

Kapitel 3

Bauelemente der CMOS-Photosensorik

Der MOS-Transistor und die mit dem CMOS-Prozess verfügbaren lichtempfindlichen Sensorelemente wie die Photodiode, der bipolare Phototransistor und der MOS-Kondensator in Inversion, stellen neben Widerstand und Kondensator die Bauelemente dar, aus denen sich optische CMOS-Sensorsysteme zusammensetzen. In diesem Kapitel werden die Eigenschaften dieser Bauelemente und deren Modellierung beschrieben. Der Grad der Komplexität der vorgestellten Modellgleichungen orientiert sich an den Bedürfnissen des Schaltungsentwicklers in der Konzeptions- und Dimensionierungsphase der Entwicklung.

Große Bedeutung bei der Konzeption und der Dimensionierung von Sensor-Ausleseschaltungen hat das Phänomen Rauschen, da dadurch die Qualität eines Sensorsystems hinsichtlich wichtiger Eigenschaften wie z.B. der Dynamik entscheidend beeinflusst wird. Die verschiedenen Rauschphänomene, die im Zusammenhang mit optischer Sensorik und analoger CMOS-Schaltungstechnik von Bedeutung sind, werden im dritten Teil dieses Kapitels als wichtige Eigenschaften der behandelten Bauelemente und den aus ihnen aufgebauten analogen Systemen vorgestellt.

3.1 MOS Feldeffekt-Transistoren

Der Drainstrom I_d eines MOS-Transistors hängt im allgemeinen in komplizierter Weise von den Potentialen an seinen vier Klemmen (Source, Drain, Gate, Bulk) ab. Für bestimmte Kombinationen von Klemmenspannungen, beziehungsweise für bestimmte Arbeitsbereiche des Transistors, werden jedoch Vereinfachungen möglich, und es können relativ gut zu handhabende Gleichungen für den Drainstrom angegeben werden. Man

unterscheidet beim MOS-Transistor in der Regel die drei Arbeitsbereiche schwache Inversion, moderate Inversion und starke Inversion [7]. Die Bezeichnung des Arbeitsbereiches entspricht jeweils dem Zustand des am stärksten invertierten Bereiches der Halbleiter-Oberfläche unterhalb der Gate-Elektrode des MOS-Transistors. Für die Praxis sind nur die beiden Arbeitsbereiche schwache- und starke Inversion von Bedeutung. Die moderate Inversion wird daher im Rahmen dieser Einführung nicht behandelt.

3.1.1 Der MOS-Transistor in starker Inversion

Die im folgenden vorgestellten Gleichungen gelten für den n-Kanal MOS-Transistor.

Die entsprechenden Gleichungen für den p-Kanal Typ erhält man durch Vorzeichenumkehr aller Spannungs- und Stromgrößen.

Der Arbeitsbereich der starken Inversion stellt sich für Gate-Source-Spannungen U_{gs} oberhalb der Schwellenspannung U_t ein. Die Schwellenspannung des MOS-Transistors ist eine Funktion der Source-Bulk-Spannung U_{sb} , es gilt:

$$U_t = U_{t0} + \gamma(\sqrt{2\Phi_f + U_{sb}} - \sqrt{2\Phi_f}) \quad (3.1)$$

Dabei bezeichnet Φ_f das Fermi-Potential¹, γ die Substrateffektkonstante (ein typischer Wert für einen NMOS-Transistor ist $0,5\sqrt{V}$) und U_{t0} die Schwellenspannung für $U_{sb} = 0V$.

Für $U_{gd} \geq U_t$ bzw. $U_{ds} \leq U_{gs} - U_t$ ist die gesamte Halbleiteroberfläche unterhalb der Gate-Elektrode invertiert, bewegliche Minoritätsladungsträger bilden einen leitenden Kanal zwischen Drain und Source. Der MOS-Transistor arbeitet im sogenannten Triodengebiet. Für den Drainstrom, der in diesem Arbeitsbereich durch Ladungsträgerdrift bestimmt wird, gilt der Zusammenhang:

$$I_d = \mu C'_{ox} \frac{W}{L} [(U_{gs} - U_t)U_{ds} - \frac{1}{2}U_{ds}^2] \quad (3.2)$$

Dabei bezeichnet μ die Beweglichkeit der Minoritätsladungsträger im Kanal für geringe Feldstärken, C'_{ox} die flächenbezogene Kapazität des Gate-Oxids, W die Gate-Weite und L die Gate-Länge.

Für $U_{gd} \leq U_t$ bzw. $U_{ds} \geq U_{gs} - U_t$ wird der leitende Kanal an seinem Drain-Ende abgeschnürt. Der MOS-Transistor arbeitet im sogenannten Sättigungsgebiet. Für den Drainstrom gilt der Zusammenhang:

$$I_d = \frac{1}{2} \mu C'_{ox} \frac{W}{L} (U_{gs} - U_t)^2 (1 + \lambda U_{ds}) \quad (3.3)$$

¹Das Fermi-Potential eines Halbleiters ergibt sich aus der Differenz zwischen der intrinsischen und der extrinsischen Fermi-Energie, es gilt $\phi_f = (E_{F,i} - E_F)/q$ [7, Seiten 19-21]. Ein typischer Wert für einen NMOS-Transistor ist $0,35V$

Die im Sättigungsgebiet mit wachsendem U_{ds} zunehmende Abschnürung des Kanals führt zu einer Reduktion der effektiven Kanallänge und damit zu einem leichten Anstieg des Drain-Stromes. Dieser Effekt (Kanallängenmodulation) wird im Modell durch Multiplikation mit dem Faktor $1 + \lambda U_{ds}$ beschrieben. λ bezeichnet hierbei den Kanallängenmodulationsfaktor.

Für kurze Gate-Längen können im Kanal des MOS-Transistors hohe elektrische Feldstärken auftreten. Die Driftgeschwindigkeit der Ladungsträger im Kanal nimmt dann mit wachsender Feldstärke weniger als proportional zu und kann schließlich für sehr hohe Feldstärken unabhängig vom elektrischen Feld werden (Driftgeschwindigkeitssättigung). Mit Hilfe des Beweglichkeitsreduktionsfaktors Θ läßt sich dieser Effekt im Modell wie folgt beschreiben [8]:

$$I_d = \frac{\mu C'_{ox}}{2(1 + \Theta(U_{gs} - U_t))} \frac{W}{L} (U_{gs} - U_t)^2 (1 + \lambda U_{ds}) \quad (3.4)$$

3.1.2 Der MOS-Transistor in schwacher Inversion

Für $U_{gs} < U_t$ und $U_{gd} < U_t$ reicht die Größe des Oberflächenpotentials ψ_s unterhalb der Gate-Elektrode nicht aus, um die durch die Drain- und Sourceseitigen pn-Übergänge hervorgerufene Potentialbarriere abzubauen. Der MOS-Transistor arbeitet in schwacher Inversion. Der Drainstrom beruht in diesem Fall auf der Diffusion von Minoritätsladungsträgern, und es ergibt sich analog zum Verhalten bipolarer Bauelemente eine exponentielle Abhängigkeit des Drainstromes von den Potentialen an den Klemmen des Transistors.

Die im Rahmen dieser Arbeit vorgestellten Schaltungskonzepte beruhen zum Teil wesentlich auf den speziellen Eigenschaften des MOS-Transistors in schwacher Inversion.

Da diese Eigenschaften in der schaltungstechnischen Literatur zum einen oft nur am Rande Erwähnung finden und zum anderen die angegebenen Modellgleichungen teilweise erheblich voneinander abweichen, werden die physikalischen Grundlagen der vorgestellten Modellgleichungen, Hinweise zu deren Herleitung, deren Genauigkeit und den Einschränkungen in bezug auf deren Verwendbarkeit für diesen Arbeitsbereich im Anhang ausführlich behandelt.

Entsprechend Gleichung (A.17) im Anhang beziehungsweise nach [7, Seite 139] kann der Drainstrom eines MOS-Transistors in schwacher Inversion durch den Zusammenhang

$$I_d = I_{\max} e^{\frac{U_{gs} - U_t}{nU_T}} \left(1 - e^{-\frac{U_{ds}}{U_T}} \right) \quad (3.5)$$

modelliert werden, wobei n den sogenannten „weak inversion slope factor“ bezeichnet, der entsprechend den Gleichungen (A.8) und (A.10) im Anhang gegeben ist durch

$$n = 1 + \frac{\gamma}{2\sqrt{U_{sb} + 1,5\Phi_f}}. \quad (3.6)$$

Für $U_{sb} = 0V$ beträgt n etwa 1,5. $U_T = kT/q$ bezeichnet die Temperaturspannung (diese hat bei Zimmertemperatur den Wert 26 mV). I_{max} bezeichnet den maximalen Drainstrom in schwacher Inversion und ist nach Gleichung (A.6) gegeben durch

$$I_{max} = \mu C'_{ox} \frac{W}{L} U_T^2 \frac{\gamma}{2\sqrt{U_{sb} + 2\Phi_f}}. \quad (3.7)$$

Es sollte an dieser Stelle darauf hingewiesen werden, daß entsprechend den Ausführungen im Anhang n , I_{max} und U_t vom Sourcepotential U_{sb} abhängen. Gleichung (3.5) sollte daher nur dann angewendet werden, wenn das Sourcepotential konstant ist.

3.1.3 Das Kleinsignalverhalten des MOS-Transistors

Im Betrieb eines Verstärkerbauelementes wie dem MOS-Transistor wird der in einem Großsignal-Arbeitspunkt am Eingang des Bauelementes anliegenden Gleichspannung ein zu verstärkendes Wechsignale überlagert. Für beliebige Eingangssignale ist die Berechnung des zeitlichen Verlaufes des Ausgangssignals aufgrund der nichtlinearen Bauelemente-Kennlinien sehr aufwendig und in der Regel nur numerisch durchführbar.

Für genügend kleine Signalamplituden lassen sich die Eingangs- und Ausgangskennlinien eines Bauelementes in einer Umgebung des Arbeitspunktes linearisieren. Das dynamische Bauelementeverhalten für kleine Signale wird dann durch die Steigung der Kennlinien im Arbeitspunkt beschrieben. Diese Vorgehensweise führt auf das sogenannte Kleinsignal-Ersatzschaltbild eines Bauelementes.

Komponenten des Kleinsignal-Ersatzschaltbildes eines MOS-Transistors (Abbildung 3.1) sind die parasitären Kapazitäten zwischen den Klemmen des Bauelementes, der differentielle Ausgangsleitwert und verschiedene spannungsgesteuerte Stromquellen, die den Einfluß kleiner Variationen der Klemmenspannungen auf den Ausgangsstrom beschreiben. Zu betonen ist, daß die Parameter des Kleinsignalersatzschaltbildes (auch die Kapazitäten [7, Seite 325]) nur für einen bestimmten Arbeitspunkt und einen beschränkten Frequenzbereich gelten.

In Abbildung 3.1 bezeichnet g_m die Transkonduktanz (Vorwärtssteilheit) des MOS-Transistors. Sie beschreibt den Einfluß kleiner Variationen der Gate-Source Spannung auf den Drainstrom:

$$g_m = \left. \frac{\partial I_d}{\partial U_{gs}} \right|_{\text{Arbeitspunkt}}. \quad (3.8)$$

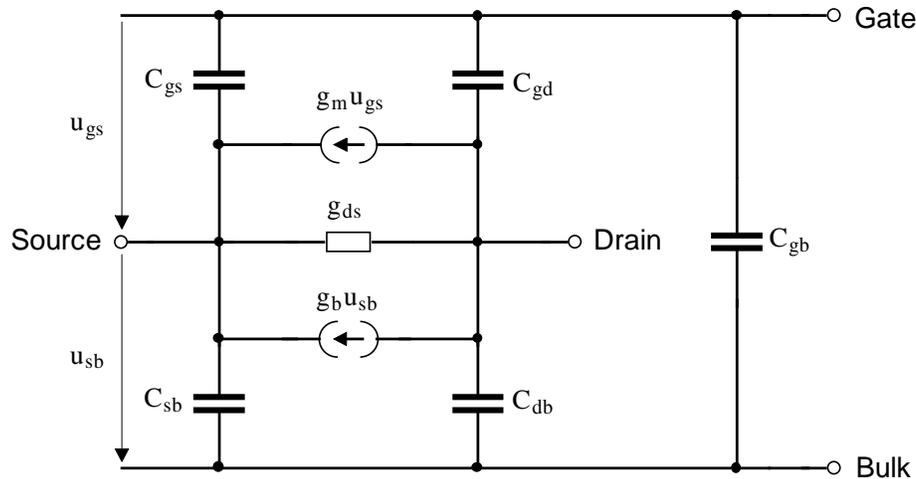


Abbildung 3.1: Kleinsignalersatzschaltbild des MOS-Transistors.

g_b bezeichnet die Substratsteilheit des MOS-Transistors und beschreibt den Einfluß kleiner Variationen der Source-Bulk Spannung auf den Drainstrom:

$$g_b = \left. \frac{\partial I_d}{\partial U_{sb}} \right|_{\text{Arbeitspunkt}} \quad (3.9)$$

g_{ds} bezeichnet den Ausgangsleitwert des MOS-Transistors und gibt die Steigung der Ausgangskennlinie im Arbeitspunkt an:

$$g_{ds} = \left. \frac{\partial I_d}{\partial U_{ds}} \right|_{\text{Arbeitspunkt}} \quad (3.10)$$

Tabelle 3.1 gibt die Berechnungsvorschriften für die wichtigsten Kleinsignalparameter des MOS-Transistors an. ΔL bezeichnet dabei die Überlappung der Gate-Elektrode über die Diffusionsgebiete (Unterdiffusionslänge).

3.1.4 Matching-Eigenschaften von MOS-Transistoren

Die Modellparameter eines MOS-Transistors in einem bestimmten Arbeitspunkt können nicht als Konstante angenommen werden. Messungen [9–11] zeigen vielmehr, daß die Parameter gleich gestalteter Transistoren in Abhängigkeit von ihrer Position auf dem Wafer zeitunabhängigen statistischen Schwankungen unterliegen. Diese auf eine räumliche Inhomogenität der zugrundeliegenden Technologieparameter sowie statistische Schwankungen der Bauelementegeometrie zurückzuführenden Variationen der Transistorparameter werden als „Mismatch“ bezeichnet. Sie stellen einen limitierenden Faktor für die Verwendbarkeit insbesondere analoger CMOS-Schaltungen dar und müssen beim Schaltungsdesign berücksichtigt werden.

Tabelle 3.1: Kleinsignalparameter des MOS-Transistors.

	<i>starke Inversion</i>		<i>schwache Inversion</i>
	<i>Sättigungs-Gebiet</i>	<i>Trioden-Gebiet</i>	<i>Sättigungs-Gebiet</i> ($U_{ds} \gg U_T$)
g_m	$\sqrt{2\mu C'_{ox} \frac{W}{L} (1 + \lambda U_{ds}) I_d}$	$\mu C'_{ox} \frac{W}{L} U_{ds}$	$\frac{I_d}{nU_T}$
g_b	$-g_m \gamma / (2\sqrt{2\Phi_f + U_{sb}})$		$-\frac{I_d}{U_T}$
g_{ds}	$\frac{\lambda I_d}{1 + \lambda U_{ds}}$	$\mu C'_{ox} \frac{W}{L} (U_{gd} - U_t)$	$\propto \frac{I_d}{L}$
C_{gs}	$\frac{2}{3} C_{ox}$	$\frac{1}{2} C_{ox}$	$W \Delta L C'_{ox}$
C_{gd}	$W \Delta L C'_{ox}$	$\frac{1}{2} C_{ox}$	$W \Delta L C'_{ox}$

Bei der Entwicklung zweidimensionaler CMOS-Bildsensoren stellt die Beherrschung des auf dem Mismatching von Transistoren beruhenden Ortsrauschens („fixed pattern noise“) eines der Hauptprobleme dar. Die Kompensation der auch bei räumlich völlig homogener Beleuchtung zwischen verschiedenen Pixeln auftretenden Differenzen im Ausgangssignal ist mit großem Aufwand verbunden, da bei der Auslese des Sensorarrays für jedes Pixel eine individuelle Korrektur der Sensorkennlinie vorgenommen werden muß.

In der Fachliteratur [10, 12, 13] werden als mögliche Ursachen für das Mismatching von MOS-Transistoren die folgenden Effekte genannt:

- *Statistische Variationen der Transistorgeometrie.* Die Geometrie der Gate-Fläche eines MOS-Transistors wird im Lithographieprozeß durch mehrere Masken bestimmt. Die Gate-Weite wird durch eine Diffusionsmaske-, die Gate-Länge durch die Polysilizium-Maske festgelegt. Da die Herstellung der einzelnen Masken und die relative Positionierung der Masken zueinander fehlerbehaftet sind, treten voneinander unabhängige Streuungen der Gate-Längen L und der Gate-Weiten W auf. Dieser Effekt verursacht Abweichungen in der Steilheit β eines Transistors.
- *Statistische Variationen der Oberflächenzustandsdichte und der Ionen-Implantation.* Da die Oberflächenzustände und die in den Kanalbereich eingebrachten Implantationsladungen nicht beliebig fein verteilt, sondern quantisiert vorliegen, ist ihre Gesamtzahl unterhalb der Gatefläche unvermeidlichen statistischen Streuungen unterworfen. Durch diesen Effekt werden Streuungen der Flachbandspannung U_{FB} und folglich der Schwellenspannung U_t verursacht.

- *Statistische Variationen der Oxiddicke.* Lokale Schwankungen in der Dicke des Gate-Oxids und damit im Kapazitätsbelag der Gate-Elektrode gegenüber der Halbleiteroberfläche verursachen Variationen der Steilheit β , der Substrateffektkonstante γ sowie der Flachbandspannung U_{FB} .
- *Statistische Variationen der Ladungsträger-Beweglichkeit im Kanal.* Diese beruhen unter anderem auf den statistischen Streuungsmechanismen von Ladungsträgern an Verunreinigungen im Halbleiter und bewirken Abweichungen in der Steilheit β eines MOS-Transistors.
- *Statistische Variationen der Dotierung im Kanal.* Streuungen in der Gesamtzahl der Dotierstoffatome in der Verarmungszone unterhalb der Gate-Elektrode sind unvermeidlich und verursachen Variationen der Substrateffektkonstante γ und des Fermi-Potentials Φ_f der Halbleiteroberfläche im Kanal.

Messungen zur Charakterisierung der Matching-Eigenschaften von MOS-Transistoren in einem $2,5\mu\text{m}$ CMOS-Prozeß, veröffentlicht von Pelgrom, Duinmaijer und Welbers im Jahre 1989 in [10], führten unter anderem zu den folgenden Ergebnissen:

- Entsprechend Gleichung (3.13) zeigte sich, daß die relativen Abweichungen des Drainstromes I_d für kleine Gate-Source Spannungen U_{gs} vor allem durch die Streuungen der Schwellenspannung U_t bestimmt werden. Dies gilt insbesondere auch für den Arbeitsbereich der schwachen Inversion.
- Es wurde bestätigt, daß die Varianzen der Schwellenspannung U_t und der Steilheit β sich in guter Näherung umgekehrt proportional zur Gate-Fläche $W \cdot L$ verhalten.
- Es zeigte sich, daß der räumliche Abstand zweier Transistoren auf dem Wafer nur eine geringe Auswirkung auf den Parameter-Mismatch der beiden Transistoren hat.
- Die Autoren vermuteten als Ursachen für die Streuungen der Schwellenspannung U_t Variationen in der Substratdotierung, der Oberflächenzustandsdichte und der Dichte Implantierter Ionen. Als mögliche Ursache für Streuungen der Steilheit β wurden Schwankungen der lokalen Ladungsträgerbeweglichkeit genannt. Nach Ansicht der Autoren haben Variationen der Geometrieparameter keinen wesentlichen Einfluß auf die Matching-Eigenschaften von MOS-Transistoren.

Eine neuere Veröffentlichung aus dem Jahre 1994 [12] kommt zu dem Ergebnis, daß Streuungen der Schwellenspannung U_t vor allem durch lokale Variationen der Dotierung im Kanal bedingt sind.

Im folgenden soll der Einfluß der Streuung einzelner Modellparameter auf den Drainstrom für verschiedene Arbeitsbereiche untersucht werden. Mit der Steilheit $\beta = \mu C'_{\text{ox}} W/L$ ist entsprechend Gleichung (3.3) der Drainstrom eines in Sättigung arbeitenden MOS-Transistors in starker Inversion bei Vernachlässigung der Kanallängenmodulation gegeben durch die vereinfachte Beziehung

$$I_d \approx \frac{\beta}{2} (U_{\text{gs}} - U_t)^2. \quad (3.11)$$

Entsprechend dem Gaußschen Fehlerfortpflanzungsgesetz ergibt sich dann die Varianz des Drainstromes zu

$$\sigma_{I_d}^2 = \left(\frac{\partial I_d}{\partial \beta} \Big|_{\text{AP}} \right)^2 \sigma_{\beta}^2 + \left(\frac{\partial I_d}{\partial U_t} \Big|_{\text{AP}} \right)^2 \sigma_{U_t}^2. \quad (3.12)$$

Diese Gleichung gilt streng genommen nur für dem Fall, daß die Abweichungen der Steilheit und der Schwellenspannung unkorreliert sind. Messungen [10] zeigen, daß diese Voraussetzung in guter Näherung erfüllt ist. Nach partieller Differentiation der Gleichung (3.11) nach β und U_t ergibt sich dann:

$$\frac{\sigma_{I_d}^2}{I_d^2} = \frac{\sigma_{\beta}^2}{\beta^2} + \frac{4\sigma_{U_t}^2}{(U_{\text{gs}} - U_t)^2}. \quad (3.13)$$

Sind die Standardabweichungen der Steilheit und der Schwellenspannung bekannt, so erlaubt diese Beziehung für einen bestimmten Arbeitspunkt die Berechnung des relativen mittleren quadratischen Fehlers des Drainstromes. Desweiteren verdeutlicht die Gleichung, daß für hohe Gate-Source Spannungen die besten Matching-Eigenschaften für den Drainstrom erzielt werden. Für den Sättigungsstrom in schwacher Inversion ergibt sich aus den Gleichungen (A.6) und (A.18) die Beziehung

$$I_d \approx K \frac{\beta}{2} e^{\frac{U_{\text{gs}} - U_t}{nU_T}}, \quad K = \gamma U_T^2 / \sqrt{U_{\text{sb}} + 2\Phi_f}. \quad (3.14)$$

Setzt man entsprechend [10] analog zu oben unkorrelierte Abweichungen von K , β und U_t voraus, so ergibt das Gaußsche Fehlerfortpflanzungsgesetz

$$\sigma_{I_d}^2 = \left(\frac{\partial I_d}{\partial K} \Big|_{\text{AP}} \right)^2 \sigma_K^2 + \left(\frac{\partial I_d}{\partial \beta} \Big|_{\text{AP}} \right)^2 \sigma_{\beta}^2 + \left(\frac{\partial I_d}{\partial U_t} \Big|_{\text{AP}} \right)^2 \sigma_{U_t}^2 \quad (3.15)$$

und daraus schließlich

$$\frac{\sigma_{I_d}^2}{I_d^2} = \frac{\sigma_K^2}{K^2} + \frac{\sigma_{\beta}^2}{\beta^2} + \frac{\sigma_{U_t}^2}{(nU_T)^2}. \quad (3.16)$$

Ein Vergleich der Gleichungen (3.13) und (3.16) zeigt, daß Variationen der Schwellenspannung in schwacher Inversion maximal zur Varianz des Drainstromes beitragen. Dieser wichtige Zusammenhang ist immer dann zu berücksichtigen, wenn der Drainstrom eines

MOS-Transistors über ein festes Gate-Potential eingestellt wird. Dieses ist unter anderem bei der Realisierung von Stromspiegeln der Fall, bei denen zwei oder mehrere Transistoren mit gleicher Gate-Source Spannung betrieben werden. In schwacher Inversion weisen die durch verschiedene Transistoren eines Stromspiegels fließenden Ströme maximale relative Abweichungen auf. Wird dagegen der Arbeitspunkt über einen festen Drainstrom eingestellt, so ergibt sich für die Varianz der Gate-Source Spannung in starker Inversion

$$\sigma_{U_{gs}}^2 = \sigma_{U_t}^2 + \frac{(U_{gs,n} - U_t)^2}{4} \frac{\sigma_\beta^2}{\beta^2}. \quad (3.17)$$

$U_{gs,n}$ bezeichnet hierbei den Nominalwert der Gate-Source Spannung im Arbeitspunkt. In schwacher Inversion gilt entsprechend

$$\sigma_{U_{gs}}^2 = \sigma_{U_t}^2 + (nU_T)^2 \left(\frac{\sigma_\beta^2}{\beta^2} + \frac{\sigma_K^2}{K^2} \right). \quad (3.18)$$

Sieht man von den relativen Abweichungen des Faktors K ab (diese liegen nach [10] in der gleichen Größenordnung wie die der Steilheit β), so ergibt ein Vergleich der Gleichungen (3.17) und (3.18), daß nun die Transistoren in schwacher Inversion die besseren Matching-Eigenschaften aufweisen. Die relativen Abweichungen der Steilheit tragen in starker Inversion mit größerem Gewicht zur Varianz der Gate-Source Spannung bei. Dies kann zum Beispiel bei der Realisierung von Differenzstufen mit geringem Offset ausgenutzt werden.

Statistische Fluktuationen der Schwellenspannung U_t und der Steilheit β beeinflussen sowohl das Groß-, als auch das Kleinsignalverhalten eines MOS-Transistors, da die Kleinsignal-Modellparameter vom jeweiligen Arbeitspunkt abhängig sind (siehe Tabelle 3.1).

3.2 Lichtempfindliche Bauelemente im CMOS-Prozeß

3.2.1 Übersicht

Der Standard CMOS-Prozeß bietet zur Realisierung lichtempfindlicher Bauelemente zwei prinzipiell verschiedene Möglichkeiten. In einem n-Wannen Standard-CMOS-Prozeß stehen zum einen die drei in Abbildung 3.2 dargestellten lichtempfindlichen pn-Übergänge n-Diffusion/p-Substrat (a), n-Wanne/p-Substrat (b) und p-Diffusion/n-Wanne (c) zur Realisierung von pn-Photodioden, sowie ein lichtempfindlicher Bipolartransistor (d) zur Verfügung. Auf der Grundlage einer Wannendiode kann desweiteren der in Abbildung 3.3

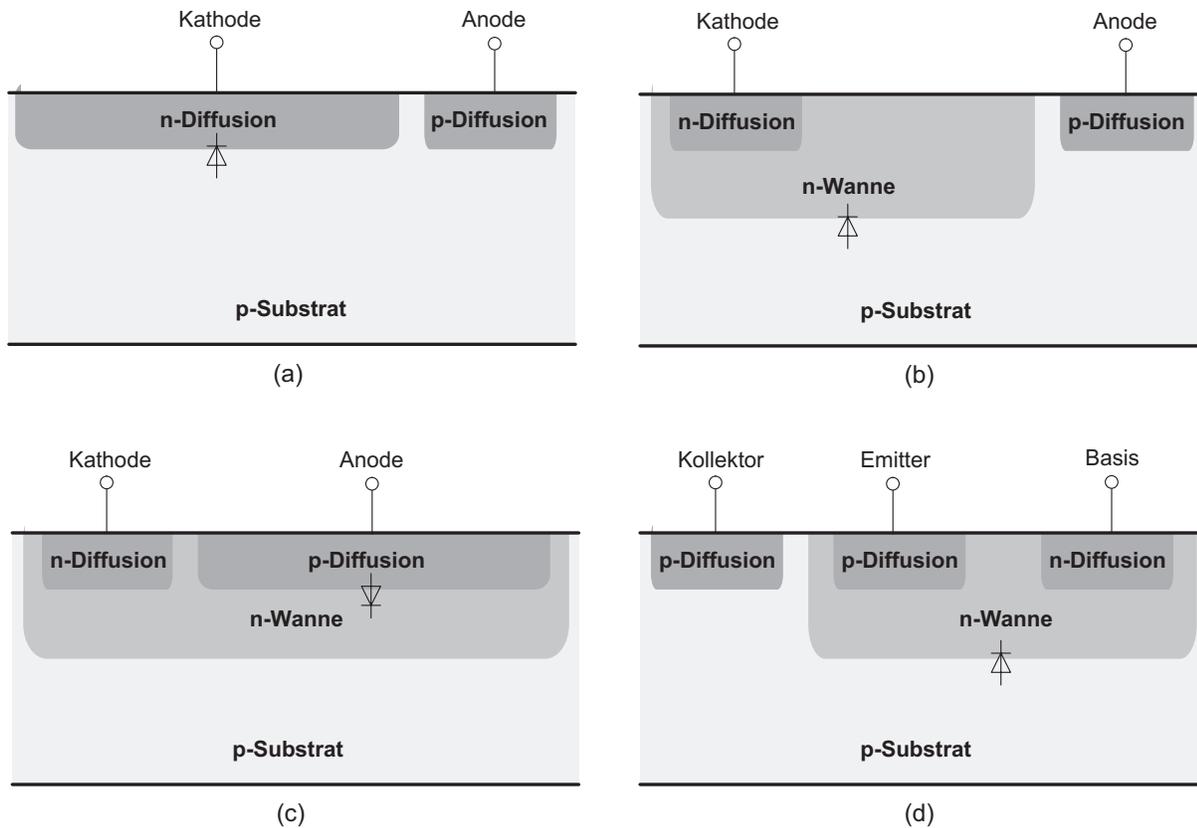


Abbildung 3.2: Im n-Wannen Standard-CMOS-Prozeß realisierbare lichtempfindliche bipolare Bauelemente. Die eingezeichneten Dioden symbolisieren jeweils den lichtempfindlichen pn-Übergang.

dargestellte Photo-MOSFET, der für die zweidimensionale Bildsensorik jedoch nicht von Bedeutung ist, realisiert werden.

In Anlehnung an die Technologie von CCD-Bildsensoren besteht zum anderen die Möglichkeit, optisch generierte Ladungen in einer Potentialwanne unterhalb der Gate-Elektrode eines MOS-Kondensators zu speichern. Diese Bauelemente, in der Fachliteratur als Photogates bezeichnet, werden in den vom Jet Propulsion Lab und der daraus hervorgegangenen Firma Photobit entwickelten Bildsensoren verwendet [14–19]. Da der lichtempfindliche Bereich dieser Sensoren von einer Polysilizium-Steuerlektrode abgedeckt wird, kommt es zu Mehrfachreflexionen an den Grenzflächen zwischen Polysilizium mit hohem Brechungsindex und Siliziumdioxid mit niedrigem Brechungsindex. Zusätzlich wird insbesondere der Blauanteil im Polysilizium teilweise absorbiert [20]. Es ergeben sich somit für das Photogate schlechtere Quantenwirkungsgrade als für die pn-Photodioden [21].

Die Realisierung eines lichtempfindlichen pnp-Bipolartransistors kann entsprechend

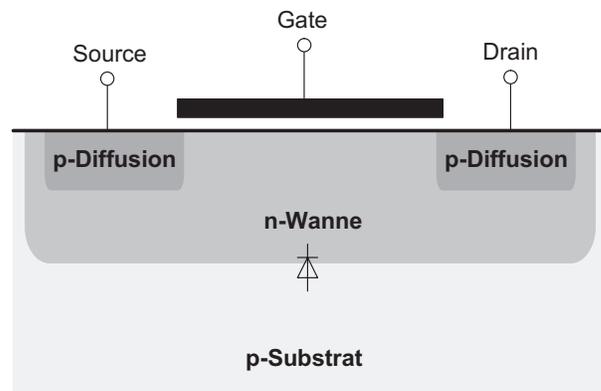


Abbildung 3.3: Im n-Wannen Standard-CMOS-Prozeß realisierbarer Photo-MOSFET.

Abbildung 3.2 in der Form p-Diffusion/n-Wanne/p-Substrat erfolgen. Die p-Diffusion dient dann als Emitter, während der Kollektor durch das p-Substrat außerhalb der Wanne gegeben ist. Die Basis (n-Wanne) wird zur Realisierung von Phototransistoren in der Regel nicht kontaktiert [22]. Der im lichtempfindlichen pn-Übergang zwischen n-Wanne und p-Substrat generierte Photostrom wird im Bauelement verstärkt und erzeugt einen um den Faktor der Stromverstärkung höheren Kollektorstrom. Da die Stromverstärkung eines Bipolartransistors vom Arbeitspunkt abhängig ist [8], ergibt sich eine nichtlineare Abhängigkeit des Kollektorstromes vom fließenden Photostrom. Messungen, die am FhG IMS an bipolaren Phototransistoren durchgeführt wurden, ergaben je nach einfallender Bestrahlungsstärke Stromverstärkungen zwischen 5 und 50 [23]. Neben einer schlechten Linearität weisen bipolare Phototransistoren aufgrund der hochohmig betriebenen Basis schlechte Eigenschaften bezüglich Geschwindigkeit und Rauschen auf. Bipolare Photosensoren werden unter anderem in den Bildsensoren der aus dem California Institute of Technology hervorgegangenen Firma Synapsis verwendet [24].

Photo-MOSFETs bestehen entsprechend Abbildung 3.3 aus einem p-Kanal MOS-Transistor mit unkontaktierter n-Wanne. Die im pn-Übergang zwischen n-Wanne und p-Substrat optisch generierten freien Elektronen driften in die Wanne, werden dort gesammelt und bewirken eine Verringerung des Wannepotentials. Das gegenüber der Source-Elektrode veränderte Wannepotential bewirkt über den Substrateffekt eine Modulation des Drainstromes. Photo-MOSFET's wurden in den am FhG IMS entwickelten Zeilenkameras mit Erfolg angewendet [25, 26]. Die Übertragungscharakteristik dieser Sensorbauelemente kann durch geeignete Wahl des Arbeitspunktes eingestellt werden. Für den Arbeitsbereich der schwachen Inversion ergibt sich ein lineares Übertragungsverhalten, während die Bauelemente in starker Inversion eine logarithmische Charakteristik aufweisen. Zur Verwendung in zweidimensionalen Bildsensoren sind sie unter anderem aufgrund

der für kleine Gateflächen problematischen Matching-Eigenschaften nicht geeignet.

Die im Rahmen dieser Arbeit realisierten CMOS-Bildsensoren verwenden als lichtempfindliche Bauelemente den pn-Übergang zwischen n-Diffusionsgebiet und p-Substrat. Dieser weist gegenüber den anderen oben beschriebenen Alternativen wegen der fehlenden Notwendigkeit einer n-Wanne den geringsten Platzbedarf auf. Die Abstände zwischen zwei benachbarten n-Wannen müssen aufgrund der CMOS-Designregeln sehr viel größer gewählt werden als für zwei benachbarte Diffusionsgebiete.

3.2.2 Strahlungsabsorption im Halbleiter

Halbleiter-Photodetektoren beruhen auf der Generation frei beweglicher Elektron-Loch Paare infolge von Photon-Elektron Wechselwirkungen. Hierbei verliert das betroffene Photon seine Energie ($W_{\text{photon}} = h\nu$) an ein Elektron, das auf ein höheres Energieniveau angehoben wird. Je nach dem, ob die freigesetzten Elektronen im Material verweilen oder die Oberfläche des Materials verlassen, spricht man vom inneren oder vom äußeren Photoeffekt. Im Zusammenhang mit den im vorigen Abschnitt vorgestellten Bauelementen ist nur der im folgenden behandelte innere Photoeffekt von Bedeutung.

Damit ein Elektron infolge der Wechselwirkung mit einem Photon vom Valenzband in einen frei beweglichen Zustand im Leitungsband angehoben werden kann, muß die Energie des beteiligten Photons mindestens dem Bandabstand des jeweiligen Halbleitermaterials entsprechen, d. h.

$$W_{\text{photon}} = h\nu \geq W_{\text{gap}} . \quad (3.19)$$

Aus dieser Bedingung ergibt sich für die in das Material eindringende Strahlung eine Grenzwellenlänge (Absorptionskante), oberhalb der keine optisch induzierten Band-Band Übergänge möglich sind. Es gilt:

$$\lambda_{\text{max}} = \frac{hc}{W_{\text{gap}}} . \quad (3.20)$$

Für Silizium ergibt sich mit $W_{\text{gap}} = 1,1\text{eV}$ bei $T = 300\text{K}$ eine Grenzwellenlänge von $1,1\mu\text{m}$. Strahlung mit einer Wellenlänge größer λ_{max} wird nicht absorbiert, das jeweilige Material erscheint für die Strahlung transparent.

Das physikalische Modell zur Beschreibung des Photoeffektes beruht auf der Wechselwirkung von Materie mit Photonen. Es ist daher für die weitere Behandlung dieses Effektes sinnvoll, die Intensität einer Strahlung bezüglich des mit ihr verbundenen mittleren Photonenstromes auszudrücken. Bezieht man die Anzahl Δn der innerhalb eines Zeitintervalls Δt auf eine Bezugsfläche ΔA auftreffenden Photonen auf das Zeitintervall

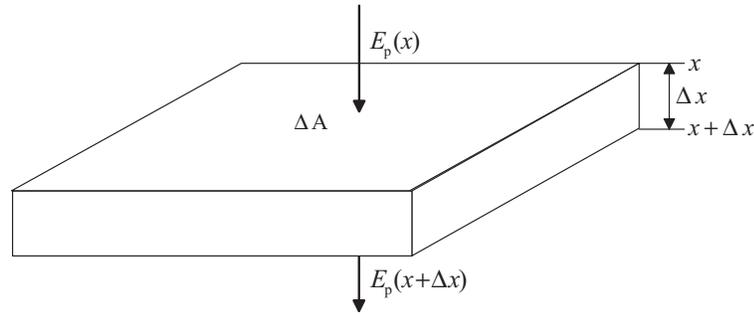


Abbildung 3.4: Zur Strahlungsabsorption im Halbleiter.

Δt und auf die Bezugsfläche ΔA , so erhält man nach DIN 5031 die Photonenbestrahlungsstärke E_p , mit der die Bezugsfläche beaufschlagt wird. Es gilt:

$$E_p = \frac{\Delta n}{\Delta t \cdot \Delta A} . \quad (3.21)$$

Berücksichtigt man, daß pro Photon eine Energiemenge $h\nu = h \cdot (c/\lambda)$ transportiert wird, so ist die spektrale Bestrahlungsstärke $E_{e,\lambda}$ mit der korrespondierenden spektralen Photonenbestrahlungsstärke $E_{p,\lambda}$ über die folgende Beziehung verknüpft:

$$E_{e,\lambda} = E_{p,\lambda} \frac{h c}{\lambda} . \quad (3.22)$$

Für spektral breitbandige Strahlung ergibt sich die Photonenbestrahlungsstärke E_p durch Integration von $E_{p,\lambda}$ über die Wellenlänge.

Im folgenden sollen die makroskopischen Gesetzmäßigkeiten der Strahlungsabsorption und darauf aufbauend die Gesetzmäßigkeiten für die optische Generation von Ladungsträgerpaaren erläutert werden. Die Anzahl der optisch generierten Elektron-Loch Paare pro Volumen- und Zeiteinheit wird im folgenden mit der Generationsrate $g(x)$ bezeichnet. In dem in Abbildung 3.4 dargestellten Volumenelement $\Delta V = \Delta A \cdot \Delta x$ werden $g(x) \cdot \Delta A \cdot \Delta x$ Ladungsträgerpaare pro Zeiteinheit generiert. Dies entspricht der Anzahl der im Volumenelement absorbierten Photonen pro Zeiteinheit beziehungsweise der negativen Änderung $\Delta\Phi_p(x)$ des Photonennflusses über der Strecke Δx . Durch Division von $\Delta\Phi_p(x)$ durch das Flächenelement ΔA ergibt sich die Änderung $\Delta E_p(x)$ der Photonenbestrahlungsstärke über Δx :

$$\Delta E_p(x) = -g(x) \cdot \Delta x \quad (3.23)$$

Für den Gradienten der Photonenbestrahlungsstärke in x-Richtung gilt damit die Beziehung:

$$\frac{dE_p(x)}{dx} = -g(x) \quad (3.24)$$

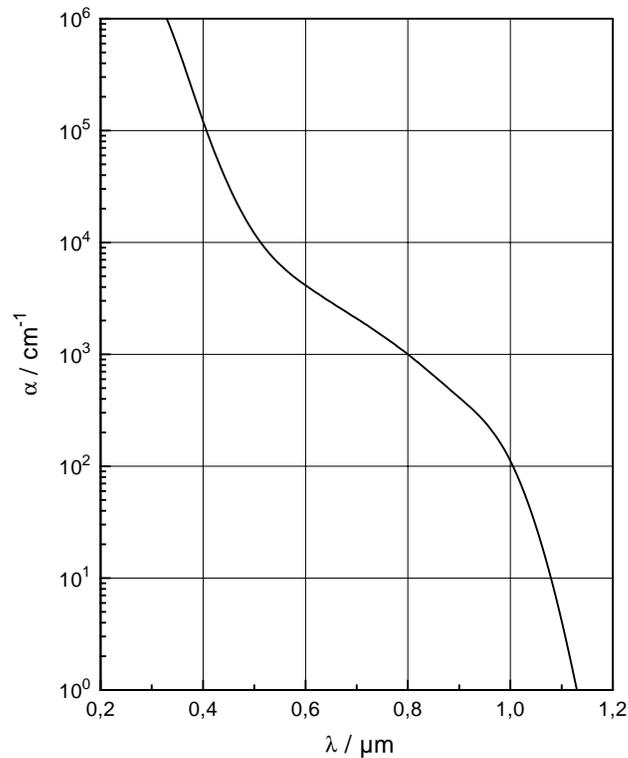


Abbildung 3.5: Absorptionskoeffizient α von Silizium für $T = 300\text{K}$.

Beim Durchtritt von Strahlung durch einen Festkörper ist die Wahrscheinlichkeit der Wechselwirkung mit einem Elektron für jedes eindringende Photon gleich. Die Anzahl der in einem Volumenelement ΔV absorbierten Photonen pro Zeiteinheit muß daher proportional der Anzahl in das Volumenelement einströmender Photonen pro Zeiteinheit sein. Auf das Volumenelement ΔV bezogen gilt damit unter Berücksichtigung von Gleichung (3.24) der folgende Zusammenhang:

$$\frac{dE_p(x)}{dx} = -\alpha E_p(x) = -g(x). \quad (3.25)$$

Der Faktor α bezeichnet den Absorptionskoeffizienten des jeweiligen Materials. Sein Kehrwert ist ein Maß für die Eindringtiefe von Strahlung in den Festkörper. Abbildung 3.5 zeigt den Verlauf des Absorptionskoeffizienten von Silizium als Funktion der Wellenlänge für eine Temperatur von 300K. Die Zunahme des Absorptionskoeffizienten mit abnehmender Wellenlänge (also mit zunehmender Photonenenergie) ist im wesentlichen darauf zurückzuführen, daß im Leitungsband die Dichte der verfügbaren freien Zustände, die ein angeregtes Elektron einnehmen kann, mit wachsender potentieller Energie zunimmt. Nimmt man an, daß Strahlung an der Stelle $x = 0$ mit der Photonenbestrahlungsstärke $E_p(0)$ in das Material eindringt, so ergibt sich durch Integration von Gleichung (3.25) das

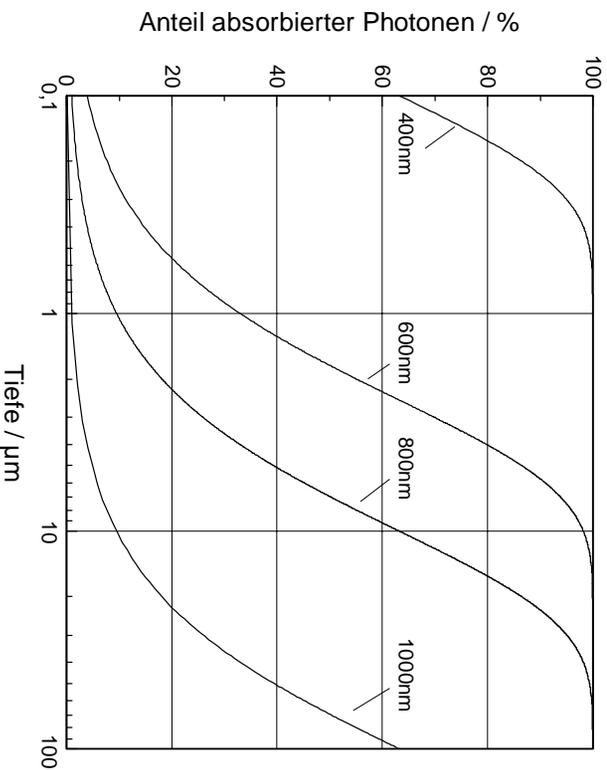


Abbildung 3.6: Prozentualer Anteil $100 \cdot G(0, d_2)/E_p(0)$ absorbierter Photonen in Abhängigkeit von der Tiefe d_2 im Material (mit der Wellenlänge als Parameter).

makroskopische Absorptionsgesetz:

$$E_p(x) = E_p(0) e^{-\alpha x} . \quad (3.26)$$

Aus den Gleichungen (3.25) und (3.26) ergibt sich für die tiefenabhängige optische Generationsrate:

$$g(x) = \alpha E_p(0) e^{-\alpha x} . \quad (3.27)$$

Die auf das bestrahlte Flächenelement ΔA bezogene Gesamtzahl $G(d_1, d_2)$ aller in einer durch $d_1 \leq x \leq d_2$ begrenzten Schicht generierten Elektron-Loch Paare pro Zeiteinheit ergibt sich durch Integration der Generationsrate $g(x)$ in den Grenzen d_1 und d_2 :

$$G(d_1, d_2) = \int_{d_1}^{d_2} \alpha E_p(0) e^{-\alpha x} dx = E_p(0) \left(e^{-\alpha d_1} - e^{-\alpha d_2} \right) . \quad (3.28)$$

Abbildung 3.6 zeigt entsprechend Gleichung (3.28) bei $d_1 = 0$ für vier verschiedene Wellenlängen den prozentualen Anteil $100 \cdot G(0, d_2)/E_p(0)$ absorbierter Photonen in Abhängigkeit von der Tiefe d_2 im Material (Silizium). Es wird unter anderem deutlich, daß bereits innerhalb der ersten $10 \mu\text{m}$ unterhalb der Siliziumoberfläche Strahlung des sichtbaren Spektralbereiches fast vollständig absorbiert wird.

3.2.3 Charakterisierung und Wirkungsweise von pn-Photodioden

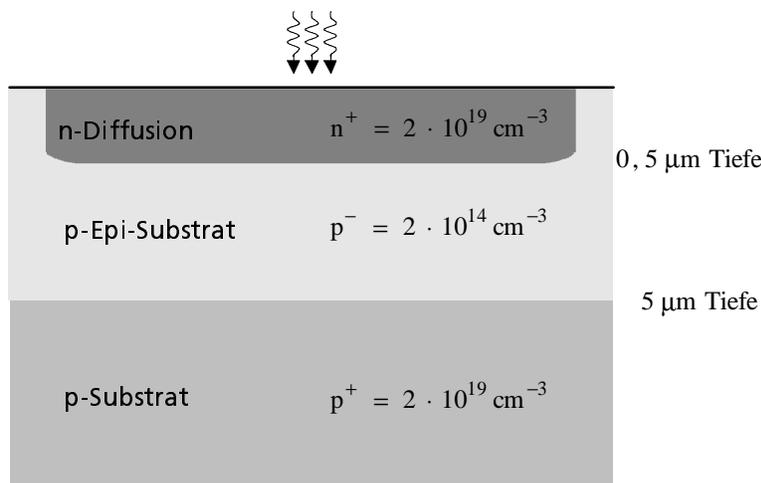


Abbildung 3.7: Aufbau einer n+/p Photodiode in CMOS-Technologie.

Im Rahmen dieses Abschnittes sollen neben einer einfachen Beschreibung der Wirkungsweise von pn-Photodioden die für ihre quantitative Charakterisierung benötigten Größen sowie die für die Entwicklung optischer Sensorsysteme wichtigsten Modellgleichungen vorgestellt werden. Abbildung 3.7 zeigt typische Werte für die Dicke und die Dotierung der verschiedenen Halbleiterschichten der im Rahmen dieser Arbeit verwendeten n-Diffusion/p-Substrat-Photodiode.

Die Effektivität der Wandlung einer einfallenden Strahlungsenergie in elektrische Energie wird in der pn-Photodiode durch verschiedene Verlustmechanismen, die teilweise Reflexion des einfallenden Strahlungsflusses an der Halbleiteroberfläche, die unvollständige Absorption eindringender Strahlungsenergie im Halbleiterinneren und die Rekombination optisch generierter Ladungsträgerpaare, begrenzt. Der Einfluß dieser Effekte kann mit Hilfe eines wellenlängenabhängigen Quantenwirkungsgrades η , der wie folgt definiert ist, beschrieben werden.

$$\eta = \frac{\text{Anzahl der zum Photostrom beitragenden Ladungsträgerpaare}}{\text{Anzahl der auf die Diodenoberfläche auftreffenden Photonen}} \quad (3.29)$$

Die Erzeugung eines Photostromes im pn-Übergang beruht auf der Trennung optisch generierter Ladungsträgerpaare im hohen elektrischen Feld der Raumladungszone (RLZ) zwischen p- und n-Leiter. Ladungsträgerpaare können daher nur dann zum Photostrom beitragen, wenn sie entweder im Bereich der Raumladungszone generiert wurden (Driftphotostrom), oder die beteiligten Minoritätsladungsträger aus den umliegenden feldfreien

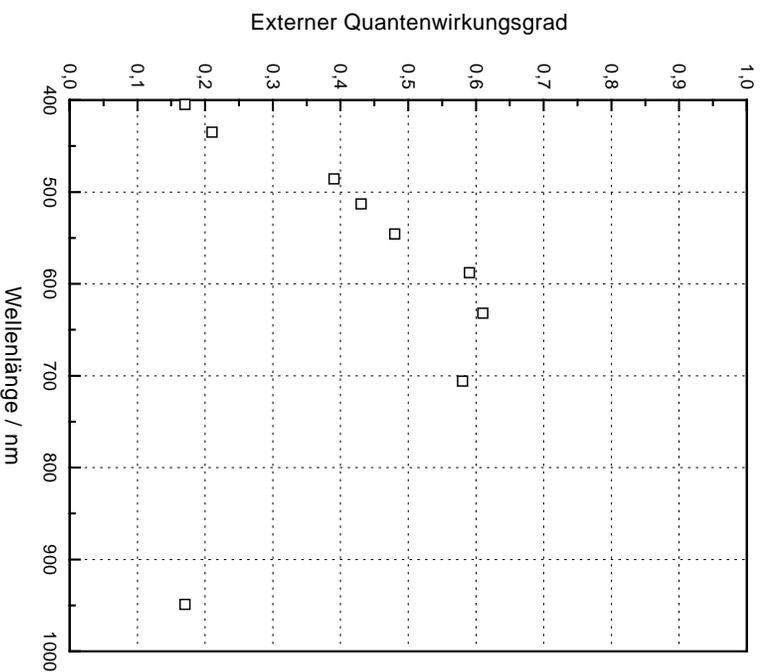


Abbildung 3.8: Gemessener externer Quantenwirkungsgrad η_{ext} einer in Sperrichtung betriebenen n+/p Photodiode.

Gebieten in die Raumladungszone diffundieren (Diffusionsphotostrom). Zur rechnerischen Bestimmung des externen, von außen meßbaren Quantenwirkungsgrades η_{ext} müssen die verschiedenen Drift- und Diffusionsgebiete getrennt betrachtet werden. Der externe Quantenwirkungsgrad ergibt sich dann aus der Summe der für die einzelnen Gebiete ermittelten lokalen Quantenwirkungsgrade. Die Berechnung des externen Quantenwirkungsgrades zur Modellierung des statischen Verhaltens einer vertikalen n-Diffusion/p-Substrat-Photodiode wird im Anhang vorgestellt. Abbildung 3.8 zeigt den gemessenen externen Quantenwirkungsgrad einer solchen Photodiode in Abhängigkeit von der Wellenlänge [23].

Ist die spektrale Abhängigkeit des externen Quantenwirkungsgrades aufgrund von Messungen oder Simulationen bekannt, so läßt sich der in der Diode fließende Photostrom I_{ph} berechnen. Dazu wird zunächst der spektrale Photostrom $I_{\text{ph},\lambda}$ betrachtet, der durch Strahlung in einem Intervall $d\lambda$ um die mittlere Wellenlänge λ generiert wird. Setzt man voraus, daß die einfallende spektrale Bestrahlungsstärke über der optisch wirksamen Diodenfläche konstant ist, so ergibt sich unter Berücksichtigung von Gleichung (3.22)

$$I_{\text{ph},\lambda} = A_{\text{D}} q E_{\text{p},\lambda} \eta_{\text{ext}}(\lambda) = A_{\text{D}} \underbrace{\frac{\lambda q \eta_{\text{ext}}(\lambda)}{h c}}_{R_{\lambda}} E_{\text{e},\lambda}, \quad (3.30)$$

wobei $E_{e,\lambda}$ die einfallende spektrale Bestrahlungsstärke, $E_{p,\lambda}$ die einfallende spektrale Photonenbestrahlungsstärke, A_D die optisch wirksame Diodenfläche und q die Elementarladung bezeichnet. Mit der spektralen Empfindlichkeit R_λ (spectral responsivity), die entsprechend Gleichung (3.30) gegeben ist durch

$$R_\lambda = \frac{dI_{ph,\lambda}}{d\Phi_{e,\lambda}} = \frac{\lambda q \eta_{ext}(\lambda)}{h c}, \quad (3.31)$$

ergibt sich schließlich der Photostrom durch Integration seiner spektralen Dichte über der Wellenlänge zu

$$I_{ph} = A_D \int_{\lambda} R_\lambda E_{e,\lambda} d\lambda. \quad (3.32)$$

Wird eine Photodiode ohne Beleuchtung betrieben, so fließt im äußeren Stromkreis beim Anlegen einer Sperrspannung ein Kurzschlußdunkelstrom, der in Diffusions-, Generations- und Tunneldunkelstrom sowie Oberflächenleckstrom unterteilt werden kann. Die Summe dieser Stromanteile hängt von sehr vielen Parametern ab und ist nur schlecht vorhersagbar. Aus Messungen ergab sich für den Dunkelstrom einer n-Diffusion/p-Substrat Photodiode ein Wert von etwa $20 \text{ aA}/\mu\text{m}^2$.

Der Gesamtstrom durch eine Photodiode in Flußrichtung ist gegeben durch die Beziehung [22]:

$$I_D = I_S \left(e^{\frac{U}{U_T}} - 1 \right) - I_{ph} - I_{dark}. \quad (3.33)$$

Der Gesamtstrom I_D setzt sich hierbei zusammen aus einem Diffusionsstrom $I_S e^{U/U_T}$ in Flußrichtung, einem Feldstrom I_S in Sperrichtung, dem Photostrom I_{ph} in Sperrichtung und einem zusätzlichen Dunkelstromanteil I_{dark} in Sperrichtung. Die einzelnen Stromanteile fließen unabhängig voneinander und sind bei der späteren Behandlung des Rauschens von Photodioden getrennt zu betrachten. Abbildung 3.9 zeigt den qualitativen Verlauf der Kennlinie einer Photodiode mit- und ohne Beleuchtung. Die Kennlinie wird mit zunehmender Beleuchtung in Richtung kleinerer Ströme verschoben.

Für die dynamischen Eigenschaften optischer Sensorsysteme und insbesondere auch für deren Rauscheigenschaften ist die Kapazität des Strahlungsdetektors von großer Bedeutung. Die flächenbezogene differentielle Sperrschichtkapazität einer Photodiode ist gegeben durch

$$C'_D = \sqrt{\frac{\epsilon_{Si} q N_A N_D}{2(N_A + N_D)} \frac{1}{U_D + U_{sperr}}} = \frac{C'_{D,0}}{\sqrt{1 + \frac{U_{sperr}}{U_D}}}, \quad (3.34)$$

wobei U_{sperr} hier die Sperrspannung der Photodiode, U_D die Diffusionsspannung des pn-Übergangs und $C'_{D,0}$ die flächenbezogene Sperrschichtkapazität für $U_{sperr} = 0 \text{ V}$ bezeichnet.

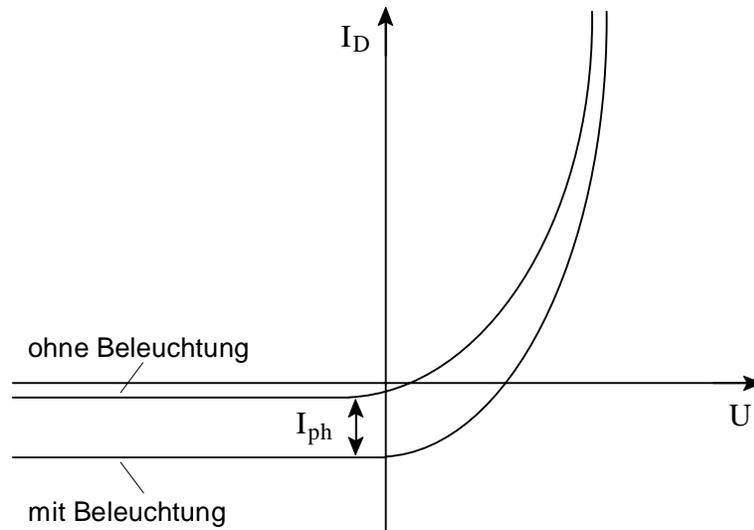


Abbildung 3.9: Kennlinie einer Photodiode.

3.3 Rauschen in CMOS-Photosensoren

Der Begriff Rauschen im hier verwendeten Sinne bezeichnet die in allen elektronischen Schaltungen auftretenden statistischen Fluktuationen von Spannungen und Strömen. Die Existenz von elektronischem Rauschen beruht auf der fundamentalen Tatsache, daß elektrische Ladung nicht räumlich kontinuierlich, sondern gequantelt in Einheiten einer Elementarladung transportiert wird. Bei der Konzeption und der Dimensionierung optischer Sensorsysteme kommt dem im Sensor und seiner Auslese auftretenden Rauschen große Bedeutung zu, da es eine fundamentale untere Grenze für die Meßgenauigkeit und für die Stärke von Signalen, die elektronisch verarbeitet werden können, darstellt.

Im Rahmen dieses Abschnittes sollen die für die Beschreibung von Rauschsignalen im Zeit- und Frequenzbereich erforderlichen Grundlagen, die verschiedenen in CMOS-Photosensoren zu berücksichtigenden Rauschmechanismen und die daraus für die Komponenten optischer Sensorsysteme resultierenden Rauschmodelle vorgestellt und erläutert werden.

3.3.1 Beschreibung von Rauschen im Zeit- bzw. Frequenzbereich

Die dem Rauschen zugrundeliegenden stochastischen Prozesse können nach [27, Seite 7] als ergodisch angenommen werden, so daß zum einen die zu deren Beschreibung benutzten

Schmittmittelwerte² (Erwartungswert, Varianz, Autokorrelationsfunktion) durch entsprechende zeitliche Mittelwerte ausgedrückt werden können und zum anderen die Rauschprozesse als stationär beziehungsweise zeitinvariant angenommen werden können. Der Mittelwert \bar{x} (Erwartungswert) einer Zufallsvariable X ist für einen ergodischen Rauschprozeß gegeben durch

$$\bar{x} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) dt . \quad (3.35)$$

Die Varianz σ^2 kennzeichnet die mittlere quadratische Abweichung der Werte der Probenfunktion $x(t)$ vom Mittelwert \bar{x} . Für einen ergodischen Zufallsprozeß entspricht sie der mittleren Leistung des fluktuierenden Signalanteils $x(t)$ und ist gegeben durch

$$\sigma^2 = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (x(t) - \bar{x})^2 dt . \quad (3.36)$$

Eine für die Charakterisierung von Rauschprozessen sehr wichtige Kenngröße ist die Autokorrelationsfunktion $R(\tau)$. Sie ist ein Maß für die Ähnlichkeit zweier im Zeitabstand τ auftretender Werte der Probenfunktion $x(t)$. Für einen ergodischen Zufallsprozeß kann unter der Voraussetzung $\bar{x} = 0$ die Autokorrelationsfunktion einer Zufallsvariablen X folgendermaßen angegeben werden:

$$R(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t + \tau) x(t) dt . \quad (3.37)$$

Die Autokorrelationsfunktion ist für stationäre Zufallsprozesse eine reelle gerade Funktion in τ und nimmt für $\tau = 0$ ihren Maximalwert an. Dieser Wert $R(0)$ ist für $\bar{x} = 0$ gleich der Varianz σ^2 des Zufallsprozesses. Die mit dem Faktor 2 multiplizierte Fouriertransformierte

$$S_F(f) = \int_{-\infty}^{\infty} R(\tau) e^{-j2\pi f\tau} d\tau \quad (3.38)$$

der Autokorrelationsfunktion $R(\tau)$ bezeichnet für positive Frequenzen f die einseitige spektrale Leistungsdichte $S(f)$ eines stationären Zufallsprozesses:

$$S(f) = \begin{cases} 2 \cdot \int_{-\infty}^{\infty} R(\tau) e^{-j2\pi f\tau} d\tau & \text{falls } f \geq 0 \\ 0 & \text{sonst,} \end{cases} \quad (3.39)$$

In der einschlägigen Fachliteratur wird im Zusammenhang mit Rauschen in elektronischen Schaltungen meist die einseitige spektrale Leistungsdichte verwendet. Der Grund hierfür

²Schmittmittelwerte ergeben sich durch Mittelung eines Ensembles von Zufallswerten, das man erhält, wenn eine parallel gemessene Schar von Probenfunktionen eines Zufallsprozesses zu einem festen Zeitpunkt abgetastet wird.

ist sicherlich, daß die einseitige spektrale Leistungsdichte physikalisch anschaulich für positive Frequenzen die Signalleistung in einem Frequenzintervall von 1 Hz angibt, während die zweiseitige spektrale Leistungsdichte jeweils nur den halben Wert dieser Signalleistung angibt. Im folgenden wird, falls nicht besonders gekennzeichnet, für Rauschberechnungen die einseitige spektrale Leistungsdichte verwendet. Berücksichtigt man, daß die Autokorrelationsfunktion $R(\tau)$ eine reellwertige und in τ gerade Funktion ist und ihre Fouriertransformierte $S_F(f)$ somit in f gerade ist, so ergibt sich für die Rücktransformation der einseitigen spektralen Leistungsdichte $S(f)$ der Zusammenhang:

$$R(\tau) = \int_{-\infty}^{\infty} S_F(f) e^{j2\pi f\tau} df = 2 \cdot \int_0^{\infty} S_F(f) e^{j2\pi f\tau} df = \int_0^{\infty} S(f) e^{j2\pi f\tau} df . \quad (3.40)$$

Für einen durch $\bar{x} = 0$ charakterisierten Rauschprozeß ergibt sich dann für die Varianz beziehungsweise die Signalleistung die wichtige Beziehung:

$$\sigma^2 = R(0) = \int_0^{\infty} S(f) df . \quad (3.41)$$

Ein Rauschprozeß wird als weiß bezeichnet, wenn die zweiseitige spektrale Leistungsdichte $S_F(f)$ über der Frequenz konstant ist. Für diesen Fall sind zwei zeitlich versetzt gemessene Signalwerte $x(t)$ völlig unkorreliert, die Autokorrelationsfunktion ist dementsprechend durch einen Dirac-Stoß gegeben. Weißes Rauschen müßte entsprechend Gleichung (3.41) eine unendlich hohe mittlere Signalleistung aufweisen und kann daher in realen physikalischen Systemen nicht vorkommen. Es dient zur Modellierung von Rauschprozessen, für die die spektrale Leistungsdichte im interessierenden Frequenzbereich nahezu konstant ist.

Sehr wichtig für die Rauschanalyse elektronischer Schaltungen ist die Übertragung stochastischer Signale über lineare zeitinvariante Systeme (LZI-Systeme). Bezeichnet $H(f)$ die Übertragungsfunktion eines LZI-Systems und $S_e(f)$ die spektrale Leistungsdichte am Eingang, so ist die spektrale Leistungsdichte $S_a(f)$ am Ausgang des Systems gegeben durch

$$S_a(f) = |H(f)|^2 S_e(f) . \quad (3.42)$$

Hilfreich für die Analyse der Übertragung von weißem Rauschen über Systeme mit einer Polstelle (Tiefpaß erster Ordnung) ist der Begriff der äquivalenten Rauschbandbreite [8]. Hierunter versteht man diejenige Bandbreite, die ein Tiefpaß mit rechteckförmiger Übertragungsfunktion haben muß, um am Ausgang die gleiche Rauschleistung hervorzurufen, wie das betrachtete System. Bezeichnet f_1 die 3-dB-Grenzfrequenz des Tiefpasses erster Ordnung, so gilt für die äquivalente Rauschbandbreite $f_{\text{äq}}$:

$$f_{\text{äq}} = \frac{\pi}{2} f_1 . \quad (3.43)$$

Für die Übertragung von weißem Rauschen über einen Tiefpaß, dessen Systemfunktion zwei Polstellen f_1 und f_2 aufweist, ergibt sich entsprechend:

$$f_{\text{äq}} = \frac{\pi}{2} \frac{1}{\frac{1}{f_1} + \frac{1}{f_2}} . \quad (3.44)$$

Das Rauschen am Ausgang einer elektronischen Schaltung setzt sich in der Regel aus Beiträgen mehrerer Rauschquellen zusammen, die als statistisch unabhängig angenommen werden können. Die Varianz der Überlagerung der verschiedenen Beiträge ergibt sich für diesen Fall aus der Summe der Varianzen der einzelnen Rauschbeiträge [28]. Es gilt:

$$\sigma_{\Sigma}^2 = \sum_n \sigma_n^2 . \quad (3.45)$$

Entsprechende Zusammenhänge gelten unter der Voraussetzung statistischer Unabhängigkeit auch für die Autokorrelationsfunktion und die spektrale Leistungsdichte eines additiv zusammengesetzten Rauschprozesses.

3.3.2 Schrotrauschen

Rauschen, das entsteht, wenn beispielsweise Regentropfen auf ein Blechdach fallen, Photonen auf einen Detektor treffen, oder Ladungsträger aufgrund ihrer kinetischen Energie die Potentialbarriere eines pn-Überganges überwinden, wird als Schrotrauschen bezeichnet. Charakteristisch für Schrotrauschen ist, daß das resultierende Signal sich aus zufällig über die Zeitachse verteilten gleichartigen Einzelimpulsen mit einer mittleren Rate λ zusammensetzt (mit der mittleren Rate λ wird hier die mittlere Anzahl von Impulsen pro Sekunde bezeichnet). Wichtig ist dabei, daß die Einzelimpulse völlig unabhängig voneinander auftreten, d.h. daß zwischen dem Auftreten zweier Impulse kein kausaler Zusammenhang bestehen darf.

Mißt man die Anzahl N der in einem Zeitintervall T auftretenden Einzelimpulse, so stellt N eine Zufallsvariable dar und ist unter der obigen Voraussetzung, daß die Einzelimpulse unabhängig voneinander auftreten, poissonverteilt [29]. Für den Mittelwert \overline{N} und die Varianz σ_N^2 von N gelten dann die Zusammenhänge:

$$\overline{N} = \lambda T \quad , \quad \sigma_N^2 = \overline{N} = \lambda T . \quad (3.46)$$

Die spektrale Leistungsdichte des Schrotrauschens kann unter der Voraussetzung einer poissonverteilten Impulsrate mit Hilfe des Theorems von Carson bestimmt werden [27,29]. Beschreibt $g(t)$ die Form eines einzelnen Impulses und $G(f)$ die Fourier-Transformierte von $g(t)$, so gilt für die einseitige spektrale Leistungsdichte $S(f)$ der Zusammenhang:

$$S(f) = 2\lambda |G(f)|^2 . \quad (3.47)$$

Betrachtet man nun das Schrotrauschen eines elektrischen Stromes I_0 , wobei zunächst davon ausgegangen wird, daß die einzelnen durch den Transport von Elementarladungen verursachten Stromimpulse gleiche Polarität aufweisen, so werden durch die Funktion $g(t)$ einzelne Stromimpulse der Fläche

$$q = \int_{-\infty}^{\infty} g(t) dt = G(0) \quad (3.48)$$

beschrieben. Mit $\lambda = I_0/q$ und der Normierung $G_n(f) = G(f)/G(0)$ erhält man aus Gleichung (3.47) für die einseitige spektrale Leistungsdichte des Rauschstromes die Beziehung:

$$S_i(f) = 2qI_0 |G_n(f)|^2, \quad (3.49)$$

wobei I_0 den Mittelwert des fließenden Stromes bezeichnet. Für Frequenzen weit unterhalb der reziproken Impulsdauer (beziehungsweise wenn man die Impulsdauer gegen Null gehen läßt) geht Gleichung (3.49) schließlich in die als Schottky-Formel bekannte Beziehung über:

$$S_i(f) \approx 2qI_0 |G_n(0)|^2 = 2qI_0, \quad (3.50)$$

An dieser Stelle ist es wichtig, darauf hinzuweisen, daß die Schottky-Beziehung in dieser Form nicht auf den in einer Halbleiter-Diode fließenden Gesamtstrom angewendet werden kann. Dieser setzt sich aus entgegengesetzt fließenden Diffusions- und Feldströmen zusammen, die unabhängig voneinander gemäß Gleichung (3.50) zum Schrotrauschen beitragen [28]. Die Ströme heben sich in bezug auf den Gleichanteil zwar teilweise auf, die sich überlagernden Stromimpulse tragen aber ungeachtet ihres Vorzeichens zum Rauschen bei, so daß zur Berechnung der durch Schrotrauschen einer Halbleiter-Diode verursachten Gesamt-Rauschleistung die Impulsraten der entgegengesetzt fließenden Ströme addiert werden müssen.

In Halbleiter-Bauelementen sind die Randbedingungen für die Entstehung von Schrotrauschen insbesondere dann gegeben, wenn der Stromfluß durch die Höhe einer Potentialbarriere gesteuert wird (Diode, Bipolartransistor, MOS-Transistor in schwacher Inversion). Aufgrund der sehr hohen elektrischen Feldstärke in der Umgebung einer Potentialbarriere bewegen sich die Ladungsträger praktisch unabhängig voneinander, da die gegenseitige Beeinflussung hier gegenüber den im Feld wirkenden Kräften vernachlässigt werden kann.

3.3.3 Photonenrauschen

Der Begriff Photonenrauschen benennt hier den im Strahlungsdetektor aufgrund der stochastischen Natur der optischen Generation von Ladungsträgerpaaren hervorgerufenen Rauschbeitrag. Photonenrauschen stellt eine fundamentale Einschränkung für die in Strahlungsdetektoren erzielbaren Rauscheigenschaften dar, da es durch die detektierte Größe selbst in das System eingebracht wird. Detektoren, die in der Weise optimiert sind, daß ihr Eigenrauschen kleiner ist als der am Ausgang durch das Photonenrauschen hervorgerufene Rauschbeitrag, werden als „photon-noise-limited“ bezeichnet.

Der Detektor (z.B. eine Photodiode) soll für die folgenden Betrachtungen als ideal angenommen werden, in dem Sinne, daß er in Abwesenheit einer einfallenden Strahlung keinen Ausgangsstrom liefert. Die durch die optische Generation von Ladungsträgerpaaren gegebenen Ereignisse können als unabhängig voneinander und zufällig (mit einer mittleren Rate λ) über die Zeitachse verteilt angenommen werden [30]. Der im Detektor erzeugte Signalstrom weist daher entsprechend der Darstellung im vorigen Abschnitt Schrotrauschen auf, und die Anzahl der in einem Zeitintervall generierten Ladungsträgerpaare ist poissonverteilt.

Im folgenden soll der limitierende Einfluß des Photonenrauschens für den Fall integrierender Photosensoren untersucht werden. Bei einem integrierenden Photosensor im hier verstandenen Sinne wird auf der Sperrschichtkapazität einer in Sperrichtung betriebenen Photodiode deren Photostrom I_{ph} während einer Integrationszeit T_{int} aufintegriert.

Bezeichnet die Zufallsvariable N die Anzahl der während eines speziellen Integrationsintervalles der Länge T_{int} auf der Sperrschichtkapazität akkumulierten Photoelektronen, so ist der Mittelwert \overline{N} von N gegeben durch:

$$\overline{N} = \frac{I_{\text{ph}} T_{\text{int}}}{q} . \quad (3.51)$$

Entsprechend Gleichung (3.46) ist unter der Voraussetzung, daß N poissonverteilt ist, die Varianz σ_N^2 von N gegeben durch:

$$\sigma_N^2 = \overline{N} = \frac{I_{\text{ph}} T_{\text{int}}}{q} . \quad (3.52)$$

Diese beim „Zählen von Photonen“ bestehende Unsicherheit stellt eine fundamentale Einschränkung für den in Photodetektoren erzielbaren Signal- zu Rauschabstand (SNR) dar, der gegeben ist durch

$$\text{SNR}_{\text{dB,max}} = 10 \text{ dB} \log \frac{\overline{N}^2}{\sigma_N^2} = 10 \text{ dB} \log \overline{N} . \quad (3.53)$$

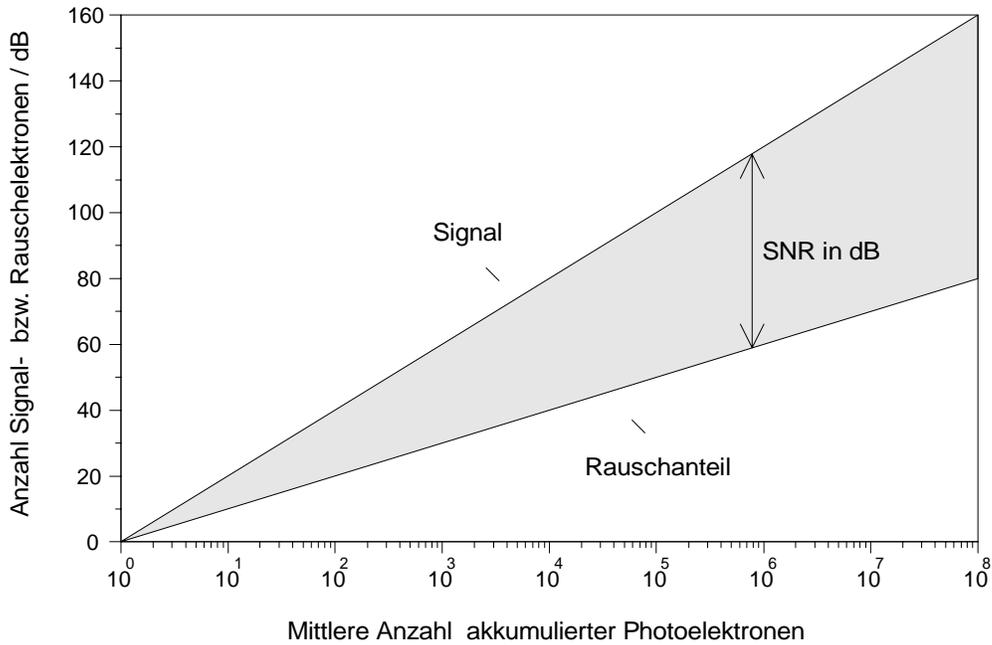


Abbildung 3.10: Durch Photonenrauschen begrenzter SNR für den Fall integrierender Photodetektoren. Die obere Gerade zeigt die mittlere Anzahl akkumulierter Photoelektronen \bar{N} in dB, gegeben durch $20\text{dB} \cdot \log(\bar{N})$. Die untere Gerade zeigt die Standardabweichung σ_N von N in dB, gegeben durch $20\text{dB} \cdot \log(\sigma_N)$.

Der erzielbare SNR steigt mit der mittleren Anzahl akkumulierter Photoelektronen um 10dB/Dekade. Die Zusammenhänge zwischen der Anzahl von Signalelektronen³, der Anzahl von Rauschelektronen⁴ und dem daraus resultierenden SNR werden in Abbildung 3.10 nochmals graphisch verdeutlicht.

Zur Realisierung hochgenauer Strahlungsdetektoren, wie sie zum Beispiel im Bereich der Medizintechnik in Computertomographen benötigt werden, ist es nicht sinnvoll, die Empfindlichkeit wie in herkömmlichen Kameras über die Belichtungszeit einzustellen, da auf diese Weise der erzielbare SNR unnötig begrenzt wird (eine Halbierung der Belichtungszeit reduziert den erzielbaren SNR um 3 dB). Die Belichtungszeit sollte dagegen in solchen Systemen unter den gegebenen Randbedingungen so groß wie möglich gewählt, und die Empfindlichkeit über die Verstärkung der Ausleseschaltung eingestellt werden.

Das in Gleichung (3.52) angegebene Ergebnis für die durch Photonenrauschen verursachte mittlere quadratische Anzahl von Rauschelektronen läßt sich mit Hilfe systemtheo-

³Unter der Anzahl von Signalelektronen wird hier die als Meßwert interessierende mittlere Anzahl akkumulierter Photoelektronen verstanden.

⁴Unter der Anzahl von Rauschelektronen wird hier die Standardabweichung der Anzahl akkumulierter Photoelektronen verstanden.

retischer Überlegungen auch direkt aus der spektralen Leistungsdichte des Rauschstromes einer Photodiode herleiten. Dabei wird sich zeigen, daß der Integration des Photostromes I_{ph} während der Belichtungszeit T_{int} eine Begrenzung der Rauschbandbreite auf die Frequenz $1/(2T_{\text{int}})$ entspricht. Dieses Ergebnis ist wichtig für den späteren Vergleich von integrierenden und nichtintegrierenden Photosensoren.

Die Integration des zeitabhängigen Photostromes läßt sich mathematisch durch eine Faltung des Photostromes mit einem Zeitfenster der Form $f(t) = \text{rect}[(t - T_{\text{int}}/2)/T_{\text{int}}]$ und eine anschließende Abtastung des Ergebnisses der Faltungsoperation beschreiben. Das Ergebnis der Faltungsoperation gibt für jeden Zeitpunkt t_0 die während der zurückliegenden Zeitspanne $t_0 - T_{\text{int}} < t < t_0$ akkumulierte Ladung an. Ist die spektrale Leistungsdichte $S_q(f)$ dieser zeitabhängigen Ladungsfunktion bekannt, so kann die beim Abtasten dieser Funktion auftretende mittlere quadratische Rauschladung (beziehungsweise die Varianz der akkumulierten Ladungsmenge) entsprechend Gleichung (3.41) durch Integration von $S_q(f)$ über der Frequenz berechnet werden. Die spektrale Leistungsdichte $S_q(f)$ ergibt sich aus den Gleichungen (3.50) und (3.42) zu

$$S_q(f) = 2 q I_{\text{ph}} T_{\text{int}}^2 \frac{\sin^2(\pi T_{\text{int}} f)}{(\pi T_{\text{int}} f)^2} . \quad (3.54)$$

Für I_0 in Gleichung (3.50) wurde dabei I_{ph} und für $H(f)$ in Gleichung (3.42) die Fouriertransformierte der Fensterfunktion $f(t)$ eingesetzt. Für die beim Abtastvorgang auftretende mittlere quadratische Rauschladung ergibt sich damit

$$\overline{Q_n^2} = 2 q I_{\text{ph}} T_{\text{int}}^2 \underbrace{\int_0^{\infty} \frac{\sin^2(\pi T_{\text{int}} f)}{(\pi T_{\text{int}} f)^2} df}_{1/(2T_{\text{int}})} = q I_{\text{ph}} T_{\text{int}} . \quad (3.55)$$

Dieses Ergebnis entspricht der Gleichung (3.52), wenn der Zusammenhang $\overline{Q_n^2} = q^2 \sigma_N^2$ berücksichtigt wird. Aus Gleichung (3.55) folgt, daß der Integration eine Bandbegrenzung des Rauschens auf die Frequenz $1/(2T_{\text{int}})$ entspricht.

3.3.4 Thermisches Rauschen

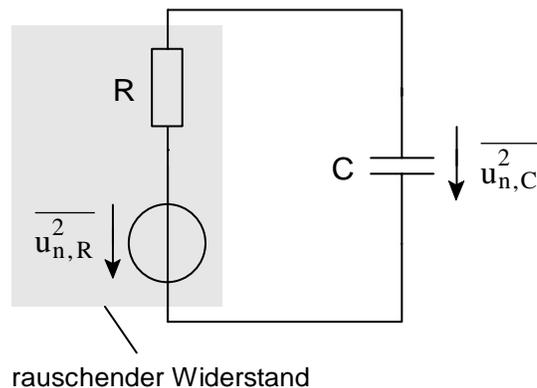


Abbildung 3.11: Zur Übertragung weißen Rauschens über einen RC-Tiefpaß. $\overline{u_{n,R}^2}$ bezeichnet die Signalleistung der Rauschspannung am Eingang des RC-Tiefpasses, $\overline{u_{n,C}^2}$ bezeichnet die Signalleistung der Rauschspannung am Ausgang des RC-Tiefpasses.

Thermisches Rauschen (Widerstandsrauschen), zum Beispiel in Form einer Rauschspannung $u_n(t)$, beruht auf der in jedem Leiter oberhalb $T = 0\text{ K}$ stattfindenden thermischen Bewegung der Ladungsträger, die zu stochastischen Fluktuationen der Ladungsverteilung im Leiter und damit zu Spannungsfluktuationen an dessen Klemmen führen. Thermisches Rauschen an einem Widerstand tritt im Gegensatz zu Schrotrauschen unabhängig davon auf, ob dieser Widerstand von einem Strom durchflossen wird oder nicht.

Unter der Annahme, daß es sich beim thermischen Rauschen eines Widerstandes um einen weißen Rauschprozeß handelt, was experimentell nachgewiesen werden kann, läßt sich die spektrale Leistungsdichte des thermischen Rauschens aus der Betrachtung der Parallelschaltung eines rauschenden Widerstandes und eines Kondensators gemäß Abbildung 3.11 mit Hilfe des Gleichverteilungssatzes herleiten. Der rauschende Widerstand ist hier im Ersatzschaltbild als die Reihenschaltung eines idealen Widerstandes und einer Rauschspannungsquelle dargestellt. Ordnet man der Rauschspannungsquelle am Eingang des dargestellten RC-Tiefpasses die spektrale Leistungsdichte $S_{u,R}(f)$ zu, so ergibt sich am Ausgang des RC-Tiefpasses für die spektrale Leistungsdichte $S_{u,C}(f)$ der Rauschspannung am Kondensator unter anderem mit Hilfe von Gleichung (3.42) die Beziehung:

$$S_{u,C}(f) = S_{u,R}(f) \left| \frac{1}{1 + j2\pi fRC} \right|^2. \quad (3.56)$$

Die äquivalente Rauschbandbreite nach Gleichung (3.43) ergibt sich mit $f_{3\text{dB}} = 1/(2\pi RC)$ zu $f_{\text{äq}} = 1/(4RC)$. Damit erhält man unter der Voraussetzung weißen Rauschens für die

Signalleistung der Rauschspannung am Kondensator:

$$\overline{u_{n,c}^2} = S_{u,R} \frac{1}{4RC} . \quad (3.57)$$

Andererseits muß aufgrund des Gleichverteilungssatzes der statistischen Physik [29], wonach ein System im thermischen Gleichgewicht pro Freiheitsgrad⁵ die mittlere thermische Energie $kT/2$ annimmt, gelten:

$$\frac{1}{2} C \overline{u_{n,c}^2} = \frac{1}{2} kT \quad \text{bzw.} \quad \overline{u_{n,c}^2} = \frac{kT}{C} . \quad (3.58)$$

Die Kombination der letzten beiden Gleichungen ergibt für die spektrale Leistungsdichte der Rauschspannung am Widerstand schließlich den als Nyquist-Beziehung bekannten Ausdruck

$$S_{u,R} = 4 kT R . \quad (3.59)$$

Wird dem rauschenden Widerstand eine äquivalente Rauschstromquelle zugeordnet, so beträgt deren spektrale Leistungsdichte

$$S_{i,R} = 4 kT \frac{1}{R} . \quad (3.60)$$

In einer allgemeineren Formulierung der Nyquist-Beziehung für passive lineare Zweipole mit der Eingangsimpedanz $\underline{Z}(\omega)$ gilt

$$S_{u,Z}(f) = 4 kT \operatorname{Re}\{\underline{Z}(2\pi f)\} . \quad (3.61)$$

3.3.5 1/f-Rauschen

Zusätzlich zum stets vorhandenen thermischen Rauschen findet sich in allen elektronischen Bauelementen eine vom fließenden Strom abhängige Rauschkomponente, deren spektrale Leistungsdichte mit $|f|^{-\alpha}$ variiert – wobei α in der Regel Werte nahe bei Eins annimmt. Dieser Rauschprozeß wird aufgrund seiner spektralen Charakteristik als 1/f-Rauschen bezeichnet. Obwohl es seit seiner ersten Beobachtung (Johnson 1925) Gegenstand intensiver Forschung gewesen ist, existiert zur Zeit noch keine einheitliche, allgemein anerkannte physikalische Theorie des 1/f-Rauschens, die in der Lage wäre, die beobachteten Rauschphänomene in allen Fällen befriedigend zu erklären [29]. Das in MOS-Transistoren beobachtete 1/f-Rauschen ist nach Ansicht der meisten Autoren [27,29] durch einen Oberflächeneffekt an der Grenzfläche zwischen Kanal und Gate-Oxid erklärbar. Die dort aufgrund von Verunreinigungen und Kristalldefekten vorhandenen Oberflächenzustände

⁵Die Anzahl der Freiheitsgrade eines Systems entspricht der Anzahl unabhängiger Variablen, die in die Gesamtenergie des Systems quadratisch eingehen. Ein Kondensator besitzt demnach einen Freiheitsgrad.

können Elektronen aus dem Kanal einfangen (bzw. wieder emittieren) und auf diese Weise Rauschen verursachen. Das $1/f$ -Rauschen von MOS-Transistoren läßt sich durch eine eingangsbezogene spektrale Leistungsdichte der folgenden allgemeinen Form beschreiben:

$$S_u(f) = \frac{K}{f^\alpha} \quad (3.62)$$

Bei der Integration einer bei $\alpha = 1$ durch $S_u(f) = K/f$ gegebenen spektralen Leistungsdichte über dem Intervall $[f_1, f_2]$ ergibt sich für die in diesem Frequenzbereich enthaltene Rauschleistung der Ausdruck

$$\overline{u_n^2} = K \ln \frac{f_2}{f_1} = K \ln 10 \log \frac{f_2}{f_1} . \quad (3.63)$$

Hieraus folgt unmittelbar die wichtige, für das $1/f$ -Rauschen charakteristische Eigenschaft, daß die in einer Dekade (bzw. einem beliebigen Frequenzintervall mit festem Verhältnis von oberer zu unterer Grenzfrequenz) enthaltene Rauschleistung konstant ist. Für $f_2 \rightarrow \infty$ als auch für $f_1 \rightarrow 0$ geht die durch Gleichung (3.63) gegebene Rauschleistung gegen unendlich. Während sich für hohe Frequenzen keine praktischen Probleme bei der Berechnung ergeben, da das reale rauschende System aufgrund parasitärer Effekte eine Tiefpaßcharakteristik aufweisen wird, ist unklar, ob eine untere Grenzfrequenz existiert, unterhalb der die $1/f$ -Charakteristik in einen geraden Verlauf übergeht, und ab der die Integration sinnvoll durchzuführen ist. Eine solche untere Grenzfrequenz konnte bisher durch Messungen nicht gefunden werden [29].

In [31] wird vorgeschlagen, als untere Grenzfrequenz die reziproke Betriebsdauer des rauschenden Systems zu wählen, da viel kleinere Frequenzen aufgrund der endlichen Beobachtungsdauer und der damit verbundenen Unschärfe im Frequenzbereich nicht meßbar sind (Einer Betriebsdauer von einem Tag würde eine untere Grenzfrequenz von 10^{-5} Hz entsprechen). Desweiteren tritt die Problematik nur auf, wenn der Exponent α exakt den Wert Eins annimmt. Für jeden Wert von α , für den gilt $0 < \alpha < 1$, nimmt die Rauschleistung für $f_1 \rightarrow 0$ dagegen endliche Werte an.

Zur vereinfachten rechnerischen Analyse der Übertragung von weißem Rauschen über beliebige Tiefpaß-Systeme erster oder zweiter Ordnung wurde die äquivalente Rauschbandbreite $f_{\text{äq}}$ eingeführt. Entsprechend dazu kann für Berechnungen, die die Übertragung von $1/f$ -Rauschen über solche Systeme betreffen, nach [32] eine äquivalente Anzahl von Dekaden ND_e eingeführt werden. Für

$$S_u(f) = \frac{K}{f} \quad \text{und} \quad |H(f)|^2 = \frac{1}{1 + \left(\frac{f}{f_1}\right)^2} \quad (3.64)$$

ergibt sich am Ausgang des Systems $H(f)$ mit der Eckfrequenz f_1 durch Integration in den Grenzen $[f_0, \infty]$ die Signalleistung

$$\overline{u_n^2} = K \frac{1}{2} \ln \left(1 + \frac{f_1^2}{f_0^2} \right) \approx K \ln \frac{f_1}{f_0} = K \ln 10 \log \frac{f_1}{f_0}, \quad (3.65)$$

wobei angenommen wurde, daß $f_1 \gg f_0$. Ein Vergleich mit Gleichung (3.63) ergibt für die äquivalente Anzahl von Dekaden:

$$\text{ND}_e = \log \frac{f_1}{f_0}. \quad (3.66)$$

Für die Übertragung von $1/f$ -Rauschen über ein System mit zwei Polstellen erhält man mit $f_1, f_2 \gg f_0$:

$$\text{ND}_e = \frac{f_2^2 \log \frac{f_1}{f_0} - f_1^2 \log \frac{f_2}{f_0}}{f_2^2 - f_1^2}. \quad (3.67)$$

Für $f_2 \gg f_1$ kann diese Beziehung durch den einfacheren Ausdruck in Gleichung (3.66) angenähert werden.

Beim Entwurf rauscharmer analoger MOS-Schaltungen ist die Berücksichtigung des $1/f$ -Rauschens wichtig, da es (insbesondere für kleine Transistoren) bis in den MHz-Bereich hinein gegenüber dem thermischen Rauschen überwiegen kann.

3.3.6 Rauschmodell für die Photodiode

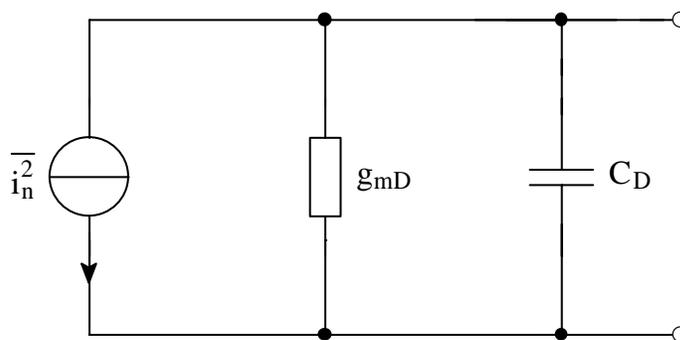


Abbildung 3.12: Rauschersatzschaltbild einer in Sperrichtung betriebenen Photodiode.

Abbildung 3.12 zeigt das Rauschersatzschaltbild einer Photodiode mit der für die Analyse des Rauschens optischer Sensorsysteme wichtigen Detektorkapazität C_D (diese ist durch

Gleichung (3.34) gegeben) und dem rauschfreien differentiellen Ausgangsleitwert g_{mD} , der mit Hilfe von Gleichung (3.33) berechnet werden kann zu

$$g_{mD} = \frac{\partial I_D}{\partial U} = \frac{I_s}{U_T} e^{\frac{U}{U_T}} . \quad (3.68)$$

Berücksichtigt man, daß das thermische Widerstandsrauschen einer Photodiode entsprechend [22] gegenüber ihrem Schrotrauschen vernachlässigt werden kann, so ergibt sich die spektrale Leistungsdichte des Rauschstromes einer in Sperrichtung betriebenen Photodiode aus Gleichung (3.33) und der Schottky-Beziehung (3.50) zu

$$S_I(f) = 2q \left[I_s \left(1 + e^{\frac{U}{U_T}} \right) + I_{ph} + I_{dark} \right] , \quad (3.69)$$

wobei in I_{dark} alle relevanten Dunkelstromanteile zusammengefaßt wurden. Die verschiedenen Dunkelstromanteile und der Photostrom tragen unabhängig voneinander zum Gesamtrauschen des Detektors bei. Für $-U \gg U_T$ kann der Exponentialterm in Gleichung (3.69) vernachlässigt werden und es gilt

$$S_I(f) \approx 2q(I_s + I_{ph} + I_{dark}) . \quad (3.70)$$

Für $U = 0$ V nimmt die spektrale Leistungsdichte des Rauschstromes ihren Maximalwert an. Aus den Gleichungen (3.68) und (3.69) ergibt sich dazu:

$$S_I(f) = 4kTg_{mD} + 2q(I_{ph} + I_{dark}) . \quad (3.71)$$

Der Term $4kTg_{mD}$ kann hier als thermisches Rauschen des differentiellen Ausgangsleitwertes der Photodiode interpretiert werden.

3.3.7 Rauschmodell für den MOS-Feldeffekttransistor

Das Rauschen eines MOS-Transistors in starker Inversion setzt sich aus einem thermischen Anteil, der durch die ungeordnete thermische Bewegung freier Elektronen im Kanalbereich hervorgerufen wird, und 1/f-Rauschen zusammen. Bei niedrigen Frequenzen überwiegt normalerweise der 1/f-Anteil. Im Ersatzschaltbild eines rauschenden MOS-Transistors (Abbildung 3.13) können die verschiedenen Rauschbeiträge einerseits eingangsbezogen entsprechend Fall (a) in einer Rauschspannungsquelle in Reihe zum Eingang – oder entsprechend Fall (b) in einer Rauschstromquelle zwischen Drain und Source eines rauschfreien MOS-Transistors zusammengefaßt werden, wobei die spektralen Leistungsdichten verschiedener unkorrelierter Rauschbeiträge zu addieren sind. Die spektrale Leistungsdichte $S_{i,a}(f)$ des Rauschstromes am Ausgang des Transistors ist mit der spektralen Leistungsdichte $S_{u,e}(f)$ der eingangsbezogenen Rauschspannungsquelle über die Beziehung

$$S_{i,a}(f) = g_m^2 S_{u,e}(f) \quad (3.72)$$

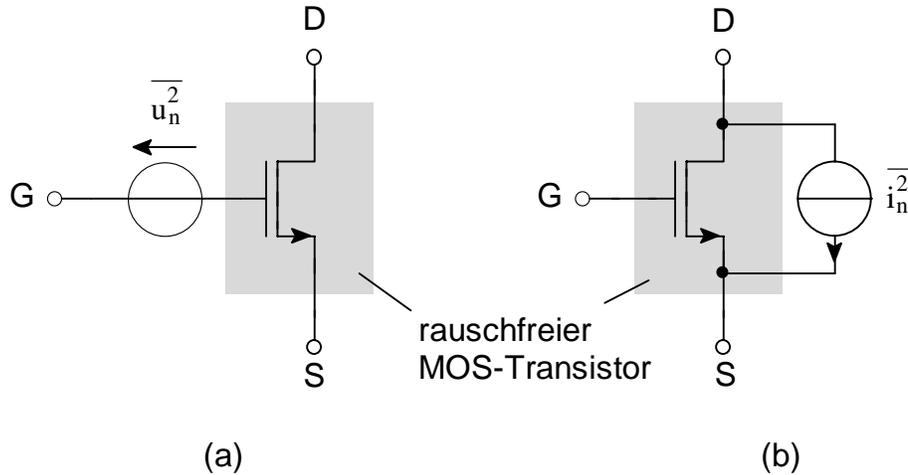


Abbildung 3.13: Vereinfachtes Kleinsignal-Rauschersatzschaltbild des MOS-Transistors. Die eingezeichneten MOS-Transistoren sind hier jeweils durch ihr Kleinsignal-Ersatzschaltbild zu ersetzen.

verknüpft. Bei der rechnerischen Analyse des Rauschens elektronischer Schaltungen ist meistens die Verwendung eingangsbezogener Rauschquellen von Vorteil. Bezeichnet man mit

$$g_{\max} = \mu C'_{\text{ox}} \frac{W}{L} (U_{\text{gs}} - U_{\text{t}}) \quad (3.73)$$

die Transkonduktanz eines MOS-Transistors in Sättigung ($U_{\text{ds}} \geq U_{\text{gs}} - U_{\text{t}}$), so ist entsprechend [27] die spektrale Leistungsdichte einer Ersatz-Rauschstromquelle, die das thermische Kanalrauschen repräsentiert, gegeben durch:

$$S_{i,a}(f) = 4kT \left(\frac{2}{3} g_{\max} \right) \quad (3.74)$$

Für $U_{\text{ds}} \ll U_{\text{gs}} - U_{\text{t}}$ (dieses ist der Fall, wenn der MOS-Transistor z.B. als Schalter betrieben wird) ist der Kanal des MOS-Transistors vollständig invertiert und der Transistor verhält sich ausgangsseitig zwischen Drain und Source wie ein ohmscher Leitwert

$$G_{\text{on}} = \mu C'_{\text{ox}} \frac{W}{L} (U_{\text{gs}} - U_{\text{t}}) = g_{\max} . \quad (3.75)$$

In diesem Arbeitsbereich gilt dann mit der Nyquist-Beziehung (3.60) für die spektrale Leistungsdichte des ausgangsseitigen Rauschstromes:

$$S_{i,a}(f) = 4kT g_{\max} . \quad (3.76)$$

Im Gegensatz zum thermischen Rauschen und zum Schrotrauschen tritt 1/f-Rauschen im MOS-Transistor unabhängig vom jeweiligen Arbeitsbereich und vom jeweils vorherrschenden Transportmechanismus (Drift oder Diffusion) auf [33]. Die eingangsbezogene

spektrale Leistungsdichte für das $1/f$ -Rauschen ist nach [8] gegeben durch

$$S_{u,e}(f) = \frac{K_f}{C'_{ox}WLf} . \quad (3.77)$$

Ein Vergleich mit Gleichung (3.62) zeigt, daß für die dort verwendeten Parameter K und α in obiger Modellgleichung $K = K_f/(C'_{ox}WL)$ beziehungsweise $\alpha = 1$ eingesetzt wurde. Der Rauschparameter K_f ist nur schlecht vorhersagbar und muß für die jeweilige Technologie und den jeweiligen Transistor-Typ durch Messungen ermittelt werden. Die so gewonnenen Werte unterliegen jedoch selbst innerhalb eines Wafers noch starken Schwankungen. Ein typischer Wert für einen NMOS-Transistor ist $5 \cdot 10^{-24}$ VAs.

Für den MOS-Transistor in starker Inversion und in Sättigung ergibt sich damit die folgende eingangsbezogene spektrale Leistungsdichte:

$$S_{u,e}(f) = 4kT \frac{2}{3} \frac{1}{g_m} + \frac{K_f}{C'_{ox}WLf} . \quad (3.78)$$

In schwacher Inversion wird das Rauschen des MOS-Transistors durch Schrotrauschen und $1/f$ -Rauschen bestimmt. Für den $1/f$ -Rauschanteil gilt die Gleichung (3.77) auch in diesem Arbeitsbereich. Der Drainstrom in schwacher Inversion kann nach Gleichung (A.17) im Anhang mit Hilfe der Beziehung

$$I_d = I_{\max} e^{\frac{U_{gs}-U_i}{nU_T}} \left(1 - e^{-\frac{U_{ds}}{U_T}}\right) = I_{\text{sat}} \left(1 - e^{-\frac{U_{ds}}{U_T}}\right) \quad (3.79)$$

modelliert werden. I_d setzt sich physikalisch zusammen aus zwei entgegengesetzt fließenden Diffusionsströmen

$$I_f = I_{\text{sat}} \quad \text{und} \quad I_r = I_{\text{sat}} e^{-\frac{U_{ds}}{U_T}} , \quad (3.80)$$

so daß $I_d = I_f - I_r$ [34]. I_{sat} bezeichnet hier den in Sättigung ($U_{ds} \gg U_T$) fließenden Drainstrom. Dementsprechend ergibt sich für das Schrotrauschen eines MOS-Transistors in schwacher Inversion mit der Schottky-Beziehung (3.50) die spektrale Leistungsdichte

$$S_{i,a}(f) = 2q(I_f + I_r) = 2qI_{\text{sat}} \left(1 + e^{-\frac{U_{ds}}{U_T}}\right) . \quad (3.81)$$

Für $U_{ds} = 0V$ kann das Schrotrauschen eines MOS-Transistors oder auch einer Diode als thermisches Rauschen des differentiellen Ausgangsleitwertes g_{ds} interpretiert werden. Mit

$$g_{ds} = \frac{I_{\text{sat}}}{U_T} e^{-\frac{U_{ds}}{U_T}} \quad (3.82)$$

gilt dann nämlich

$$S_{i,a}(f) = 4qI_{\text{sat}} = 4kTg_{ds} . \quad (3.83)$$

3.3.8 Rauschmodell für den Operationsverstärker

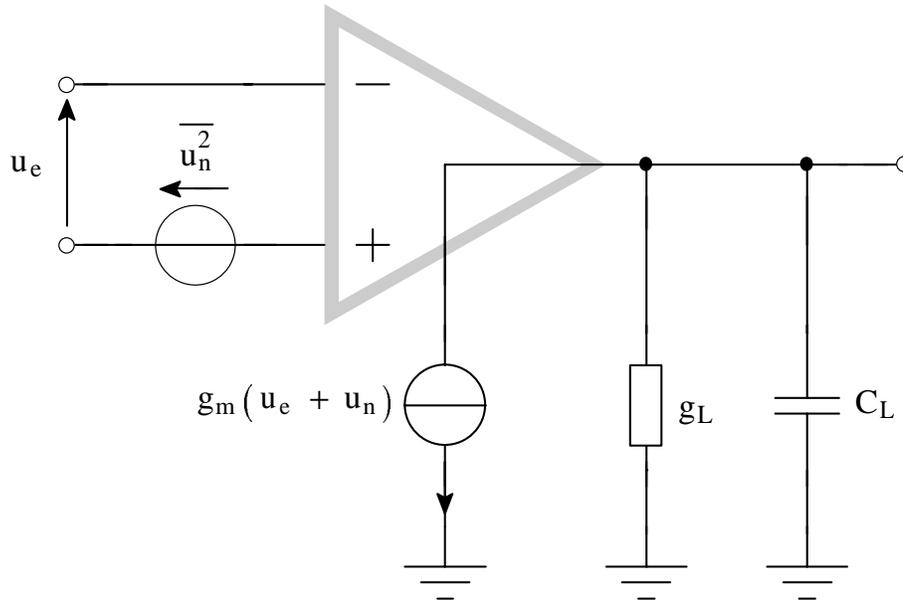


Abbildung 3.14: Kleinsignal-Ersatzschaltbild eines rauschenden Operationsverstärkers. g_m bezeichnet die Transkonduktanz der Eingangstransistoren, g_L den differentiellen Ausgangsleitwert und C_L die kapazitive Last am Ausgang des Operationsverstärkers.

Die Verstärkung und damit auch das Rauschverhalten eines bezüglich seiner Rauscheigenschaften gut entworfenen Operationsverstärkers wird vor allem durch die Dimensionierung der Eingangstransistoren der Differenzstufe bestimmt [8, 35]. Diese sollten einerseits entsprechend Gleichung (3.77) eine möglichst große Gatefläche besitzen, um den $1/f$ -Rauschanteil zu minimieren, und andererseits eine wesentlich höhere Transkonduktanz als alle anderen Transistoren im Signalpfad aufweisen, um die Übertragung verstärkereigenen Rauschens auf den Ausgang gering zu halten.

Da beide Eingangstransistoren der Differenzstufe gleichermaßen zum Gesamttrauschen des Verstärkers beitragen, genügt es, im eingangsbezogenen Rausch-Ersatzschaltbild entsprechend Abbildung 3.14 einen Eingang mit einer in Reihe geschalteten Rauschspannungsquelle zu versehen. Die spektrale Leistungsdichte dieser Rauschspannungsquelle ist dann entsprechend Gleichung (3.78) gegeben durch

$$S_{u,e}(f) = 2 \cdot \left(4kT \frac{2}{3} \frac{1}{g_m} + \frac{K_f}{C'_{ox} WL f} \right), \quad (3.84)$$

wobei g_m die Transkonduktanz eines der Eingangstransistoren- und WL dessen Gatefläche bezeichnet.