistry and physics*. Taylor and Francis, Boca Raton, FL, 2007.
- [57] A. Heigl and G. Wachutka. Quantum-Corrected Simulation of Complementary Nanowire Tunneling Transistors of 5 nm Gate-Length. In *Proceedings of the International Conference on Advanced Semiconductor Devices and Microsystems (AS-DAM)*, pages 115–118, Smolenice, Slovakia, 16.–18. Oktober 2008.
- [58] P.M. Solomon, K.W. Guarini, Y. Zhang, K. Chan, E.C. Jones, G.M. Cohen, A. Krasnoperova, M. Ronay, O. Dokumaci, H.J. Hovel, J.J. Bucchignano, C. Cabral Jr., C. Lavoie, V. Ku, D.C. Boyd, K. Petrarca, J.H. Yoon, I.V. Babich, J. Treichler, P.M. Kozlowski, J.S. Newbury, C.P. D’Emic, R.M. Sicina, J. Benedict, and H.-S.P. Wong. Two gates are better than one. *IEEE Circuits and Devices*, 19:48–62, 2003.
- [59] A. Heigl and G. Wachutka. Detailed Analysis of Quantum-Effects in Nanowire Tunneling Transistors with Different Channel-Profiles. In *Proceedings of the NSTI Nanotechnology Conference and Trade Show (Nanotech)*, volume 1, pages 590–593, Houston, TX, USA, 3.–7. Mai 2009.
- [60] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York, 1964.
- [61] A.I. Nikishov and V.I. Ritus. Asymptotic representations for some functions and integrals connected with the Airy function. arXiv:math-ph/0501062v1, 2005.
- [62] A. Heigl, A. Schenk, and G. Wachutka. Correction to the Schenk Model of Band-to-Band Tunneling in Silicon Applied to the Simulation of Nanowire Tunneling Transistors. In *Proceedings of 13th International Workshop on Computational Electronics (IWCE)*, pages 96–99, Beijing, China, 27.–29. Mai 2009.
- [63] W.G. Vandenberghe, A.S. Verhulst, G. Groeseneken, B. Sorée, and W. Magnus. Analytical model for point and line tunneling in a tunnel field-effect transistor. In *Proceedings of the International Conference on the Simulation of Semiconductor Processes and Devices (SISPAD)*, pages 137–140, Hakone, Japan, 26.–29. September 2008.
- [64] E.O. Kane. The theory of tunneling. *Journal of Applied Physics*, 32(1):83–91, 1961.
- [65] A.S. Verhulst, B. Sorée, D. Leonelli, W.G. Vandenberghe, and G. Groeseneken. Modeling the single-gate, double-gate and gate-all-around tunnel field-effect transistor. *Journal of Applied Physics*, 107(2):024518, 2010.

- [66] B. Hoeneisen and C.A. Mead. Fundamental limitations in microelectronics–I. MOS technology. *Solid-State Electronics*, 15(7):819–829, 1972.
- [67] C. Hu. Gate oxide scaling limits and projection. In *Proceedings of the International Electron Devices Meeting (IEDM)*, pages 319–322, 1996.
- [68] A.I. Kingon, J.-P. Maria, and S.K. Streiffer. Alternative dielectrics to silicon dioxide for memory and logic devices. *Nature*, 406:1032–1038, 2000.
- [69] J. Welser, J.L. Hoyt, and J.F. Gibbons. Electron mobility enhancement in strained-Si n-type metal-oxide-semiconductor field-effect transistors. *IEEE Electron Device Letters*, 15(3):100–102, 1994.
- [70] P.-F. Wang, Th. Nirschl, D. Schmitt-Landsiedel, and W. Hansch. Simulation of the Esaki-tunneling FET. *Solid-State Electronics*, 47(7):1187–1192, 2003.
- [71] K.K. Bhuwarka, J. Schulze, and I. Eisele. A simulation approach to optimize the electrical parameters of a vertical tunnel FET. *IEEE Transactions on Electron Devices*, 52(7):1541–1547, 2005.
- [72] M. Sterkel. *Ultra-Kurzkanal Tunnel-Feldeffekt-Transistoren auf Silizium- und SOI-Substraten*. Dissertation, TU München, 2008.
- [73] A. Heigl and G. Wachutka. Simulation of Quantum-Ballistic Nanoswitches. In *Proceedings of the 11th International Workshop on Computational Electronics (IWCE)*, pages 191–192, Wien, Austria, 26.–29. Mai 2006.
- [74] A. Heigl and G. Wachutka. Simulation of quantum-ballistic nanoswitches. *Journal of Computational Electronics*, 6(1):97–100, 2007.
- [75] M. Fulde, A. Heigl, M. Wirnshofer, K.v. Arnim, T. Nirschl, M. Sterkel, G. Knoblinger, W. Hansch, G. Wachutka, and D. Schmitt-Landsiedel. Fabrication, Optimization and Application of Complementary Multiple-Gate Tunnelings FETs. In *Proceedings of the 2nd IEEE International Nanoelectronics Conference (INEC)*, pages 579–584, Shanghai, China, 26.–29. März 2008.
- [76] F. Schwabl. *Quantenmechanik*. Springer-Verlag, Berlin, Heidelberg, New York, 1998.
- [77] F. Schwabl. *Quantenmechanik für Fortgeschrittene*. Springer-Verlag, Berlin, Heidelberg, New York, 1997.
- [78] F. Schwabl. *Statistische Mechanik*. Springer-Verlag, Berlin, Heidelberg, New York, 2000.
- [79] H.R. Schwarz. *Numerische Mathematik*. B.G. Teubner Verlag, Stuttgart, 1997.
- [80] Ch. Kanzow. *Numerik linearer Gleichungssysteme*. Springer-Verlag, Berlin, Heidelberg, New York, 2005.

- [81] A. Meister. *Numerik linearer Gleichungssysteme*. Vieweg-Verlag, Wiesbaden, 2005.
- [82] S. Datta. *Quantum Transport - Atom to Transistor*. Cambridge University Press, Cambridge, 2005.
- [83] J.D. Jackson. *Klassische Elektrodynamik*. Walter de Gruyter Verlag, Berlin, New York, 2006.
- [84] C. Cohen-Tannoudji, B. Diu, and F. Laloe. *Quantenmechanik - Band 1*. Walter de Gruyter Verlag, Berlin, New York, 1999.
- [85] C. Cohen-Tannoudji, B. Diu, and F. Laloe. *Quantenmechanik - Band 2*. Walter de Gruyter Verlag, Berlin, New York, 2008.
- [86] J.W. Demmel, S.C. Eisenstat, J.R. Gilbert, X.S. Li, and J.W.H. Liu. A supernodal approach to sparse partial pivoting. *SIAM J. Matrix Analysis and Applications*, 20(3):720–755, 1999.
- [87] T.A. Davis and I.S. Duff. An Unsymmetric-Pattern Multifrontal Method for Sparse LU Factorization. *SIAM Journal on Matrix Analysis and Applications*, 18(1):140–158, 1997.
- [88] T.A. Davis and I.S. Duff. A combined unifrontal/multifrontal method for unsymmetric sparse matrices. *ACM Trans. Math. Softw.*, 25(1):1–20, 1999.
- [89] T.A. Davis. Algorithm 832: UMFPACK V4.3—an unsymmetric-pattern multifrontal method. *ACM Trans. Math. Softw.*, 30(2):196–199, 2004.
- [90] T.A. Davis. A column pre-ordering strategy for the unsymmetric-pattern multifrontal method. *ACM Trans. Math. Softw.*, 30(2):165–195, 2004.
- [91] P.R. Amestoy, T.A. Davis, and I.S. Duff. An Approximate Minimum Degree Ordering Algorithm. *SIAM Journal on Matrix Analysis and Applications*, 17(4):886–905, 1996.
- [92] G. Karypis and V. Kumar. A Fast and High Quality Multilevel Scheme for Partitioning Irregular Graphs. *SIAM Journal on Scientific Computing*, 20(1):359–392, 1998.
- [93] O. Schenk and K. Gärtner. Solving unsymmetric sparse systems of linear equations with PARDISO. *Future Generation Computer Systems*, 20(3):475–487, 2004.
- [94] Y. Saad. *Iterative methods for sparse linear systems*. Society for Industrial and Applied Mathematics, Philadelphia, 2003.
- [95] B.M. Irons and R.C. Tuck. A version of the Aitken accelerator for computer iteration. *International Journal of Numerical Methods in Engineering*, 1:275–277, 1969.

- [96] A. Frommer. BiCGStab(l) for families of shifted linear systems. *Computing*, 70:87–109, 2003.
- [97] V. Simoncini and D.B. Szyld. Recent computational developments in Krylov subspace methods for linear systems. *Numerical Linear Algebra with Applications*, 14(1):1–59, 2007.
- [98] B. Jegerlehner. Krylov space solvers for shifted linear systems. arXiv:hep-lat/9612014v1, 1996.
- [99] D.B.M. Klaassen, J.W. Slotboom, and H.C. de Graaff. Unified apparent bandgap narrowing in n- and p-type silicon. *Solid State Electronics*, 35(2):125–129, 1992.
- [100] G. Wachutka. Rigorous thermodynamic treatment of heat generation and conduction in semiconductor device modelling. *IEEE Transactions on Computer Aided Design*, 9(11):1141–1149, 1990.
- [101] C. Canali, G. Majni, R. Minder, and G. Ottovani. Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature. *IEEE Transactions on Electron Devices*, 22(11):1045–1047, 1975.
- [102] A. Wettstein, A. Schenk, and W. Fichtner. Quantum device-simulation with the density-gradient model on unstructured grids. *IEEE Transactions on Electron Devices*, 48(2):279–284, 2001.
- [103] R.E. Bank and D.J. Rose. Global approximate Newton methods. *Numerische Mathematik*, 37(2):279–295, 1981.
- [104] R.E. Bank, D.J. Rose, and W. Fichtner. Numerical methods for semiconductor device simulation. *IEEE Transactions on Electron Devices*, 30(9):1031–1041, 1983.
- [105] W. Hansch, P. Borthen, J. Schulze, C. Fink, T. Sulima, and I. Eisele. Performance improvement in vertical surface tunneling transistors by a boron surface phase. *Japanese Journal of Applied Physics*, 40:3131–3136, 2001.
- [106] J. Koga and A. Toriumi. Room temperature negative differential conductance in three-terminal silicon surface tunneling device. In *Technical Digest of the International Electron Devices Meeting (IEDM)*, pages 265–268, San Francisco, CA, USA, 1996.
- [107] K. Boucart and A.M. Ionescu. Double-Gate Tunnel FET With High- κ Gate Dielectric. *IEEE Transactions on Electron Devices*, 54(7):1725–1733, 2007.
- [108] F. Balestra, S. Cristoloveanu, M. Benachir, J. Brini, and T. Elewa. Double-gate silicon-on-insulator transistor with volume inversion: A new device with greatly enhanced performance. *IEEE Electron Device Letters*, 8(9):410–412, 1987.
- [109] R.W. Keyes. High-mobility FET in strained silicon. *IEEE Transactions on Electron Devices*, 33(6):863, 1986.

-
- [110] F.O. Heinz. *Simulation approaches for nanoscale semiconductor devices*. Dissertation, ETH Zürich, 2004.
- [111] M. Luisier. *Quantum transport beyond the effective mass approximation*. Dissertation, ETH Zürich, 2007.
- [112] C.S. Lent and D.J. Kirkner. The quantum transmitting boundary method. *Journal of Applied Physics*, 67(10):6353–6359, 1990.
- [113] L.V. Keldysh. Diagram technique for non-equilibrium processes. *Soviet Physics Journal of Experimental and Theoretical Physics*, 20(1018), 1965.
- [114] P.C. Martin and J. Schwinger. Theory of Many-Particle Systems. I. *Physical Review*, 115:1342–1373, 1959.

Danksagung

Mein abschließender Dank gilt all jenen, die auf unterschiedliche Weise zum Gelingen dieser Arbeit beigetragen haben:

An erster Stelle bedanke ich mich bei meinem Doktorvater Prof. Dr. Gerhard Wachutka für die interessante Themenstellung und die Betreuung meiner Arbeit. Durch seine wertvollen Anregungen entstanden immer wieder neue interessante Ansätze und Ideen. Im Lauf meiner Zeit am Lehrstuhl für Technische Elektrophysik schätzte ich besonders die Freiheit zur wissenschaftlichen Entfaltung und die Möglichkeit aktuelle Ergebnisse vor internationalem Publikum zu präsentieren.

Des Weiteren danke ich Prof. Dr. Walter Hansch vom Institut für Physik der Universität der Bundeswehr für das rege Interesse an meiner Arbeit und die bereitwillige Übernahme des Zweitgutachtens.

Ein großer Dank gilt auch Prof. Dr. Andreas Schenk vom Institut für integrierte Systeme an der Eidgenössischen Technischen Hochschule Zürich für die fachlichen Diskussionen in den Bereichen phononenunterstütztes Band-zu-Band Tunneln und Quantentransport.

Vielen Dank auch an alle meine ehemaligen und noch immer aktiven Mitarbeiter des Lehrstuhls für Technische Elektrophysik für die fachlich unterhaltsamen Diskussionen in der Kaffeerunde und die immer freundschaftliche und entspannte Atmosphäre. Hierbei möchte ich besonders Stefan Schaub und Martin Nießner hervorheben.

Ein besonderer Dank gilt meinen unermüdlichen Korrekturlesern für ihre wertvollen und auch bisweilen kritischen Anmerkungen, die zum Gelingen dieser Arbeit entscheidend beigetragen haben.

Nicht zuletzt danke ich meiner Familie und meinen Freunden, die mich in allen Phasen meines Lebens immer tatkräftig unterstützt haben.