

SPAD-basierte Sensoren für die laufzeitbasierte Distanzmessung bei hoher Hintergrundlichtintensität

Von der Fakultät für Ingenieurwissenschaften
Abteilung Elektrotechnik und Informationstechnik
der Universität Duisburg-Essen

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaft

genehmigte Dissertation

von

Maik Beer

aus

Bad Friedrichshall

Gutachter: Prof. Dr.-Ing. Rainer Kokozinski

Gutachter: Prof. Dr. rer. nat. Bernd Jähne

Tag der mündlichen Prüfung: 15.11.2018

Zusammenfassung

Die dreidimensionale Erfassung der Umgebung bietet vielfältige Möglichkeiten in vielerlei Anwendungen. Im Automobilbereich wird sie heute in Fahrerassistenzsystemen wie Notbremssystemen zur Erhöhung der Sicherheit eingesetzt. Auch für eines der großen Zukunftsthemen der Branche, dem autonomen Fahren, spielt sie eine entscheidende Rolle. In den letzten Jahren erlangte die Nutzung elektromagnetischer Wellen im sichtbaren oder infraroten Bereich für die berührungslose Distanzmessung, das sogenannte Light Detection and Ranging, zunehmende Bedeutung. Dieses Verfahren basiert auf der Messung der Zeit, die ein ausgesendetes Lichtsignal zum Zielobjekt und wieder zurück benötigt. Zur Minimierung der emittierten Lichtleistung sind hochempfindliche Photodetektoren nötig, weshalb in der vorliegenden Arbeit auf Single-Photon Avalanche Dioden basierende Sensoren betrachtet werden. Der Einsatz dieser Systeme in einer Umgebung mit hohem Hintergrundlicht stellt eine der größten Herausforderungen für den Einsatz im Automobil dar.

Im Rahmen der Arbeit wurde ein Laufzeitsensor für Anwendungen im Automobil entwickelt, wobei die Reduktion des negativen Einflusses hoher Hintergrundstrahlung im Vordergrund stand. Hierfür wurde zunächst die Statistik der bekannten Verfahren bei Verwendung von Single-Photon Avalanche Dioden betrachtet und Modelle zur Bestimmung der Messvarianz entwickelt. Auf Grundlage dieser Modelle wurden im weiteren Konzept zur Reduktion des Hintergrundlichteinflusses auf die Distanzmessung entworfen. Jedes Konzept wurde theoretisch unter Zuhilfenahme von statistischen Berechnungen sowie mittels Simulationen untersucht und bewertet. Die Betrachtung der verschiedenen Konzepte ließ den Schluss zu, dass eine direkte Laufzeitmessung in Verbindung mit der Erfassung zeitlich korrelierter Photonen das beste Ergebnis verspricht. Basierend auf diesen Erkenntnissen wurde ein entsprechender Doppelzeilen-sensor mit 192 Pixeln je Zeile, wobei jedes 4 Dioden enthält, in einem 0,35- μm -CMOS-Prozess entwickelt und charakterisiert. Die Messergebnisse zeigen eine hohe Übereinstimmung mit den theoretischen Berechnungen und bestätigen die Machbarkeit einer Messung bei hohem Hintergrundlicht.

Inhaltsverzeichnis

Abkürzungsverzeichnis.....	ix
Formelzeichen	xi
1 Einleitung.....	1
1.1 Stand der Technik.....	2
1.2 Ziel und Aufbau der Arbeit.....	6
2 Single-Photon Avalanche Diode	9
2.1 Einführung	9
2.2 Aufbau.....	11
2.3 Funktionsweise	14
2.4 Kenngrößen	16
2.4.1 Photonendetektionseffizienz.....	17
2.4.2 Rauschen	24
2.4.3 Zeitlicher Jitter	29
2.5 Quenching und Reset	32
2.5.1 Passives Quenching und Reset.....	32
2.5.2 Aktives Quenching und Reset.....	34
2.5.3 Passives Quenching und aktiver Reset.....	35
2.5.4 Time-Gating.....	36
3 Verfahren der optischen Entfernungsmessung.....	37
3.1 Interferometrie.....	37
3.2 Triangulation.....	39
3.3 Laufzeitmessung.....	40
3.3.1 Direkte Laufzeitmessung	40
3.3.2 Indirekte Laufzeitmessung.....	41
4 Statistik laufzeitbasierter Messverfahren.....	45
4.1 Photonenstatistik	45
4.1.1 Kohärentes Licht.....	46
4.1.2 Thermisches Licht.....	49
4.1.3 Chaotisches Licht	51
4.2 Indirekte Laufzeitmessung	51
4.2.1 Statistik.....	52

4.2.2	Varianz.....	57
4.2.3	Vergleich der Modulationsarten.....	69
4.3	Direkte Laufzeitmessung.....	72
4.3.1	Statistik.....	72
4.3.2	Hintergrundstrahlung.....	74
4.3.3	Histogramm.....	75
4.3.4	Varianz.....	80
5	Modellbildung und Simulation.....	81
5.1	Lasergenerierte Photonenrate	82
5.2	Augensicherheit	84
5.3	Hintergrundstrahlung.....	87
5.4	Monte-Carlo-Simulation	89
5.5	Modellverifikation	90
5.6	Vergleich direktes und indirektes Verfahren	93
6	Konzepte zur Unterdrückung von Hintergrundlicht	97
6.1	Photonenkorrelation.....	97
6.1.1	Modell.....	98
6.1.2	Bewertung.....	102
6.1.3	Adaptive Koinzidenz	104
6.2	Mehrfache Ereignisdetektion.....	108
6.2.1	Signal-Rausch-Verhältnis	109
6.2.2	Bewertung.....	112
6.3	Erhöhte Fensteranzahl	113
6.3.1	Varianz.....	114
6.3.2	Bewertung.....	118
6.4	Verfahrensauswahl	120
7	SPAD-basierter Zeilensensor für automobiler Anwendungen	121
7.1	Assistenzsysteme im Automobil.....	121
7.2	Sensorkonzeptionierung	123
7.2.1	Messkonzept	123
7.2.2	Dynamikbereich.....	124
7.2.3	Sichtfeld und Auflösung.....	124

7.3	Schaltungsentwurf.....	125
7.3.1	Pixelschaltung.....	126
7.3.2	Time-to-Digital-Converter	131
7.3.3	Datenauslese.....	134
7.3.4	Layout.....	135
7.3.5	Kenngößen.....	137
7.4	Kamerasystem	138
7.4.1	Laserquelle	139
7.4.2	Benutzerinterface	140
7.4.3	Datenauswertung.....	141
7.5	Charakterisierung.....	157
7.5.1	Single-Photon Avalanche Diode.....	157
7.5.2	Time-to-Digital-Converter	160
7.5.3	Photonenkorrelation.....	168
7.5.4	Kamerasystem.....	170
7.5.5	Zusammenfassung.....	181
8	Zusammenfassung und Ausblick	183
	Literaturverzeichnis.....	185
	Veröffentlichungen	191
	Anhang	193
A	Nichtideale Signalform im indirekten Verfahren	193
B	Schärfentiefe und Auflösungsvermögen des Sensors.....	197

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AEL	Accessible Emission Limit
APD	Avalanche Photodiode
CCD	Charge-Coupled Device
CMOS	Complementary Metal-Oxide-Semiconductor
CW	Continuous Wave
DCR	Dark Count Rate
DLL	Delay-Locked-Loop
DNL	Differentielle Nichtlinearität
EHP	Elektron-Loch-Paar
FOV	Field-of-View
FPAT	First-Photon Arrival Time
GUI	Graphical User Interface
IAT	Interarrival Time
INL	Integrale Nichtlinearität
LiDAR	Light Detection and Ranging
LSB	Least Significant Bit
MTF	Modulation Transfer Function
PDE	Photon Detection Efficiency
PDF	Probability Density Function
PEB	Premature Edge Breakdown
PLL	Phase-Locked-Loop
PM	Pulsmodulation
RLZ	Raumladungszone

RMS	Quadratischer Mittelwert
SBR	Signal-to-Background-Ratio
SNR	Signal-to-Noise-Ratio
SoC	System-on-Chip
SPAD	Single-Photon Avalanche Diode
TAT	Trap-Assisted-Tunneling
TDC	Time-to-Digital-Converter

Formelzeichen

Lateinische Schriftzeichen

Symbol	Einheit	Bedeutung
A	m^2	Fläche
A_F	m^2	Durchflutete Fläche der Messapertur
A_{Pix}	m^2	Pixelfläche
a_{Hist}	–	SNR des Histogramms
C_4	–	Wellenlängenkorrekturfaktor
C_6	–	Quellenausdehnungskorrekturfaktor
d	m	Distanz
d_{max}	m	Maximale Messdistanz
D	m	Durchmesser der Empfangsapertur
E	V/m	Elektrische Feldstärke
E_C	J	Leitungsbandenergie
E_F	J	Fermienergie
E_G	J	Bandlückenenergie
E_n	J	Energieniveau des n -ten Zustandes
E_P	J	Zulässige Pulsenergie
E_{PC}	J	Zulässige gewichtete Pulsenergie
E_{Ph}	J	Photonenenergie
E_V	J	Valenzbandenergie
f	m	Brennweite
$f_{\#}$	–	Blendenzahl
G_{th}	$1/\text{m}^3\text{s}$	Thermische Ladungsträgergenerationsrate
I_{lu}	A	Latchup-Strom

m^*	kg	Effektive Ladungsträgermasse
n	–	Koinzidenztiefe
n_{Air}	–	Brechungsindex von Luft
n_i	1/m ³	Intrinsische Ladungsträgerkonzentration
n_{Si}	–	Brechungsindex von Silizium
N_A	1/m ³	Dotierstoffkonzentration im p-Gebiet
N_B	–	Anzahl der Hintergrundereignisse
N_{Bin}	–	Anzahl der Klassen des Histogramms
N_D	1/m ³	Dotierstoffkonzentration im n-Gebiet
N_i	–	Zählwert des i -ten Zeitfensters
N_L	–	Anzahl der lasergenerierten Ereignisse
N_P	–	Pulsanzahl
N_{Ph}	–	Photonenanzahl
N_S	–	Anzahl der Subpixel / SPADs für Koinzidenz
$N_{T,\text{eff}}$	1/m ³	Effektive Störstellenkonzentration im pn-Übergang
N_W	–	Anzahl der Messfenster
N_Z	–	Anzahl der akkumulierten Messzyklen je Frame
p	–	Erfolgswahrscheinlichkeit
p_i	–	Erfolgswahrscheinlichkeit der i -ten Klasse
P	W	Strahlungsleistung
P_0	W	Strahlungsleistung am Positionsursprung
$P_{1,n}$	1/s	PDF des ersten Ereignisses bei n SPADs
P_{abs}	W	Absorbierte Strahlungsleistung
P_B	–	Binomialverteilung
$P_{IA,n}$	1/s	PDF der Zwischenankunftszeit bei n SPADs

$P_{k,n}$	1/s	PDF des k -ten Ereignisses bei n SPADs
P_L	W	Optische Quellenleistung
P_n	–	Auslösewahrscheinlichkeit durch Elektronen
P_{np}	–	Kombinierte Auslösewahrscheinlichkeit
P_p	–	Auslösewahrscheinlichkeit durch Löcher
P_P	–	Poisson-Verteilung
P_{PT}	W	Zulässige mittlere Pulskettenleistung
P_{Ph}	1/s	PDF der Photonendetektion
r	1/s	Ereignisrate
r_B	1/s	Ereignisrate der Hintergrundstrahlung
$r_{B,Est}$	1/s	Schätzwert der Ereignisrate der Hintergrundstrahlung
$r_{B,T}$	1/s	Effektive Ereignisrate des Hintergrundes unter Totzeit
r_{Coin}	1/s	Koinzidenzereignisrate
r_{LB}	1/s	Summenereignisrate des Signal und Hintergrundes
$r_{LB,T}$	1/s	Effektive Summenereignisrate unter Totzeiteinfluss
r_L	1/s	Ereignisrate des reflektierten Lasersignals
$r_{L,T}$	1/s	Effektive Ereignisrate des Signals unter Totzeiteinfluss
r_T	1/s	Effektive Ereignisrate unter Totzeiteinfluss
R	–	Reflexionsfaktor
R_Q	Ω	Quenching-Widerstand
R_T	–	Reflexionsfaktor des Zielobjektes
S_B	W/m ²	Sonnengenerierte Bestrahlungsstärke
S_L	W/m ²	Quellengenerierte Bestrahlungsstärke
S_S	W/m ²	Bestrahlungsstärke auf dem Sensor

t	s	Zeit
T	K	Absolute Temperatur
T_{Air}	–	Transmissionsfaktor der Luft
T_{Bin}	s	Breite der Klassen des Histogramms
T_{CW}	s	Periodendauer
T_{F}	s	Filterbreite
T_i	s	Dauer des i -ten Zeitfensters
T_{O}	–	Transmissionsfaktor der Optik
T_{P}	s	Pulsdauer
T_{T}	s	Totzeit
$T_{\text{T,NV}}$	s	Nicht verlängerbare Totzeit
$T_{\text{T,V}}$	s	Verlängerbare Totzeit
T_{TOF}	s	Lichtlaufzeit
T_{W}	s	Zeitfensterbreite
U_{BD}	V	Durchbruchspannung
U_{ex}	V	Überschussspannung
U_{R}	V	Sperrspannung
U_{S}	V	Anodenspannung der SPAD
U_{th}	V	Schwellspannung
x	m	Ortsvariable
x_{dn}	m	Ausdehnung der RLZ im n-Gebiet
x_{dp}	m	Ausdehnung der RLZ im p-Gebiet
x_{n}	m	Ausdehnung des n-Gebietes
x_{p}	m	Ausdehnung des p-Gebietes

Griechische Schriftzeichen

Symbol	Einheit	Bedeutung
α	1/m	Absorptionskoeffizient
α_H	rad	Horizontaler Quellenöffnungswinkel
α_V	rad	Vertikaler Quellenöffnungswinkel
β	1/m	Ionisationskoeffizient
β_n	1/m	Ionisationskoeffizient für Elektronen
β_p	1/m	Ionisationskoeffizient für Löcher
γ	rad	Quellenausdehnung
Γ_{DC}	–	Tastverhältnis
ε_{Si}	–	Relative Permittivität von Silizium
η_A	–	Abbildungsverhältnis
η_{FF}	–	Füllfaktor
η_L	–	Wirkungsgrad der Lambertschen Reflexion
η_{PDE}	–	Photonendetektionseffizienz
ϑ	s	Koinzidenzzeit
Θ	–	Heaviside-Funktion
κ	m ²	Einfangquerschnitt der Gitterdefekte
λ	m	Wellenlänge
μ	–	Erwartungswert
μ_i	–	Erwartungswert im i -ten Zeitfenster / Klasse
μ_T	–	Erwartungswert eines Zählprozesses mit Totzeit
ν	1/s	Frequenz der elektromagnetischen Strahlung
ν_{th}	m/s	Thermische Ladungsträgergeschwindigkeit

ρ	Ω/m^2	Spezifischer Widerstand
σ^2	–	Varianz
σ_i^2	–	Varianz des i -ten Zeitfensters / Klasse
σ_T^2	–	Varianz eines Zählprozesses mit Totzeit
τ	s	Zeitkonstante
φ	rad	Phasenwinkel
ω_P	J/m^2	Pulsenergiedichte

Physikalische Konstanten

Symbol	Einheit	Wert	Bedeutung
c	m/s	299792458	Lichtgeschwindigkeit im Vakuum
h	Js	$6,626 \cdot 10^{34}$	Plancksches Wirkungsquantum
k_B	J/K	$1,381 \cdot 10^{-23}$	Boltzmann-Konstante
q	As	$1,602 \cdot 10^{-19}$	Elementarladung
ϵ_0	As/Vm	$8,854 \cdot 10^{-12}$	Permittivität des Vakuums

1 Einleitung

Für viele Anwendungen wie Sicherheitsüberwachung, optische Geschwindigkeits- und Entfernungsbestimmung, Sicherheitssysteme im Automobil, Robotik und Architektur spielt neben der Lichtintensität einer Szene oder eines Objektes die Tiefeninformation eine entscheidende Rolle [1]. Für die berührungslose, dreidimensionale Erfassung können Schallwellen oder elektromagnetische Wellen eingesetzt werden. Mit der Entwicklung des CCDs von George Smith und Willard Boyle in den AT&T Bell Labs im Jahre 1969 [2] und der Entwicklung von CMOS-basierten Bildsensoren in den Jet Propulsion Labs in den frühen 1990ern von Eric Fossum [3] wurde der Weg für die Erfassung von dreidimensionalen Bildern unter Verwendung von elektromagnetischen Wellen im sichtbaren Spektrum geebnet. Grundsätzlich lassen sich drei Verfahren zur Gewinnung der Tiefeninformation einer Szene mittels optischen Wellen unterscheiden: Interferometrie, Triangulation und Laufzeitbestimmung [4]. Für Entfernungen, wie sie in den oben genannten Anwendungen gefordert werden, bietet die Messung der Laufzeit das einfachste Funktionsprinzip [5]. Bei diesem wird die Entfernung eines Objektes aus der Ausbreitungsgeschwindigkeit und der Zeit, welche die Lichtwelle von der Emission über die Reflexion am Objekt zurück zum Detektor benötigt, berechnet. Dieses Verfahren wird analog zum bekannteren Radio Detection and Ranging (Radar), bei welchem Radiowellen zur Distanzmessung verwendet werden, als Light Detection and Ranging (LiDAR) bezeichnet. Grundsätzlich werden dabei zwei Methoden unterschieden: Die

direkte Messung, bei welcher die Zeit zwischen dem Senden und Empfangen eines Lichtpulses mittels einer elektronischen Stoppuhr gemessen wird [5], und die indirekte Messung, bei welcher die Laufzeit aus der zeitlichen Verschiebung eines intensitätsmodulierten Signals berechnet wird [4].

Um die erforderliche optische Leistung der Strahlungsquelle gering zu halten, werden Sensoren mit sehr hoher Empfindlichkeit gefordert. Aufgrund ihrer Fähigkeit zur Detektion einzelner Photonen [6] und der zeitlichen Auflösung im Pikosekundenbereich [7], eignen sich Single-Photon Avalanche Dioden (SPAD) besonders für die direkte Laufzeitmessung. SPADs sind Avalanche Photodioden (APD), die über ihre negative Durchbruchspannung hinaus vorgespannt werden. In diesem Bereich, auch bekannt als Geiger-Betrieb, reicht ein einzelnes Photon aus, das im aktiven Bereich der Diode ein Elektron-Loch-Paar (EHP) erzeugt, um den Lawineneffekt in Gang zu setzen, der zu einem makroskopischen Stromfluss durch die Diode führt [8]. In der Vergangenheit wurden SPADs mittels speziellen Technologien, welche auf die Bedürfnisse der Sensoren ausgerichtet waren, hergestellt. In den letzten Jahren ist es jedoch gelungen leistungsfähige SPADs in Standard-CMOS-Technologie herzustellen [9]. Dies reduziert die Kosten und ermöglicht es, die Steuer- und Auswerteelektronik für die SPADs direkt in die einzelnen Pixel der Sensorarrays zu integrieren [10]. Mit entsprechender Beschaltung können SPADs auch als Photonenzähler und so zur indirekten Messung der Laufzeit eingesetzt werden [6].

1.1 Stand der Technik

Verschiedene Forschungsgruppen weltweit beschäftigen sich mit der Entwicklung von SPAD-basierten Sensoren zur laufzeitbasierten Distanzmessung. Diese Sensoren werden dabei für verschiedene Anwendungszwecke optimiert und regelmäßig auf Konferenzen oder in Fachzeitschriften veröffentlicht.

Die Forschung des Toyota Central R&D Labs in Japan um Christiano Niclass beschäftigt sich mit Scanner-basierten Sensoren zur direkten Laufzeitmessung. Dabei wird ein Laser mittels eines rotierenden Spiegels abgelenkt und die Laufzeit des Laserpulses mit Time-to-Digital-Convertern (TDC) erfasst. Vorteile dieses Ansatzes ist ein sehr hoher Füllfaktor von 70 % durch Makropixel sowie

Tabelle 1.1: Stand der Technik

Parameter	Einheit	[11]	[12]	[13]	[14]	[1], [15]	[16]	[17]	[18]
		Niclass	Aktia	Villa	Bronzi	Bellisai	Portaluppi	Perenzoni	Vornicu
		2014	2017	2014	2014	2013	2018	2017	2014
Verfahren	-	Direkt	Direkt	Direkt	CW	PM	Direkt	Direkt	Direkt
Methode	-	Scanning	Scanning	Flash	Flash	Flash	Flash	Flash	Flash
Technologie	nm	180	180	350	350	350	180	150	180
Pixelarray	-	1 x 16	32 x 4	32 x 32	64 x 32	32 x 32	32 x 32	64 x 64	64 x 64
SPAD Größe	µm	21	10 x 15 ¹	30	30	20	32	121	12
SPAD Form	-	Quadrat ²	Rechteck ²	Kreis	Kreis	Kreis	Quadrat ²	Quadrat ²	Kreis
SPADs je Pixel	-	24	8	1	1	1	4	8	1
Füllfaktor	%	70	28	3,14	3,14	3,14	9,6	26,5	3,5
Dunkelzählrate	cps	2,65k	-	120	100	2,5k	-	6,8k	42k
Zeitauflösung	ps	416,7	208	312	-	-	50	250	150
Messbereich	ns	853,3	-	320	-	150	204,8	6400	297
Messdistanz	m	100	20	5	20	14	-	367	0,74
Genauigkeit	%	0,14	0,55	0,014	2,1	1,92	-	0,37	1,35
Reflektanz	%	9	10	-	70	70	-	-	-
Bildrate	Hz	10	-	200	25	10	-	7,68	38,5
Hintergrund	klx	70	75	-	0,45	0,45	-	-	-
Laserleistung	mW	21	-	90	400	750	-	-	0,16

¹ Schätzwert aus Abbildung² Abgerundete Ecken

eine hohe Genauigkeit von 0,14 % bei einer maximalen Messdistanz von 128 m. Weiterhin erlaubt das Scannen des Lasers eine hohe Leistungsdichte am Zielobjekt, was zusammen mit der Detektion von zeitlich korrelierten Einzelphotonen eine Messung bei hohem Hintergrundlicht von 70 klx erlaubt [11]. Entscheidender Nachteil für die vorliegende Arbeit ist die Verwendung von mechanischen Komponenten und der Ansatz eines scannenden Verfahrens. Eine andere Veröffentlichung des Toyota R&D Labs beschreibt einen Sensor für das direkte Verfahren, bei welchem das Scannen des Lasers mittels eines mikroelektromechanischen Systems (MEMS) erfolgt. Auch hier kommen Makropixel zur Verbesserung des Füllfaktors und zur Detektion zeitlich korrelierter Einzelphotonen zum Einsatz [12]. Die Forschergruppe des Polytechnikums in Mailand um Franco Zappa entwickelt auf den am Fraunhofer IMS entwickelten SPADs basierende solid state Sensoren für biologische und medizinische Anwendung, aber auch für die dreidimensionale Erfassung der Umwelt. Im Rahmen des MiSPiA Projekts (Microelectronic Single Photon 3D Imaging Arrays for low-light high-speed Safety and Security Applications, 2010 bis 2013) wurden verschiedene Sensorarrays für die direkte und indirekte Laufzeitmessung entwickelt. Der Sensor für die direkte Messung mit einer Auflösung von 32 x 32 Pixel hat einen Füllfaktor von lediglich 3,14 %, was dem Flash-Ansatz geschuldet ist. Für eine vollständig parallele Arbeitsweise ist es erforderlich, die notwendige Elektronik zur Ansteuerung und Auswertung der SPAD in jedem Pixel zu integrieren [13]. Die auf dem indirekten Messprinzip beruhenden Sensoren erreichen bei einer Array-größe von 64 x 32 bzw. 32 x 32 Pixel einen identischen Füllfaktor [14], [15]. Dies zeigt, die Elektronik für das direkte und indirekte Verfahren benötigt näherungsweise dieselbe Fläche. Gleichzeitig zeigen die indirekten Verfahren eine geringere Genauigkeit gegenüber dem direkten Verfahren mit Scanner. Ein anderer 32 x 32 Pixel Sensor erreicht dank der Nutzung eines TDCs für vier SPADs einen Füllfaktor von 9,6 %. Auch hier wird die Detektion von korrelierten Photonen zur Unterdrückung hoher Hintergrundstrahlung eingesetzt [16]. Der in [17] präsentierte 64 x 64 Pixel Sensor erreicht einen Füllfaktor von 26,5 % durch Makropixel mit jeweils 8 SPADs. Hierbei kommt zudem die Detektion korrelierter Einzelphotonen mit variabler Anzahl zur Anwendung. Die Forscher des Instituts für Mikroelektronik in Sevilla zeigen ein Array für die direkte

Laufzeitmessung mit 64×64 Pixel, einem Füllfaktor von 3,5 % und nominaler Auflösung von 145 ps bei einer Messdistanz von 44 m [18]. Eine Demokamera für den Nahbereich mit diesem Sensor zeigt eine Auflösung von 1 cm bei einer Messdistanz von bis zu 66 cm [19].

Neben den genannten Publikationen diverser Forschungsgruppen existieren auch bereits kommerzielle Produkte im Bereich LiDAR. So bietet Velodyne LiDAR Inc. Scanner-basierte Systeme für Automobil-, Industrie- oder Überwachungsanwendungen an, welche jedoch relativ groß und kostenintensiv sind [20]. LeddarTech Inc. bietet solid state Flash LiDAR Module für verschiedene Anwendungen im Automobil- und Industriebereich an. Aus den Datenblättern ist jedoch nicht ersichtlich, auf welchem Verfahren die Sensoren basieren und unter welchen Bedingungen die angegebenen Reichweiten und Genauigkeiten erzielt wurden. Zuletzt sei noch Quanergy Systems Inc. genannt, auch diese bieten Flash und Scanner-basierte LiDAR Systeme für verschiedene Anwendungen. Jedoch lassen sich auch hier keine näheren Informationen zur Technologie der angebotenen Systeme finden.

Prinzipiell zeigt ein Blick auf die Übersicht in Tabelle 1.1, dass das direkte Messverfahren eine bessere Distanzauflösung bietet als das indirekte Verfahren, da kurze Laserpulse eine höhere optische Leistung und damit ein höheres Signal-Rausch-Verhältnis (SNR) erlauben [21]. Aus den beschriebenen Publikationen lässt sich der geringe Füllfaktor als ein Problem aktueller Arrays erkennen. Dies ist der Komplexität der Schaltung, welche zur Laufzeitmessung oder Einzelphotonenzählung im Pixel integriert werden muss, geschuldet. Ein Ansatz zu dessen Lösung ist die gemeinsame Nutzung der Ressourcen oder die Auslagerung dieser Komponenten [22]. Makropixel können ebenfalls den Füllfaktor steigern, vergrößern aber auch das Pixel insgesamt, wodurch dieser Ansatz für hochauflösende Sensoren ungeeignet ist. Des Weiteren wird an rückseitig beleuchteten SPADs, bei welchen die Steuer- und Auswerteelektronik auf einem zweiten, separaten Wafer untergebracht ist [23], sowie an Strahlungskonzentration durch Mikrolinsen auf jedem Pixel geforscht [24]. Ein weiteres großes Problem ist die durch die Sonne und andere künstliche Strahlungsquellen erzeugte Hintergrundstrahlung, welche bei realen Szenen immer vorhanden ist und die Präzision der Messverfahren negativ beeinflusst [25].

1.2 Ziel und Aufbau der Arbeit

Das Ziel der Arbeit besteht im Entwurf eines SPAD-basierten Sensors zur laufzeitbasierten Distanzmessung, welcher das Problem des geringen Füllfaktors sowie wie der hohen Hintergrundlichtabhängigkeit angeht. Ansätze hierfür lassen sich aus dem Stand der Technik entnehmen. So lässt sich der Füllfaktor durch Makropixel oder einen Zeilensensor erhöhen, während die Detektion zeitlich korrelierter Einzelphotonen eine Verbesserung der Resistenz gegen hohes Hintergrundlicht verspricht.

Die Arbeit ist wie folgt aufgebaut: In Abschnitt 2 wird das Detektorelement, die SPAD, vorgestellt. Hierfür werden zunächst der Aufbau und die Funktionsweise der Diode erläutert sowie die Kenngrößen und deren physikalischer Ursprung betrachtet. Abschließend werden die typischen Beschaltungen der SPAD zur Erzeugung digital verwertbarer Ausgangssignale beschrieben. In Abschnitt 3 werden die drei grundlegenden Funktionsprinzipien der optischen Entfernungsmessung: Interferometrie, Triangulation und Laufzeitbestimmung betrachtet sowie deren Funktionsweisen und Besonderheiten erläutert. In Abschnitt 4 wird die für das direkte und indirekte Laufzeitverfahren relevante Statistik untersucht. Da beide Prinzipien auf der Detektion einzelner Photonen beruhen, wird zunächst die Photonenstatistik verschiedener Lichtquellen betrachtet. Darauf basierend werden die Laufzeitverfahren hinsichtlich ihrer zu erwartenden systematischen und zufälligen Fehler untersucht. Diese Betrachtungen dienen im Weiteren der Auswahl des für die Zielanwendung geeigneten Verfahrens. Abschnitt 5 beschreibt das Modell und die Simulationsmethodik der Laufzeitverfahren, welche zur Leistungsabschätzung und Simulation verschiedener Verfahrensmodifikationen entwickelt wurden. Das Modell basiert auf der Berechnung der Photonendetektionsraten der SPAD für gegebene System- und Umgebungsparameter sowie einer Monte-Carlo-Simulation auf Basis der berechneten Raten und der Photonenstatistik. In Abschnitt 6 werden verschiedene Konzepte zur Lösung des beschriebenen Problems der Hintergrundabhängigkeit vorgestellt. Da ein Zeilensensor entworfen wird, wird das Problem des geringen Füllfaktors nicht weiter verfolgt. Die Konzepte basieren sowohl auf dem direkten als auch auf dem indirekten Verfahren. Jedes Konzept wird

unter Zuhilfenahme der Statistik auf sein Potential untersucht und bewertet. Abschließend erfolgt die Auswahl des vielversprechendsten Verfahrens für den Entwurf des Sensors. Abschnitt 7 behandelt den Entwurf sowie die Charakterisierung des SPAD-basierten Zeilensensors. Zunächst werden die Anforderungen an den Sensor auf Grundlage der betrachteten Anwendung festgelegt. Darauf folgt der Entwurf des Sensors mit dessen Aufbau und der Beschreibung verschiedener funktionsrelevanter Schaltungskomponenten. Neben dem Sensor selbst wird auch das Sensorsystem samt Laserquelle, Ansteuerung und Datenauswertung sowie Visualisierung erläutert. Abschließend werden die Resultate der Charakterisierung vorgestellt, mit den theoretischen Erwartungen verglichen und abschließend bewertet. Abschnitt 8 enthält eine kurze Zusammenfassung der Arbeit mit den Ergebnissen sowie einen Ausblick auf zukünftige Schritte zur weiteren Optimierung der Laufzeitmessung.

2 Single-Photon Avalanche Diode

In diesem Abschnitt werden die Funktionsweise und die physikalischen Eigenschaften der SPADs, welche für die im Rahmen dieser Arbeit betrachteten Sensoren für die Umwandlung von optischen in elektrische Signale dienen, vorgestellt. Zunächst werden der Aufbau der SPAD und die Funktionsweise anhand der Strom-Spannungs-Kennlinie und des Bänderdiagramms erläutert. Im Weiteren werden relevante Kenngrößen der Diode und für das Design eines Sensors zu beachtende Rauschquellen untersucht. Abschließend werden grundlegende Schaltungsprinzipien zum Betrieb der SPAD aufgezeigt.

2.1 Einführung

Eine Photodiode besteht aus einem pn-Übergang, der in Sperrrichtung mit Spannung beaufschlagt ist. Dadurch bildet sich am Übergang zwischen p- und n-Gebiet eine Zone ohne freie Ladungsträger, die Raumladungszone (RLZ), aus. Trifft ein Photon mit ausreichend hoher Energie auf diese Zone, wird dieses absorbiert und ein EHP generiert. Durch das elektrische Feld in der RLZ driftet das Elektron in das n-Gebiet und das Loch in das p-Gebiet und wird den Kontakten der Photodiode zugeführt. Sind Anode und Kathode leitend verbunden, kommt so ein Stromfluss in Abhängigkeit der Bestrahlungsstärke zustande. Zur Vergrößerung des photosensitiven Bereichs wird in der Regel eine intrinsische Schicht zwischen p- und n-Bereich platziert.

Standard Photodioden generieren maximal ein EHP je eingefallenes Photon. Für Anwendungen mit geringen Lichtstärken oder kurzen Belichtungszeiten resultiert daraus ein geringes elektrisches Signal, welches im Rauschen nur schwer auszumachen ist. Für ein höheres SNR und damit eine bessere Detektion des Lichtes wird eine interne Verstärkung der durch Photonen erzeugten Ladungsträger benötigt. In einer APD wird dies durch einen Bereich innerhalb der Diode mit hohem elektrischem Feld erreicht. Das Feld beschleunigt die Ladungsträger derart, dass ihre Energie ausreicht, um durch Kollisionen mit Atomen weitere EHPs zu erzeugen, welche wiederum beschleunigt werden und weitere Paare generieren. Die Stärke des elektrischen Feldes ist so ausgelegt, dass nur Elektronen ausreichend Energie erhalten, um weitere EHPs zu erzeugen. In diesem Betriebsfall ist die Verstärkung konstant, er wird als linearer Betrieb bezeichnet [26]. Die Verstärkung hängt von der Vorspannung ab und ergibt sich aus empirischen Messungen zu

$$G \approx \frac{1}{1 - \left(\frac{U_R}{U_{BD}}\right)^3} \quad (2.1)$$

mit der Sperrspannung U_R und der Durchbruchspannung U_{BD} [27].

Die SPAD ist eine APD, die über die Durchbruchspannung hinaus vorgespannt wird. Aufgrund der ähnlichen Funktionsweise des Geiger-Müller-Zählers für radioaktive Strahlung wird dies als Geiger-Betrieb bezeichnet [8]. In diesem Fall ist das elektrische Feld hoch genug, um auch Löcher derart zu beschleunigen, dass deren kinetische Energie zur Erzeugung weiterer EHPs ausreicht [26]. Hierbei reicht ein einzelnes Photon aus, um den Lawineneffekt auszulösen. Dabei steigt der Strom durch die Diode in kurzer Zeit auf einen makroskopischen Wert an. Da sich die Diode dabei im Durchbruchbereich ihrer Kennlinie befindet, fließt der Strom solange, bis die Spannung durch externe Beschaltung unter die Durchbruchspannung abgesenkt wird [8]. Dies ist notwendig, um ein weiteres Photon detektieren zu können [28]. Ziel der externen Beschaltung ist es, den Ladungsfluss so gering wie möglich zu halten, um Afterpulsing, Erhitzung der Diode und Crosstalk zu vermeiden [8].

2.2 Aufbau

Grundsätzlich besteht eine SPAD aus zwei Schichten von unterschiedlich dotiertem Silizium. Silizium ist ein Element der vierten Hauptgruppe des Periodensystems und hat demnach vier Valenzelektronen, die für Bindungen mit anderen Atomen zur Verfügung stehen. Es bildet im Gitter eine Diamantstruktur, wobei die vier Valenzelektronen jeweils mit einem Elektron eines benachbarten Atoms eine kovalente Bindung eingehen [26]. Am absoluten Temperaturnullpunkt sind alle Elektronenpaarbindungen intakt und keine freien Ladungsträger im Silizium vorhanden, weshalb es als idealer Isolator wirkt. Mit steigender Temperatur schwingen die Atome, wodurch manche Elektronen ausreichend Energie gewinnen um die Bindung aufzubrechen und sich frei im Kristall bewegen können [27]. In intrinsischem Silizium ist die Zahl freier Elektronen und Löcher identisch, da jedes freie Elektron einen leeren Platz in einer Bindung (Loch) hinterlässt. Um mehr freie Ladungsträger einer Art zu erhalten, wird Silizium dotiert. Dabei werden Atome der dritten oder fünften Hauptgruppe in das intrinsische Silizium eingebracht. Ein Atom der fünften Hauptgruppe liefert ein zusätzliches Elektron, welches keine Bindung eingehen kann und als freier Ladungsträger im Kristall zur Verfügung steht. Das Material wird als n-dotiert bezeichnet, da ein Überschuss an negativ geladenen Elektronen besteht. Mit Atomen der dritten Hauptgruppe dotiertes Silizium wird als p-dotiert bezeichnet, da mehr freie Löcher zur Verfügung stehen [27].

Die Energieniveaus der gebundenen und freien Elektronen werden als Valenz- bzw. Leitungsband bezeichnet, da sich die im einzelnen Atom festen Energieniveaus aufgrund der Wechselwirkung im Kristallgitter aufteilen. Zwischen beiden Bändern, d. h. zwischen der oberen Grenze der Valenzbandenergie E_V und der unteren Grenze der Leitungsbandenergie E_C , befindet sich die Bandlücke, die Elektronen aus dem Valenzband überwinden müssen, um ins Leitungsband zu gelangen. Die Besetzung der einzelnen Zustände in den Bändern folgt der Fermi-Dirac-Statistik. Bei welcher Energie die Hälfte der Zustände besetzt sind, wird durch die Fermienergie E_F angegeben. In intrinsischem Silizium liegt diese Energie näherungsweise in der Mitte der Bandlücke. Durch Dotierung verschiebt sich diese Energie. Wird Silizium n-dotiert, steigt die Zahl freier

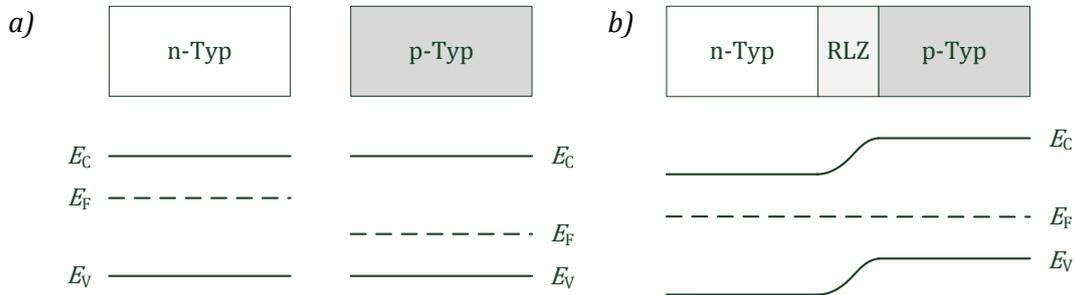


Abbildung 2.1: Bändermodell der n- und p-dotierten Gebiete vor (a) und nach (b) Kontakt nach [26]

Elektronen und die Besetzungswahrscheinlichkeit der Zustände steigt an, wodurch die sich Fermienergie zu höherer Energie verschiebt. Analog verschiebt sie sich bei p-Dotierung hin zu niedrigerer Energie [27].

Wird n-dotiertes und p-dotiertes Silizium zusammengebracht, diffundieren die freien Elektronen vom n-Gebiet in das p-Gebiet und rekombinieren dort mit den vorhandenen Löchern. Durch den Abfluss der Elektronen ändert sich die Ladung im n-Gebiet von neutral in positiv. Analog wird das p-Gebiet durch den Elektronenzufluss negativ geladen. Das sich durch die unterschiedliche Ladung ergebende elektrische Feld wirkt der Diffusion entgegen, bis sich beide Vorgänge im Gleichgewicht befinden. Dabei entsteht an der Kontaktstelle ein Gebiet ohne freie Ladungsträger, die sogenannte RLZ. In diesem Gleichgewichtszustand ist die Fermienergie über dem gesamten Halbleiter konstant. Die entstandene Potentialbarriere ist abhängig von der Lage der Fermienergien in den dotierten Schichten, welche wiederum von den jeweiligen Dotierstoffkonzentrationen abhängen [27].

Für die vorliegende Arbeit werden SPADs in Form integrierter Dünnschichtdioden verwendet. Diese werden auf einem p-dotierten Siliziumsubstrat integriert. Im einfachsten Fall wird dafür eine stark n-dotierte Schicht implantiert. SPADs benötigen für ihre Funktion ein homogenes elektrisches Feld über der gesamten photoaktiven Fläche. Bei dieser einfachen Variante ist das nicht gegeben, da das elektrische Feld am Rand der Implantierung ansteigt und zu vorzeitigem Durchbruch (Premature Edge Breakdown, PEB) führen kann. Ursächlich für diese Felderhöhung ist das größere Volumen des p-Gebietes, das an

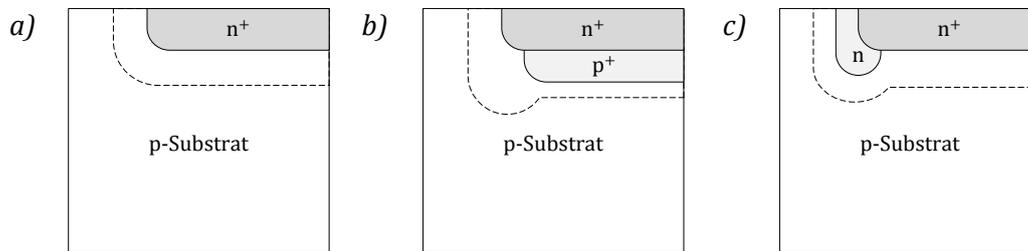


Abbildung 2.2: Querschnitt einer integrierten SPAD (a), Prävention von PEB durch Anreicherung (b) und Guard-Ring (c) nach [4]

den Rändern dem n-Gebiet gegenüber steht. Durch die höhere Zahl an freien Löchern reduziert sich in diesem Bereich die Ausdehnung der RLZ, was bei gleicher Spannung eine Erhöhung des Feldes bewirkt [27]. Um PEB zu verhindern, wurden verschiedene Techniken entwickelt. Eine Methode ist das Implantieren einer höher dotierten p-Schicht unterhalb der n-Schicht mit einer geringeren Ausdehnung. Damit wird das hohe elektrische Feld auf das Zentrum der Diode begrenzt. Eine weitere Möglichkeit ist das Einbringen eines die n-Schicht umgebenden Rings mit schwächerer n-Dotierung. Dieses als Guard-Ring bezeichnete Konstrukt schwächt das elektrische Feld am Rand ab [8].

Die hohe Dotierung der implantierten n-Schicht wird benötigt, um die Dicke möglichst gering zu halten, sodass einfallende Photonen nicht vor dem Erreichen der RLZ absorbiert werden. Weiterhin wird die Ausdehnung der RLZ in das n-Gebiet reduziert und so das elektrische Feld in diesem Bereich erhöht. Die im Rahmen des MiSPiA-Projektes vom Fraunhofer IMS und dem Polytechnikum Mailand entwickelte und in 0,35- μm -CMOS-Technologie gefertigte SPAD baut auf einem p-Substrat auf. Durch die hohe p-Dotierung und die Anreicherung in der n-Wanne wird das aktive Gebiet definiert [29]. Um PEB zu vermeiden, wird das elektrische Feld durch höhere Dotierung des n-Gebietes im Zentrum des aktiven Gebietes konzentriert und durch niedrigere p-Dotierung am Rand des aktiven Gebietes abgeschwächt [9]. Durch das Integrieren der Diode in eine n-Wanne erreicht man eine Isolation der Diode vom Substrat. Freie Ladungsträger aus dem Substrat können so nicht in den pn-Übergang der SPAD gelangen und dort einen Durchbruch auslösen. Im Betrieb liegt das Substrat auf Masse und die Kathode der SPAD auf einem hohen positiven Potential, genau

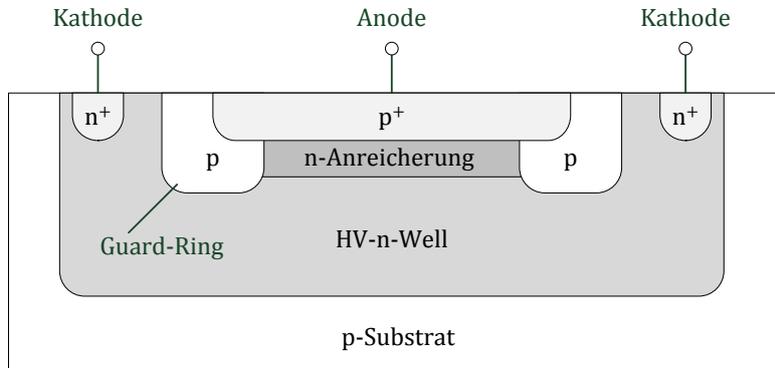


Abbildung 2.3: Querschnitt der im Rahmen des MiSPiA-Projektes entwickelten SPAD in 0,35- μm -CMOS-Technologie nach [29]

wie die SPAD ist die parasitäre Diode damit in Sperrrichtung vorgespannt. Aufgrund der geringeren Dotierung des Substrates und der Wanne gegenüber der SPAD, ist das elektrische Feld geringer und es findet kein Durchbruch statt [29].

2.3 Funktionsweise

Trifft ein Photon mit der Energie E_{ph} auf Silizium, kann es im Halbleitermaterial absorbiert und ein EHP erzeugt werden. Voraussetzung für Absorption ist, dass die Photonenenergie ausreicht, um ein Elektron aus dem Valenz- in das Leitungsband zu heben, also mindestens der Bandlückenenergie des Halbleiters $E_G = E_C - E_V$ entspricht. Im Falle von Silizium mit einer Bandlückenenergie von 1,12 eV [30] ergibt sich eine maximale Wellenlänge von

$$\lambda = \frac{hc}{E_G} = 1108 \text{ nm} . \quad (2.2)$$

Durch das in der RLZ herrschende elektrische Feld wird ein dort erzeugtes EHP getrennt. Das Elektron wird in Richtung des n-Gebietes und das Loch zum p-Gebiet beschleunigt [26]. Für den Betrieb als SPAD wird die APD oberhalb ihrer negativen Durchbruchspannung U_{BD} betrieben. Bei dieser Spannung ist das elektrische Feld innerhalb der RLZ stark genug, um sowohl Elektronen als auch Löcher derart zu beschleunigen, dass deren kinetische Energie zur Generation

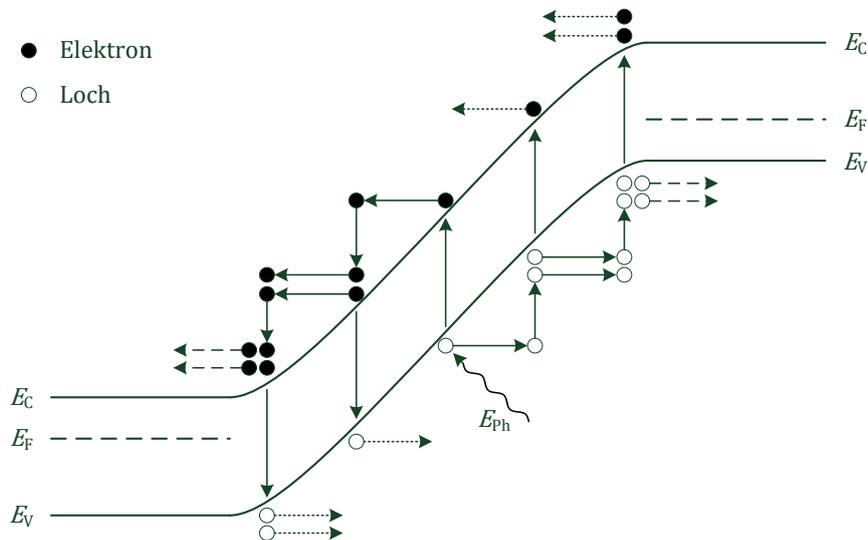


Abbildung 2.4: Photonen induzierter Lawinendurchbruch eines pn-Übergangs unter stark negativer Vorspannung nach [26]

weiterer EHPs durch Stoßionisation ausreicht [8]. Der prinzipielle Vorgang eines durch ein einfallendes Photon verursachten Lawinendurchbruchs ist in Abbildung 2.4 dargestellt. Das einfallende Photon hebt ein Elektron aus dem Valenz- in das Leitungsband. Anschließend wird das Elektron in Richtung des n-Gebietes beschleunigt, bis es ausreichend Energie hat, um durch Stoßionisation ein weiteres EHP zu erzeugen. Bei der Kollision verliert es zunächst seine kinetische Energie. Nun werden beide Elektronen wieder beschleunigt und erzeugen ein drittes und viertes EHP. Da im Falle einer SPAD auch Löcher EHPs durch Stoßionisation erzeugen können, entstehen rechts von der ursprünglichen Einfallsstelle des Photons weitere freie Elektronen, welche den Durchbruch weiter in Gang halten. Das Resultat ist ein konstanter Stromfluss durch die Diode, welcher nur durch die Reduktion des elektrischen Feldes aufgrund der generierten freien Ladungsträger begrenzt wird [31]. Abbildung 2.5 zeigt die qualitative Kennlinie einer Siliziumdiode. Befindet sich die SPAD im Wartezustand auf ein einfallendes Photon, fließt kein Strom durch die Diode und die Spannung liegt über der negativen Durchbruchspannung U_{BD} (Punkt 1). Wird durch ein einfallendes Photonen oder andere Ursachen ein Lawinendurchbruch ausgelöst, steigt der Strom innerhalb kurzer Zeit auf einen makroskopischen Wert an. In

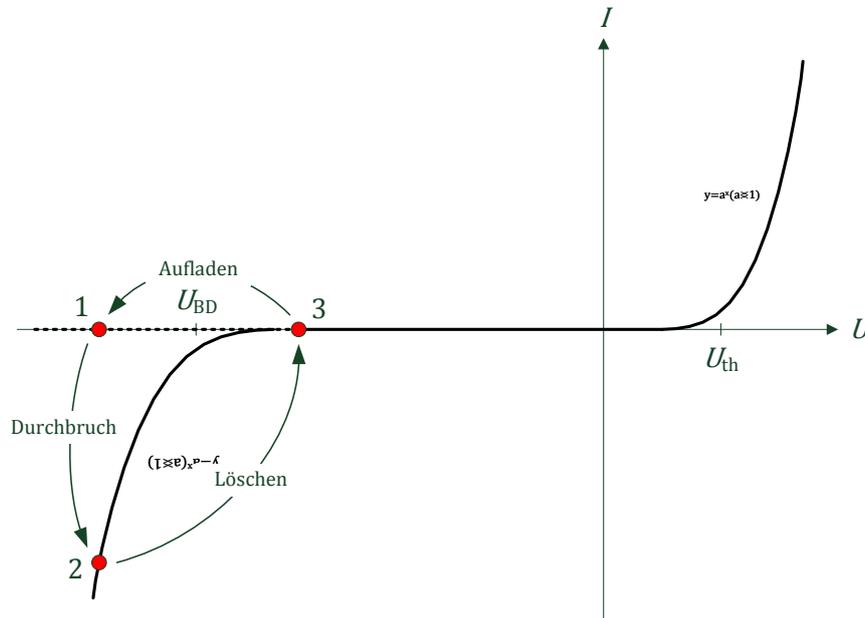


Abbildung 2.5: Kennlinie einer Siliziumdiode mit Betriebspunkten der SPAD bei Detektion eines Photons nach [32]

der Kennlinie geht der aktuelle Betriebspunkt auf die bekannte Kennlinie einer Diode über (Punkt 2). Die Spannung bleibt dabei unverändert. Im nächsten Schritt wird der Strom durch die Diode durch Absenken der Spannung unter die Durchbruchspannung mithilfe externer Beschaltung gelöscht. Der Betriebspunkt wandert auf der Kennlinie entlang hin zu geringerer Spannung (Punkt 3). Um ein weiteres Photon detektieren zu können, wird anschließend die Spannung über der SPAD wieder auf den Ursprungswert angehoben (Punkt 1). Dieser Zyklus wiederholt sich für jedes detektierte Photon [32].

2.4 Kenngrößen

Zur Beurteilung einer SPAD können verschiedene Kenngrößen betrachtet werden. Eine Kenngröße ist die Photonendetektionseffizienz (Photon Detection Efficiency, PDE), welche die Wahrscheinlichkeit für das Erkennen eines einfallenden Photons angibt. Die zeitliche Auflösung einer SPAD gibt die Varianz an,

mit der die Ankunftszeit eines Photons bestimmt werden kann [8]. Ein weiterer wichtiger Punkt ist das Rauschen, welches in 2.4.2 genauer betrachtet wird.

2.4.1 Photonendetektionseffizienz

Die PDE gibt die Wahrscheinlichkeit an, dass ein in das photosensitive Gebiet der SPAD einfallendes Photon einen Lawinendurchbruch auslöst, der anschließend von der weiteren Beschaltung erfasst werden kann. Damit ein Photon erkannt wird, muss das Photon in das aktive Gebiet eindringen, dort absorbiert werden, ein EHP erzeugen und der freie Ladungsträger einen Lawinendurchbruch auslösen [8].

2.4.1.1 Reflexion an der Oberfläche

Die erste Hürde, die ein Photon überwinden muss, ist das Eindringen aus der Luft in das Silizium des photoaktiven Elements. Der Reflexionsfaktor berechnet sich aus den Brechungsindizes beider Materialien des Übergangs und ist für schwache Absorption gegeben zu

$$R = \left(\frac{n_{\text{Si}} - n_{\text{Air}}}{n_{\text{Si}} + n_{\text{Air}}} \right)^2 \quad (2.3)$$

mit dem Brechungsindex der Luft n_{Air} und des Siliziums n_{Si} [33]. Im Falle des hier betrachteten Photodetektors befindet sich herstellungsbedingt über dem Silizium der Diode eine dünne Isolationsschicht aus Siliziumdioxid (SiO_2). Die Reflexion bestimmt sich bei senkrechtem Lichteinfall zu

$$R = \frac{n_{\text{SiO}_2}^2 (n_{\text{Air}} - n_{\text{Si}})^2 \cos^2(\varphi) + (n_{\text{Air}} n_{\text{Si}} - n_{\text{SiO}_2}^2)^2 \sin^2(\varphi)}{n_{\text{SiO}_2}^2 (n_{\text{Air}} + n_{\text{Si}})^2 \cos^2(\varphi) + (n_{\text{Air}} n_{\text{Si}} + n_{\text{SiO}_2}^2)^2 \sin^2(\varphi)} \quad (2.4)$$

mit dem Brechungsindex des Siliziumdioxids n_{SiO_2} und der Phasenverschiebung φ in der SiO_2 -Schicht. Diese bestimmt sich aus dem Verhältnis der Schichtdicke zur Wellenlänge im Siliziumdioxid multipliziert mit 2π [34].

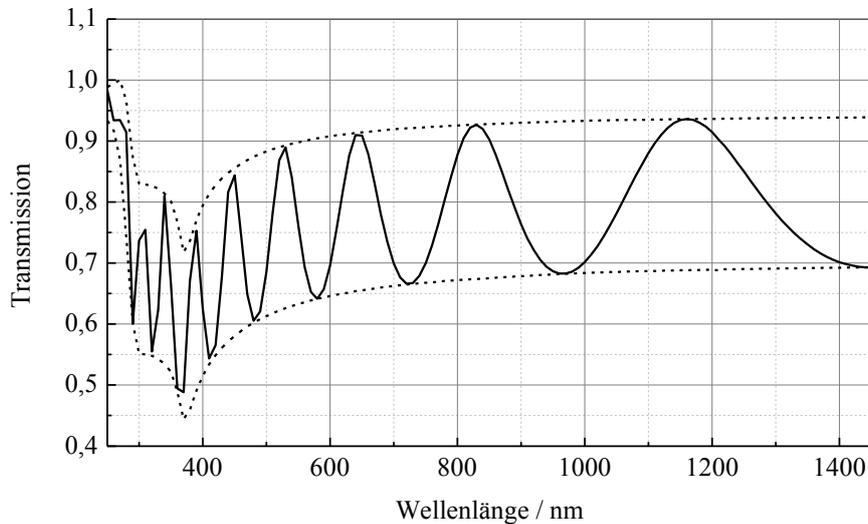


Abbildung 2.6: Transmission zwischen Luft und Silizium aus (2.4)

Durch die Wellenlängenabhängigkeit der Phasenverschiebung zeigt sich in der in Abbildung 2.6 dargestellten Transmission aus (2.4) eine sinusförmige Änderung über der Wellenlänge. Weiterhin zeigen der Maximal- und Minimalwert einen Anstieg der Reflexion hin zu geringeren Wellenlängen, was sich durch den Anstieg der Brechungsindizes von Silizium und Siliziumdioxid in diesem Bereich erklären lässt. Um die Reflexion bei einer bestimmten Wellenlänge gering zu halten, kann die Dicke der SiO_2 -Schicht entsprechend gewählt werden. Eine vollständige Eliminierung der Reflexion lässt jedoch nicht erreichen, da der Brechungsindex von SiO_2 nicht dem notwendigen idealen Index nach

$$n_{\text{ideal}} = \sqrt{n_{\text{Air}} n_{\text{Si}}} \quad (2.5)$$

entspricht [34]. Setzt man (2.5) in (2.4) ein und wählt weiterhin die Schichtdicke gleich einem Viertel der Wellenlänge im SiO_2 , werden beide Summanden in Zähler null und die Reflexion wird vollständig eliminiert. Um trotz des nichtidealen Brechungsindex eine Reduktion der Reflexionsverluste zu erhalten, können zusätzliche Schichten, sogenannte Antireflexionsbeschichtungen, auf die Photodioden aufgebracht werden. Beschichtungen aus mehreren Schichten erlauben zudem eine Reduktion der Reflexion in einem breiten Wellenlängenbereich [34].

2.4.1.2 Absorption im Halbleiter

Damit ein einfallendes Photon ein EHP erzeugen kann, muss dessen Energie E_{Ph} größer als die Bandlückenenergie E_G sein, da nur dann ein Elektron vom Valenz- in das Leitungsband gehoben werden kann. Im Falle von Silizium mit einer Bandlückenenergie von 1,12 eV [30] ergibt sich nach (2.2) eine maximale Wellenlänge von 1108 nm. Der wellenlängenabhängige optische Absorptionskoeffizient α beschreibt über das Lambert-Beersche-Gesetz die Abschwächung einer Strahlung beim Durchlaufen eines Halbleiters entsprechend

$$P(x) = P_0 \exp(-\alpha x) \text{ für } x \geq 0 \quad (2.6)$$

mit der Tiefe im Halbleiter x und der optischen Leistung an der Oberfläche des Halbleiters P_0 [30]. Um die in einem Teilstück dx absorbierte Leistung zu erhalten, wird (2.6) zunächst von eins subtrahiert, woraus sich die ortsabhängige absorbierte Leistung ergibt, und anschließend nach x differenziert. Es folgt

$$P_{\text{abs}}(x) = \frac{d(1 - P(x))}{dx} = P_0 \alpha \exp(-\alpha x) . \quad (2.7)$$

Für die Bestimmung der PDE ist nur der im Teilabschnitt dx absorbierte relative Leistungsanteil von der einfallenden Leistung P_0 relevant, weshalb (2.7) durch P_0 dividiert wird. Als ort- und wellenlängenabhängiger Faktor ergibt sich

$$\frac{P_{\text{abs}}(x)}{P_0} = \alpha \exp(-\alpha x) . \quad (2.8)$$

Der wellenlängenabhängige Absorptionskoeffizient von Silizium, dessen Kehrwert die Absorptionstiefe ist, ist in Abbildung 2.7 dargestellt.

2.4.1.3 Lawinendurchbruch

Ist das elektrische Feld innerhalb der RLZ hoch genug, wird ein freier Ladungsträger derart beschleunigt, dass seine kinetische Energie ausreicht, um durch Stoßionisation weitere EHPs zu erzeugen. Die so generierten Ladungsträger werden wiederum beschleunigt und erzeugen weitere EHPs. Dieser Vorgang sorgt für einen makroskopischen, sich selbst erhaltenden Stromfluss durch die

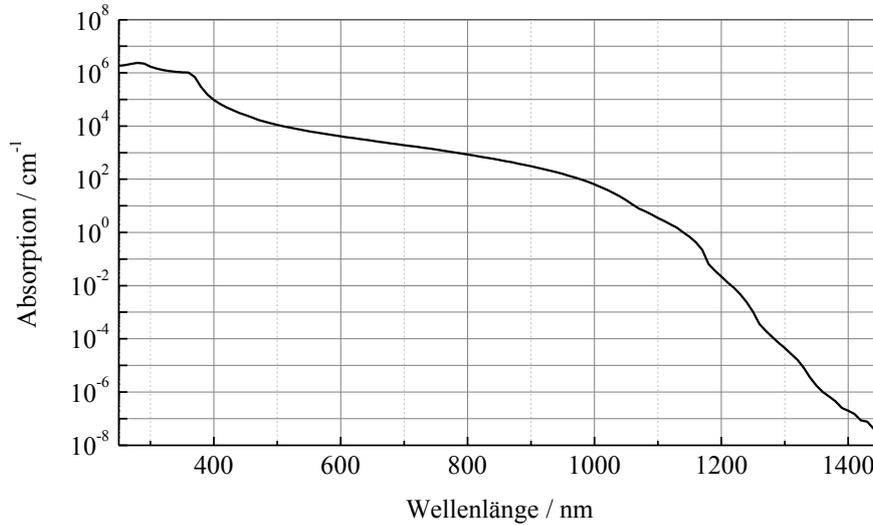


Abbildung 2.7: Wellenlängenabhängiger Absorptionskoeffizient in Silizium

SPAD [8]. Im Weiteren wird eine SPAD betrachtet, die aus dem Übergang eines stark dotierten p-Gebietes und eines schwächer dotierten n-Gebietes besteht. Die Wahrscheinlichkeit, dass entweder ein Elektron oder ein Loch eine Lawine auslöst ist

$$P_{np}(x) = P_n(x) + P_p(x) - P_n P_p(x) \quad (2.9)$$

wobei P_n die Wahrscheinlichkeit für das Auslösen durch ein Elektron und P_p durch ein Loch ist. Betrachtet man ein sich in positive x -Richtung bewegendes Elektron an der Position $x - \Delta x$ in der RLZ, gibt es zwei Möglichkeiten, dass dieses eine Lawine auslöst: Zum einen kann es eine Lawine auslösen, nachdem es den Punkt x durchlaufen hat und zum anderen kann es auf dem Weg von $x - \Delta x$ nach x durch Stoßionisation ein EHP generieren, das eine Lawine auslöst. Für die Wahrscheinlichkeit ergibt sich

$$P_n(x - \Delta x) = P_n(x) + \beta_n \Delta x P_{np}(x) - P_n(x) \beta_n \Delta x P_{np}(x). \quad (2.10)$$

Der zweitgenannte Fall ergibt sich aus der Wahrscheinlichkeit nach (2.9) multipliziert mit der zurückgelegten Strecke Δx und dem Ionisationskoeffizienten für Elektronen β_n [35]. Der Ionisationskoeffizient ist die Anzahl durch Stoßionisation erzeugter EHPs pro Strecke. Die Differentiation von (2.10) liefert



Abbildung 2.8: pn-Übergang bei negativer Vorspannung nach [27]

$$\frac{dP_n}{dx} = -\beta_n(P_n + P_p - P_n P_p)(1 - P_n). \quad (2.11)$$

Aufgrund der Polarität des elektrischen Feldes kann ein Elektron, das sich im n-dotierten Gebiet befindet, nicht in die RLZ eindringen und dort eine Lawine auslösen [36]. Es gilt die Randbedingung

$$P_n(x_{dn}) = 0 \quad (2.12)$$

mit der Ausdehnung der RLZ im n-Gebiet x_{dn} . Eine analoge Betrachtung für Löcher liefert die Gleichung

$$\frac{dP_p}{dx} = \beta_p(P_n + P_p - P_n P_p)(1 - P_p) \quad (2.13)$$

mit dem Ionisationskoeffizienten für Löcher β_p und der Randbedingung

$$P_p(-x_{dp}) = 0 \quad (2.14)$$

mit der Ausdehnung der RLZ im p-Gebiet x_{dp} . Zur Bestimmung von P_n und P_p werden (2.11) und (2.13) durch numerische Integration simultan gelöst. Dabei wird die Randbedingung (2.14) sowie ein beliebiger Wert für $P_n(-x_{dp})$ mit $0 < P_n(-x_{dp}) < 1$ verwendet. In mehreren Iterationen wird $P_n(-x_{dp})$ solange variiert, bis die Randbedingung (2.12) erfüllt ist [35]. Die Ionisationskoeffizienten sind abhängig von der elektrischen Feldstärke und gegeben zu

$$\beta = A \exp\left(-\frac{B}{E}\right) \quad (2.15)$$

mit den Parametern A und B sowie der elektrischen Feldstärke E [27]. Die beiden Parameter sind aus Messungen bekannt. Die elektrische Feldstärke kann aus der Spannung über der RLZ und den Dotierstoffkonzentrationen der p- und n-Schicht berechnet werden. In einem stufenförmigen pn-Übergang ist der Verlauf des elektrischen Feldes der RLZ gegeben durch

$$E(x) = (\Theta(x) - 1) \frac{qN_A}{\varepsilon_0 \varepsilon_{Si}} (x + x_{dp}) + \Theta(x) \frac{qN_D}{\varepsilon_0 \varepsilon_{Si}} (x - x_{dn}) \quad (2.16)$$

mit der Elementarladung q , der Dotierstoffkonzentration N_A bzw. N_D , der Permittivität des Vakuums ε_0 und der relativen Permittivität von Silizium ε_{Si} [27]. $\Theta(x)$ bezeichnet die Heaviside-Funktion. Die Ausdehnung der RLZ im n-Gebiet kann über

$$x_{dn} = \sqrt{\frac{2\varepsilon_0 \varepsilon_{Si} (U_b + U_R)}{q} \frac{N_A}{N_A N_D + N_D^2}} \quad (2.17)$$

mit der Potentialbarriere des pn-Übergangs U_b und der angelegten Spannung in Sperrrichtung U_R bestimmt werden [27]. Analog lässt sich die Ausdehnung im p-Gebiet durch Vertauschen von N_A und N_D berechnen. Aus (2.11) bis (2.17) kann nun der Verlauf der Wahrscheinlichkeit für das Auslösen eines Lawinendurchbruchs in Abhängigkeit der Position x bestimmt werden.

Die bisherige Betrachtung beschreibt die Wahrscheinlichkeit für das Auslösen einer Lawine durch einen freien Ladungsträger innerhalb der RLZ. Wird ein Photon im n- oder p-Gebiet außerhalb der RLZ absorbiert, kann ebenfalls eine Lawine ausgelöst werden. In diesem Fall ist die Wahrscheinlichkeit geringer, da der Minoritätsladungsträger zunächst durch Diffusion in die RLZ gelangen muss. Die Wahrscheinlichkeit, dass ein Ladungsträger die RLZ erreicht, ist bei der in einer Wanne integrierten SPAD nach Abbildung 2.3 linear zur Entfernung von der RLZ mit den beiden Randbedingungen $P_n(-x_p) = 0$ sowie $P_p(x_n) = 0$ mit der Ausdehnung des n- bzw. p-Gebietes x_n bzw. x_p [36].

Um den Verlauf der Wahrscheinlichkeit zu berechnen, wird eine stufenförmiger pn-Übergang mit den Dotierstoffkonzentrationen $N_A = 10^{19} \text{ cm}^{-3}$ und $N_D = 5 \cdot 10^{16} \text{ cm}^{-3}$, eine Spannung der Potentialbarriere von 0,65 V und eine

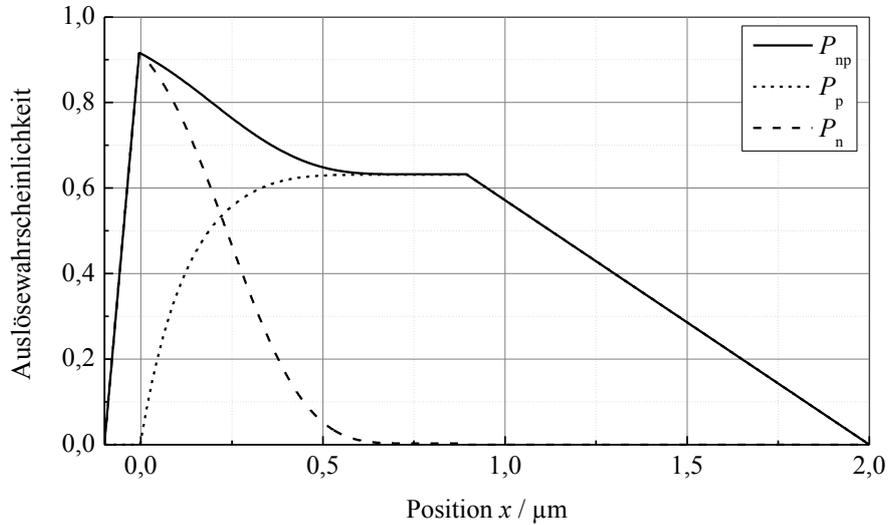


Abbildung 2.9: Auslösewahrscheinlichkeit einer Lawine durch Elektronen, Löcher und beide zusammen in der RLZ sowie den angrenzenden n- und p-Gebieten

Sperrspannung von 30 V angenommen. Die Randbedingung $P_n(x_{dn}) = 0$ ist für $P_n(-x_{dp}) = 0,9156$ erfüllt. Abbildung 2.9 zeigt den Verlauf der Wahrscheinlichkeit für das Auslösen einer Lawine über der Ausdehnung der RLZ von 900 nm mit $x_{dp} = 4,5$ nm und $x_{dn} = 895,5$ nm. Zur Darstellung des linearen Verlaufs außerhalb der RLZ wurde $x_p = 100$ nm und $x_n = 2$ µm angenommen.

2.4.1.4 Photonendetektionseffizienz

Zur Bestimmung der PDE werden im Folgenden die in den Kapiteln 2.4.1.1 bis 2.4.1.3 positions- und wellenlängenabhängigen Größen zusammengefasst. Für die PDE als Funktion der Wellenlänge gilt

$$\eta_{\text{PDE}}(\lambda) = (1 - R(\lambda)) \int_0^{x_p+x_n} \frac{P_{\text{abs}}(x, \lambda)}{P_0} P_{\text{np}}(x - x_p) dx. \quad (2.18)$$

Der berechnete wellenlängenabhängige Verlauf der PDE in Abbildung 2.10 zeigt qualitativ und quantitativ eine gute Übereinstimmung mit den in [9] gemessenen Werten der SPAD nach Abbildung 2.3.

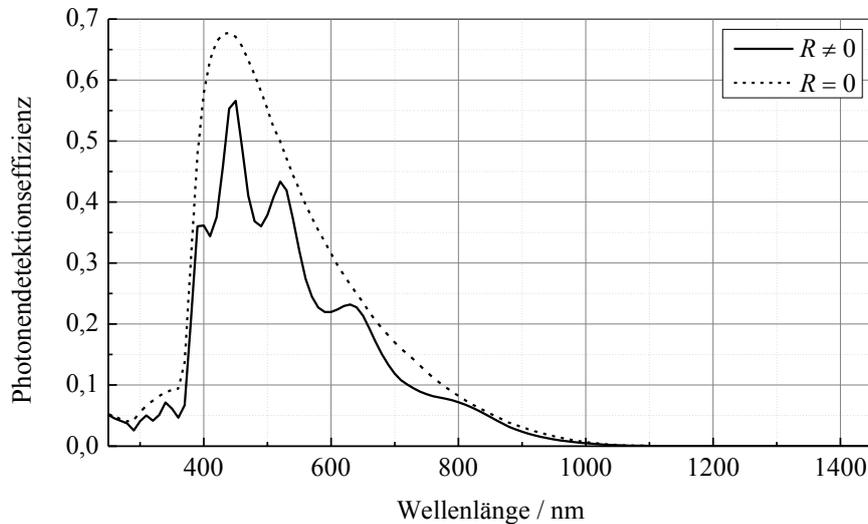


Abbildung 2.10: PDE mit und ohne Oberflächenreflexion

2.4.2 Rauschen

Das Ausgangssignal der SPAD ist ein makroskopischer Strompuls [37], weshalb das Rauschen der Diode und weiterer Beschaltung in Form von zufälligen Schwankungen der Spannungs- und Stromsignale vernachlässigt werden kann. Für die Qualität des Detektors ist entscheidend, dass Lawinendurchbrüche nur durch einfallende Photonen ausgelöst werden. Jedoch gibt es weitere Effekte, welche einen Puls auslösen können. Diese Effekte lassen sich in korreliertes und unkorreliertes Rauschen unterteilen.

2.4.2.1 Unkorreliertes Rauschen

Unkorreliertes Rauschen wird in der sogenannten Dunkelzählrate (Dark Count Rate, DCR) zusammengefasst, da diese Pulse ohne Belichtung auftreten [32]. Die DCR gibt an, wie viele Pulse ohne die Beteiligung eines eingestrahlt Photons pro Zeit ausgelöst werden. Zu den Ursachen von unkorreliertem Rauschen zählen Ladungsträgerdiffusion aus neutralem Gebiet, thermische Generation von EHPs und Band-zu-Band-Tunneln [8].

2.4.2.1.1 Ladungsträgerdiffusion aus neutralem Gebiet

Durch den Konzentrationsunterschied freier Ladungsträger zwischen der RLZ und des angrenzenden p- und n-Gebietes können durch thermische Generation erzeugte freie Minoritätsladungsträger in die RLZ diffundieren. Dort angelangt driften sie aufgrund des vorherrschenden elektrischen Feldes zur gegenüberliegenden Seite der RLZ [27]. Auf ihrem Weg durch die RLZ können diese durch Stoßionisation einen Lawinendurchbruch auslösen. Die flächenbezogene Generationsrate für schmale pn-Übergänge ist gegeben durch

$$G_{\text{diff}} = n_i^2 \left(\frac{D_n}{(x_p - x_{dp})N_A} + \frac{D_p}{(x_n - x_{dn})N_D} \right) \quad (2.19)$$

mit der intrinsischen Ladungsträgerkonzentration n_i und den Diffusionskoeffizienten von Elektronen D_n und Löchern D_p . Im Falle schmaler pn-Übergänge ist die Diffusionslänge der freien Ladungsträger durch die Ausdehnung der Dotiergebiete x_n bzw. x_p begrenzt [27]. Nach [38] ist der Anteil der auf diese Weise generierten Ladungsträger gering gegenüber der thermischen Generation innerhalb der RLZ, welche im Folgenden behandelt wird.

2.4.2.1.2 Thermische Generation von Ladungsträgern

Die Ursachen der thermischen Generation von freien Ladungsträgern innerhalb des pn-Übergangs einer SPAD entsprechen denen des Dunkelstromes bei normalen Photodioden. Ist ein pn-Übergang im thermodynamischen Gleichgewicht, entspricht die Generationsrate der Rekombinationsrate, sodass die Elektronen- und Löcherkonzentrationen im Halbleiter über der Zeit konstant sind [39]. Ist eine externe Spannung angelegt, sorgt das elektrische Feld für eine Trennung der erzeugten EHPs und so zu einem Stromfluss. Im Fall der SPAD kann ein einzelner freier Ladungsträger zum Auslösen des Lawinendurchbruchs, welcher nicht von einem durch ein einfallendes Photon verursachten Durchbruch unterschieden werden kann, führen [8].

Shockley-Read-Hall-Generation ist der dominierende Effekt thermisch generierter Ladungsträger [38]. Dabei erfolgt die Generation über Energiezustände innerhalb der Bandlücke des Halbleiters [40]. Diese Zustände ermöglichen

einem Ladungsträger mit geringerer thermischer Energie den Übergang vom Valenzband in das Leitungsband in zwei Schritten. Diese Art der Generation ist abhängig von der Konzentration der Energiezustände in der Bandlücke und damit von der Qualität des Herstellungsprozesses [8]. Die Rate der thermischen Generation lässt sich mit

$$G_{\text{th}} = \frac{\kappa v_{\text{th}} N_{\text{T}} n_{\text{i}}}{2 \cosh\left(\frac{E_{\text{i}} - E_{\text{T}}}{k_{\text{B}} T}\right)} \quad (2.20)$$

aus dem Einfangquerschnitt für Elektronen und Löcher κ , der thermischen Ladungsträgergeschwindigkeit v_{th} , der Störstellenkonzentration N_{T} , des intrinsischen Energieniveaus E_{i} , dem Energieniveau der Störstellen E_{T} , der Boltzmann-Konstanten k_{B} und der absoluten Temperatur T bestimmen. Da die Energie der Störstellen typischerweise im Bereich des intrinsischen Energieniveaus liegt und die Wahrscheinlichkeit für eine Ladungsträgergeneration am höchsten ist, wenn das Energieniveau der Störstelle in der Mitte zwischen Valenz- und Leitungsband liegt, wird eine effektive Störstellenkonzentration $N_{\text{T,eff}}$ definiert [27]. Aus (2.20) folgt damit

$$G_{\text{th}} = \frac{1}{2} \kappa v_{\text{th}} N_{\text{T,eff}} n_{\text{i}}. \quad (2.21)$$

Um die Temperaturabhängigkeit der Rate nach (2.21) zu verdeutlichen, wird die thermische Ladungsträgergeschwindigkeit

$$v_{\text{th}} = \sqrt{\frac{3k_{\text{B}} T}{m^*}} \quad (2.22)$$

mit der effektiven Ladungsträgermasse m^* [27] und die intrinsische Ladungsträgerkonzentration

$$n_{\text{i}} = \sqrt{N_{\text{C}} N_{\text{V}}} \exp\left(-\frac{E_{\text{G}}}{2k_{\text{B}} T}\right) = C T^{\frac{3}{2}} \exp\left(-\frac{E_{\text{G}}}{2k_{\text{B}} T}\right) \quad (2.23)$$

mit der effektiven Zustandsdichte im Leitungs- und Valenzband N_C bzw. N_V und der Bandlückenenergie E_G in (2.21) eingesetzt. Der Faktor C ist eine materialabhängige Konstante und hat für Silizium den Wert $2,34 \cdot 10^{17} \text{ cm}^{-3} \text{ K}^{-3/2}$ [27]. Für die thermische Ladungsträgergeneration ergibt sich

$$G_{\text{th}} = \frac{1}{2} \kappa \sqrt{\frac{3k_B}{m^*}} N_{T,\text{eff}} C T^2 \exp\left(-\frac{E_G}{2k_B T}\right). \quad (2.24)$$

Die Gleichung nach (2.24) liefert die Generationsrate pro Volumen. Für die tatsächliche Rate an thermisch erzeugten Ladungsträgern muss diese mit dem Volumen der Raumladungszone multipliziert werden [27].

2.4.2.1.3 Band-zu-Band-Tunneln

Die Quantenmechanik lehrt, dass Elektronen in eine endlich hohe Potentialbarriere eindringen können. Liegt die maximale Eindringtiefe über der Breite der Barriere, kann das Elektron diese mit einer bestimmten Wahrscheinlichkeit vollständig durchdringen. Dieser Effekt wird als Tunneln bezeichnet und lässt sich aus der Schrödingergleichung ableiten [41]. In SPADs kann dies zu freien Ladungsträgern und so zum Auslösen von Lawinendurchbrüchen führen. Da die Tunnelwahrscheinlichkeit mit sinkender Breite der Barriere steigt und diese mit steigender Sperrspannung sinkt, ist dieser Effekt bei hohen Spannungen besonders ausgeprägt. Das direkte Band-zu-Band-Tunneln steigt ab einer elektrischen Feldstärke von ca. 1 MV/cm stark an und wird relevant [27].

Analog zur thermischen Generation von Ladungsträgern kann es durch Störstellen innerhalb der Bandlücke bereits bei geringeren Feldstärken zu signifikanter Ladungsträgergeneration kommen. Die Störstellen teilen die Potentialbarriere in zwei Barrieren geringerer Breite auf und erhöhen die Tunnelwahrscheinlichkeit exponentiell [38]. Dies wird als störstellenunterstütztes Tunneln (Trap-Assisted-Tunneling, TAT) bezeichnet. Die Generationsrate durch TAT ergibt sich nach [42] durch Multiplikation der thermischen Generation nach (2.24) mit dem Faktor

$$\Gamma = 2\sqrt{3}\pi \frac{|E|}{F_\Gamma} \exp\left(\frac{E}{F_\Gamma}\right)^2 \quad (2.25)$$

mit dem lokalen elektrischen Feld E und mit

$$F_\Gamma = \frac{2\pi\sqrt{24m^*(k_B T)^3}}{qh}, \quad (2.26)$$

wobei h das Plancksche Wirkungsquantum ist.

2.4.2.2 Korreliertes Rauschen

Korreliertes Rauschen beschreibt das Auslösen von Pulsen aufgrund von zeitlich oder räumlich getrennten Lawinendurchbrüchen. Zu dieser Art des Rauschens zählen Afterpulsing, welches durch einen zeitlich getrennten Puls verursacht wird, und Crosstalk, wo die räumliche Trennung innerhalb eines Arrays von SPADs eine Rolle spielt [8].

2.4.2.2.1 Afterpulsing

Während eines Lawinendurchbruchs in einer SPAD können freie Ladungsträger die im Silizium vorhandenen Störstellen besetzen und diese nach einer endlichen Lebensdauer wieder verlassen. Erfolgt das Verlassen der Störstellen in einer Phase, in welcher ein erneuter Durchbruch der SPAD möglich ist, kann hierdurch ein zweiter Puls ohne das Vorhandensein eines Photons ausgelöst werden [8]. Die Anzahl besetzter Störstellen steigt mit der Ladung, die während eines Durchbruchs durch die Diode fließt. Zur Reduktion von Afterpulsing kann im ersten Schritt durch möglichst schnelles Löschen eines Durchbruchs die fließende Ladung auf ein Minimum gesenkt werden. Eine weitere Reduktion lässt sich dadurch erreichen, dass die SPAD nach dem Löschen der Lawine eine gewisse Zeit inaktiv gehalten wird. Ladungsträger, die während dieser Zeit die Störstellen verlassen, können keinen weiteren Durchbruch auslösen [43]. Da die Lebensdauer besetzter Störstellen einer Exponentialverteilung folgt, kann bereits eine vergleichsweise geringe Zeit der Inaktivität zu einer signifikanten Reduktion von Afterpulsing führen [38]. Durch die Zeit der Inaktivität reduziert

sich weiterhin die maximale Zählrate der SPAD, weshalb ein geeigneter Kompromiss zwischen der Reduktion von Afterpulsing und der Zählrate eingegangen werden muss [44].

2.4.2.2.2 Crosstalk

Durch Interferenzen zwischen SPADs in einem Array kann der Lawinendurchbruch in einer SPAD zum Auslösen des Durchbruchs in einer benachbarten Diode führen. Die Ursachen für das Übersprechen können elektrischer oder optischer Natur sein. Bei elektrischem Übersprechen wandern während des Durchbruchs erzeugte freie Ladungsträger von der einen SPAD zur anderen und lösen dort einen Durchbruch aus. Durch geeignete Barrieren zwischen den SPADs lassen sich die freien Ladungsträger einsammeln und das elektrische Übersprechen reduzieren. Optisches Übersprechen beschreibt das Auslösen eines Durchbruchs durch in einer benachbarten SPAD erzeugte Photonen [8]. Während eines Durchbruchs werden Photonen hauptsächlich durch strahlende Rekombination und Zustandsübergänge im Valenz- bzw. Leitungsband erzeugt [45]. Einfluss auf das Übersprechen haben die Größe der Dioden, der Abstand zwischen Aktivgebieten und deren Volumen. Analog zum Afterpulsing wird Übersprechen auch durch eine Minimierung der während eines Durchbruchs fließenden Ladung reduziert [8].

2.4.3 Zeitlicher Jitter

Der zeitliche Jitter beschreibt die Schwankung der Zeit zwischen dem Einfallen eines Photons in das photoaktive Gebiet der SPAD und der Flanke des Ausgangspulses. Ursächlich für diese Schwankungen sind die Tiefe der Absorption des Photons im Silizium sowie das statistische Verhalten beim Aufbau des Lawinendurchbruchs [8]. Zur Minimierung des Jitters kann der Grenzwert des nachfolgenden Komparators möglichst gering gewählt werden. In SPAD Arrays erhöht dies jedoch elektrischen Crosstalk, weshalb auf Kosten des Jitters höhere Grenzwerte verwendet werden [46].

2.4.3.1 Absorptionstiefe im Silizium

Die tiefenabhängige Absorptionswahrscheinlichkeit eines einfallenden Photons im Silizium folgt einem exponentiellen Verlauf. Es besteht somit die Möglichkeit, dass ein Photon außerhalb der RLZ absorbiert wird. Dabei kann der Minoritätsladungsträger die RLZ erreichen und eine Lawine auslösen. In solch einem Fall ist die Zeitverzögerung zwischen dem Eintreffen des Photons und dem Auslösen der Lawine länger, als wenn das Photon innerhalb der RLZ absorbiert wird. Durch diesen Effekt kommt es zu einem exponentiellen Abfall des Zeitverhaltens [31]. Im Fall der im Rahmen von MiSPiA entworfenen SPAD wird dieser exponentielle Abfall durch das Anlegen des Anodenpotentials an das Substrat reduziert. Dadurch wird die parasitäre Diode zwischen n-Wanne und Substrat stark negativ vorgespannt, wodurch sich eine breite RLZ ausbildet und die neutrale Zone der n-Wanne verkleinert wird [9].

Auch wenn ein EHP innerhalb der RLZ generiert wird, sind statistische Schwankungen der Zeit, bis ein Ladungsträger den Bereich des hohen elektrischen Feldes erreicht, vorhanden. Weiterhin muss ein Ladungsträger zunächst eine gewisse Strecke im elektrischen Feld zurücklegen, bis dieser ausreichend Energie für das Erzeugen eines weiteren EHPs durch Stoßionisation hat. Zudem unterliegt die Ionisation selbst einer statistischen Wahrscheinlichkeit nach

$$P_{\text{ion}}(x) = \beta \exp(\beta(x - d_{\text{ion}})) \text{ für } x > d_{\text{ion}} \quad (2.27)$$

mit dem makroskopischen Ionisationskoeffizienten β , der Position x und der notwendigen Distanz zum Erreichen ausreichender Energie d_{ion} . Erst ab einer Anzahl von etwa 100 Ladungsträgern innerhalb der RLZ mitteln sich diese statistischen Schwankungen aus und der Stromanstieg folgt einem exponentiellen Verlauf [31].

2.4.3.2 Aufbau des Lawinendurchbruchs

Der Aufbau des Lawinendurchbruchs kann in zwei Phasen betrachtet werden. In der ersten Phase bildet sich ausgehend vom Einfallspunkt des Photons durch Stoßionisation ein exponentiell steigender Strom aus, da in dieser Phase der leitende Bereich zu klein ist um das elektrische Feld der RLZ zu beeinflussen. In

der zweiten Phase breitet sich der Durchbruch aus und reduziert das elektrische Feld, wodurch der Strom sättigt. Die Ausbreitung kann durch laterale Diffusion in benachbarte Gebiete oder durch Emission von Sekundärphotonen erfolgen. Die Zeit dafür schwankt mit der Position des eingefallenen Photons und der statistischen Verteilung der lateralen Diffusion [46].

Die in [46] vorgestellte Kenngröße zur Beschreibung des durch den Aufbau des Lawinendurchbruchs verursachten Jitters ergibt sich aus dem Produkt der Stromdichte J der durchflossenen Fläche A und der Ausbreitungsgeschwindigkeit der selbigen Fläche nach

$$PT_{\text{FoM}} \propto \frac{dI}{dt} = J \frac{dA}{dt}. \quad (2.28)$$

Die Stromdichte bestimmt sich aus dem Quotienten der Überschussspannung U_{ex} und dem spezifischen Widerstand ρ . Die Ausbreitungsgeschwindigkeit der stromdurchflossenen Fläche ergibt sich zu

$$\frac{dA}{dt} = 4\pi \sqrt{\frac{D}{\tau}} \Delta d \quad (2.29)$$

mit dem effektiven Diffusionskoeffizienten D , der Zeitkonstanten τ und der Dicke des stromdurchflossenen Rings des aktiven Gebietes Δd . Die Ringdicke berücksichtigt, dass im statischen Fall der Strom aufgrund des geringeren Widerstandes hauptsächlich im Randbereich des leitenden Gebietes, welches als kreisförmig angenommen wird, fließt [46]. Da die Kenngröße nur der Vergleichbarkeit dient, wird hier nur die Proportionalität entsprechend

$$PT_{\text{FoM}} = \frac{1}{\rho\sqrt{\tau}} \quad (2.30)$$

berücksichtigt.

2.5 Quenching und Reset

Um nach dem Erfassen eines Photons ein nächstes Photon detektieren zu können, muss die Spannung über der SPAD unter die Durchbruchspannung gesenkt und anschließend wieder angehoben werden. Die für diese Aktion benötigte Zeit wird als Totzeit bezeichnet und begrenzt die maximale Rate, mit der einfallende Photonen gezählt werden können [8]. Weiterhin lässt sich durch die Begrenzung des fließenden Stromes Afterpulsing, Crosstalk und Erhitzung der Diode reduzieren [47]. Für das Löschen des Lawinenstromes und das anschließende Wiederaufladen lassen sich verschiedene Schaltungskonzepte unterscheiden.

2.5.1 Passives Quenching und Reset

Ein hochohmiger Widerstand in Reihe zur SPAD stellt die einfachste Form einer solchen Schaltung dar [44]. Wird durch das Einfallen eines Photons ein Lawinendurchbruch ausgelöst (wellenförmiger Pfeil in Abbildung 2.11 (b)), steigt der Strom auf einen Spitzenwert an, der sich aus dem Quotienten von Überschussspannung und Widerstand der Diode ergibt. Dieser Strom entlädt die parasitäre Kapazität der Anode, wodurch die Spannung am Knoten S exponentiell ansteigt. Der Stromfluss durch eine SPAD nach dem Auslösen einer Lawine ist bis zu einer bestimmten Höhe selbsterhaltend [43], dieser Wert wird als Latchup-Strom I_{lu} bezeichnet. Da die Spannung über der Diode nicht unter die Durchbruchspannung fällt, stellt sich ein statischer Stromfluss von

$$I_{SPAD} \approx \frac{U_{ex}}{R_Q} \quad (2.31)$$

mit der Überschussspannung U_{ex} und dem Reihenwiderstand R_Q durch die Diode ein [44]. Zur Löschung des Lawinendurchbruchs muss dieser Strom den Latchup-Strom unterschreiten. Für die Wahl des Widerstandswertes gilt daher

$$R_Q > \frac{U_{ex}}{I_{lu}}. \quad (2.32)$$

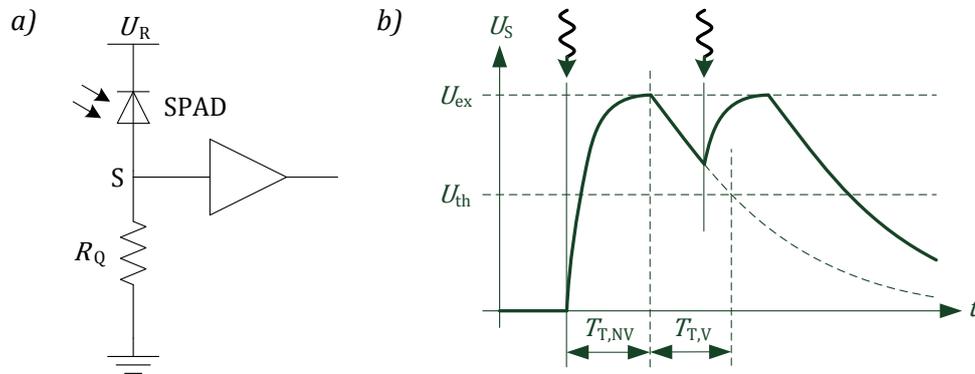


Abbildung 2.11: Beschaltung der SPAD (a) nach [44] und Spannungsverlauf (b) für passives Quenching nach [48]

Weiterhin fällt der Strom exponentiell ab, weshalb der Widerstand für eine definierte Löschzeit größer als der durch (2.32) festgelegte Minimalwert gewählt werden sollte [43]. Nach dem Löschen wird die parasitäre Kapazität der Diode über den Widerstand wieder auf den ursprünglichen Spannungswert aufgeladen. Aufgrund des hohen Widerstandswertes benötigt das Aufladen eine relativ lange Zeit, wodurch die Leistungsfähigkeit des Sensors begrenzt wird [37]. Fällt, wie in Abbildung 2.11 (b) gezeigt, während des Aufladens ein Photon in den aktiven Bereich der SPAD ein, kann ein erneuter Durchbruch ausgelöst werden. Passiert dies, bevor die Spannung an Knoten S die Schwellenspannung der Auswerteschaltung U_{th} unterschreitet, wird dieses Photon nicht registriert und die Totzeit verlängert sich, da der Strom erneut gelöscht werden muss. Daraus resultiert eine Abnahme der Zählrate für steigende Photonenraten gemäß

$$r_{T,P} = \frac{r \exp(-rT_{T,V})}{1 + rT_{T,NV}} \quad (2.33)$$

mit der idealen Photonendetectionsrate in totzeitfreien Fall r , dem Anteil der Totzeit, innerhalb dessen keine Verlängerung der Totzeit stattfindet, $T_{T,NV}$ und dem Anteil, in welchem eine Verlängerung erfolgt, $T_{T,V}$ [48].

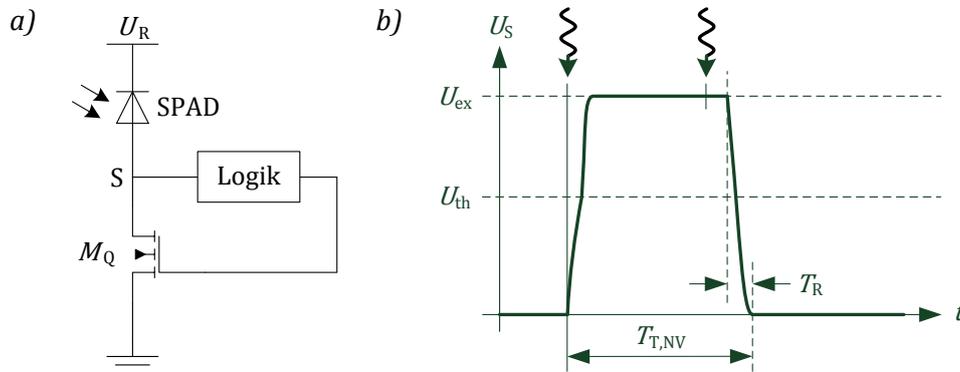


Abbildung 2.12: Beschaltung der SPAD (a) nach [44] und Spannungsverlauf (b) für aktives Quenching nach [48]

2.5.2 Aktives Quenching und Reset

Die logische Weiterentwicklung ist ein variabler Widerstand, der während des Löschens des Lawinenstromes einen hohen Wert und während des Aufladens einen geringen Wert aufweist. Zur Reduktion der Totzeit wurden aktive Schaltungen für Quenching und Reset entwickelt. Diese verwenden statt des passiven Widerstandes ein aktives Bauelement, in der Regel einen Transistor, zur Steuerung des Vorgangs [47]. Abbildung 2.12 (a) zeigt das Beschaltungsschema für aktives Quenching. Im Ruhezustand ist der Transistor M_Q leitend und die Spannung an Knoten S liegt auf Masse. Wird ein Photon detektiert, registriert die Logik einen Spannungsanstieg an S und sperrt den Transistor, wodurch der Stromfluss gelöscht wird. Nach einer festgelegten Zeit, der sogenannten Hold-Off-Zeit zur Reduktion von Afterpulsing, wird der Transistor wieder in den leitenden Zustand geschaltet. Damit kann die Kapazität der SPAD über einen geringen Widerstand und so deutlich schneller als im passiven Fall aufgeladen werden [44]. Fällt, wie in Abbildung 2.12 (b) gezeigt, während der Hold-Off-Phase ein Photon ein, führt dies zu keiner Verlängerung der Totzeit, da die SPAD in diesem Zeitbereich inaktiv ist. Während der Ladephase hingegen löst ein einfallendes Photon die SPAD aus. Dadurch, dass der Transistor in dieser Zeit leitend ist, kommt es zu einem dauernden Stromfluss bis zum Ende der Ladephase, wo das Photon von der Schaltung registriert und der Stromfluss gestoppt

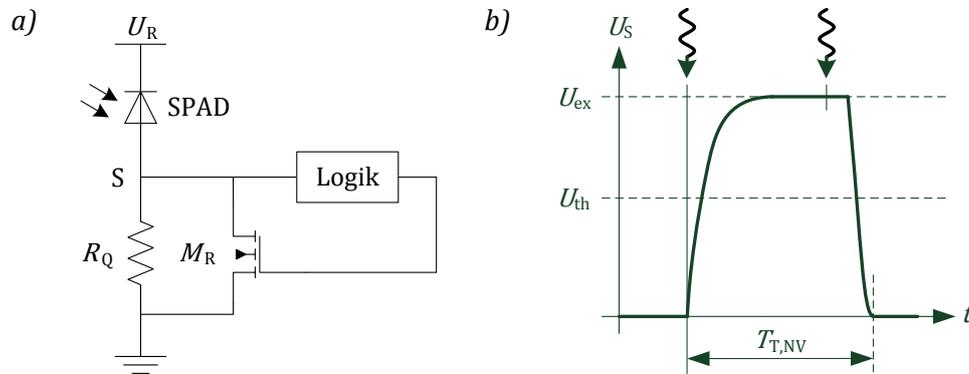


Abbildung 2.13: Beschaltung der SPAD (a) nach [44] und Spannungsverlauf (b) für kombiniertes Quenching

wird. Zur Reduktion der Leistungsaufnahme muss die aktive Ladephase daher so kurz wie möglich sein [47]. Für die gemessene Zählrate gilt

$$r_{T,A} = \frac{r}{1 + rT_{T,NV}}, \quad (2.34)$$

was sich aus (2.33) für $T_{T,V} = 0$ ergibt. Aufgrund der Verwendung des Transistors als variabler Widerstand wird diese Beschaltung auch als Variable Load Quenching Circuit bezeichnet. Ein Nachteil des vorgestellten Konzeptes ist die durch die Logik verursachte Verzögerung beim Löschen des Lawinenstromes, was die fließende Ladung und damit die Wahrscheinlichkeit für Afterpulsing erhöht [44]. Dem kann durch eine entsprechende Hold-Off-Zeit entgegenge wirkt werden.

2.5.3 Passives Quenching und aktiver Reset

Um die Vorteile beider Verfahren zu erhalten, können diese miteinander kombiniert werden [44]. Durch passives Quenching lässt sich die fließende Ladung durch die SPAD und damit Afterpulsing reduzieren, während aktives Aufladen eine Reduktion der Totzeit und der Wahrscheinlichkeit für das Einfallen von Photonen während des Aufladens erlaubt. Abbildung 2.13 zeigt das Beschaltungsschema sowie den Spannungsverlauf für diesen Fall. Hierbei ist der Transistor M_R in Ruhezustand gesperrt. Beim Auslösen einer Lawine wird der Strom

unmittelbar durch den Widerstand R_Q gelöscht. Nach verstrichener Hold-Off-Zeit wird der Transistor leitend geschaltet und die SPAD damit über einen geringen Widerstand aufgeladen. Nach dem Aufladen wird M_R wieder in den Sperrzustand versetzt, um beim nächsten Photon ein schnelles Löschen des Lawinenstromes zu ermöglichen.

2.5.4 Time-Gating

Eine weitere Variante neben dem aktiven und passiven Quenching ist das Time-Gating. Bei dieser Methode wird die SPAD für eine gewisse Dauer aktiviert und anschließend wieder deaktiviert. Da die SPAD nur zu Beginn eines jeden Zeitfensters aufgeladen wird, kann in jedem aktiven Zeitintervall nur ein Photon erfasst werden. Ein Vorteil des Verfahrens ist die Möglichkeit einer gemeinsamen Logik für das Laden mehrerer SPADs. Zum Ende eines aktiven Fensters wird geprüft, ob ein Photon registriert wurde. Das Löschen des Lawinenstromes muss in jeder SPAD individuell erfolgen, um einen dauerhaften Stromfluss über die Dauer des Fensters zu vermeiden [44].

3 Verfahren der optischen Entfernungsmessung

In diesem Abschnitt werden die verschiedenen Verfahren zur Entfernungsbestimmung mittels optischer Wellen vorgestellt sowie auf die Besonderheiten bei der Verwendung von SPADs hingewiesen. Es lassen sich die drei Verfahren Interferometrie, Triangulation und Laufzeitmessung zur Distanzbestimmung mit elektromagnetischen Wellen im Bereich des sichtbaren sowie infraroten und ultravioletten Spektrum nutzen [4].

3.1 Interferometrie

Die Interferometrie gehört zur Gruppe der aktiven Distanzbestimmung und nutzt die Lichtlaufzeit zur Bestimmung der Distanz zwischen Sensor und Messobjekt. Dabei beruht die Messung auf der Interferenz der reflektierten mit einer Referenzwelle, weshalb die Genauigkeit wesentlich von der Kohärenz der Lichtquelle abhängt. Da die Methode auf der Wellenlänge des Lichtes beruht, ist Interferometrie nur für Entfernungen bis zu wenigen Zentimetern möglich [4].

Abbildung 3.1 zeigt das Prinzip eines Michelson Interferometers [49] zur Bestimmung der Distanz d . Ein Teil des Laserstrahls wird am halbdurchlässigen Spiegel reflektiert und nach oben abgelenkt. Dieser wird anschließend an S_1 reflektiert und durch den Spiegel zum Sensor transmittiert. Der zweite Teil des

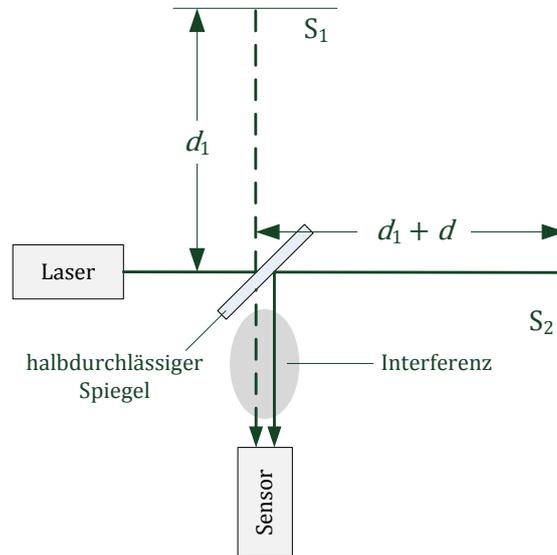


Abbildung 3.1: Prinzip des Michelson Interferometers nach [49]

Laserstrahl wird durch den Spiegel transmittiert und nach der Reflexion an S_2 vom Spiegel nach unten reflektiert. Dort interferieren beide Strahlen miteinander. Der zurückgelegte optische Weg beider Strahlen unterscheidet sich dabei um $2d$ und entscheidet über konstruktive oder destruktive Interferenz. In realen Interferometern muss zusätzlich der durch den Spiegel verursachte Gangunterschied berücksichtigt werden. Die messbare Intensität des Interferenzsignals lässt sich durch Addition beider Wellen bestimmen. Unter der Annahme einer gleichen Aufteilung der Amplitude durch den Spiegel ergibt sich

$$I(d) = \frac{1}{\lambda} \int_0^{\lambda} (\sin(kx) + \sin(k(x + 2d)))^2 dx = 1 + \cos(2kd) \quad (3.1)$$

mit der Wellenlänge λ und der Kreiswellenzahl k . Dabei zeigt sich, dass eine eindeutige Zuordnung zwischen der Intensität I und d nur für $0 \leq d < \lambda/2$ möglich ist. Die messbare Distanz ist bei diesem Verfahren durch die Wellenlänge der verwendeten Strahlung bestimmt. Im Allgemeinen kann das Prinzip der Interferometrie auch mit sehr langwelliger Strahlung wie Radiowellen angewandt werden, wodurch große Entfernungen erfasst werden können.

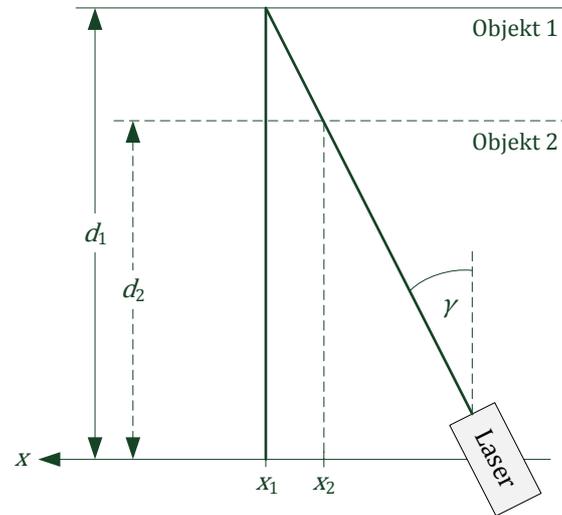


Abbildung 3.2: Prinzip der aktiven Triangulation nach [50]

3.2 Triangulation

Das Verfahren der Triangulation beruht auf den schon vor über 2000 Jahren von den Griechen verwendeten Prinzipien der Trigonometrie [50] und kann aktiv oder passiv angewendet werden. Bei letzterem wird ein Objekt mittels zwei Kameras, welche räumlich voneinander getrennt sind, aufgenommen. Aus dem Abstand der beiden Kameras und dem Winkel lässt sich die Objektdistanz mittels geometrischer Berechnung bestimmen. Hierfür müssen korrespondierende Punkte in beiden Bildern einander zugeordnet werden, was hohe Anforderungen an die nachfolgende Berechnung stellt. Vorteile dieser Methode sind die hohe Auflösung und die Nichtnotwendigkeit einer speziellen Lichtquelle [51]. Bei der aktiven Triangulation lassen sich verschiedene Ausprägungen nach der verwendeten Lichtquelle unterscheiden. Beim Laserscannen wird mittels eines Lasers ein Lichtpunkt auf das zu messende Objekt projiziert und von einem Sensor unter einem bekannten Winkel erfasst. Je nach Entfernung des Objektes ändert sich die Position des Lichtpunktes auf dem Sensor, woraus sich mit bekanntem Abstand und Winkel zwischen Lichtquelle und Sensor der Abstand

berechnen lässt [52]. Abbildung 3.2 zeigt das Prinzip dieser Art der Triangulation. Für die Objektdistanz d gilt

$$d = x \tan(\gamma) \quad (3.2)$$

mit dem bekannten Winkel zwischen Laser und Sensorfläche γ und dem Nullpunkt der x -Achse an der Position des Lasers. Bei anderen Varianten wird eine Linie oder ein Muster auf die Zielszene projiziert [53]. Vorteile dieser Varianten sind, dass nicht jeder Punkt einzeln abgetastet werden muss und so die Zeit zur Aufnahme des Bildes und die Anforderungen an die Mechanik zur Ablenkung des Laserstrahls reduziert werden [50]. Ein die Zielszene vollständig abdeckendes Muster macht eine Mechanik sogar überflüssig, erhöht aber die Anforderungen an die nachfolgende Berechnung der Tiefeninformation.

3.3 Laufzeitmessung

Zur aktiven Messung größerer Distanzen werden auf CCD/CMOS-Photodioden basierende Sensoren verwendet. Bei diesem Verfahren wird die Laufzeit des Lichtes von der Aussendung über die Reflexion bis zum Empfang erfasst [4]. Die Welleneigenschaft des Lichtes spielt bei den betrachteten Entfernungen keine Rolle und wird für die Betrachtung des grundsätzlichen Funktionsprinzips vernachlässigt. Es lässt sich das direkte und indirekte Verfahren unterscheiden.

3.3.1 Direkte Laufzeitmessung

Beim direkten Verfahren zur Laufzeitmessung wird mit der Aussendung eines Lichtpulses eine elektronische Stoppuhr in Sensor gestartet und mit dem Empfang des reflektierten Pulses wieder gestoppt. Aus der gemessenen Laufzeit T_{TOF} berechnet sich die Objektdistanz d nach

$$d = \frac{c}{2} T_{\text{TOF}} \quad (3.3)$$

mit der Lichtgeschwindigkeit c [4]. Der Faktor $1/2$ ergibt sich aus der Tatsache, dass das Licht die Entfernung zwischen dem Sensor und Messobjekt zweimal

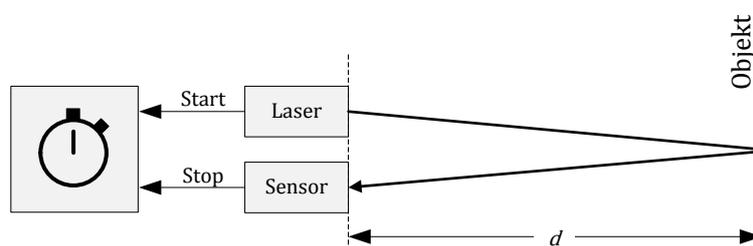


Abbildung 3.3: Prinzip der direkten Laufzeitmessung nach [4]

zurücklegt. Für eine präzise Distanzmessung ist eine hohe zeitliche Auflösung der Stoppuhr erforderlich [6]. Um die Resistenz gegen Hintergrundlicht zu erhöhen, werden in der Regel mehrere Zeitstempel in einem Histogramm gesammelt, aus welchem anschließend die Laufzeit ermittelt wird.

3.3.2 Indirekte Laufzeitmessung

Beim indirekten Verfahren wird intensitätsmoduliertes Licht ausgesendet und die Laufzeit aus der zeitlichen Verschiebung zwischen gesendetem und empfangenem Signal bestimmt. Grundsätzlich werden zwei Modulationsarten unterschieden: Pulsmodulation (PM) und Sinusmodulation (Continuous Wave, CW) [4]. In beiden Varianten erfolgt die Bestimmung der zeitlichen Verschiebung durch Erfassung der reflektierten Lichtmenge in mehreren Zeitfenstern, was im Falle von SPADs durch Zählung der einfallenden Photonen erfolgt.

3.3.2.1 Pulsmodulation

Bei dieser Variante der indirekten Laufzeitmessung wird ein Lichtpuls mit der Dauer T_p ausgesendet und das reflektierte Signal in mehreren Zeitfenstern integriert. Für die Anordnung der Zeitfenster existieren verschiedene Varianten, wobei die hier betrachtete Variante mit drei Fenstern die geringste Varianz verspricht [1], [25]. Wie in Abbildung 3.4 dargestellt, sind die Fenster bündig nacheinander angeordnet und entsprechen in ihrer Dauer der Pulsbreite.

Das erste Zeitfenster beginnt mit dem Aussenden des Lichtpulses, sodass in diesem in Abhängigkeit der Entfernung des Zielobjektes eine Teilmenge des reflektierten Laserpulses integriert wird. Das zweite Zeitfenster integriert den

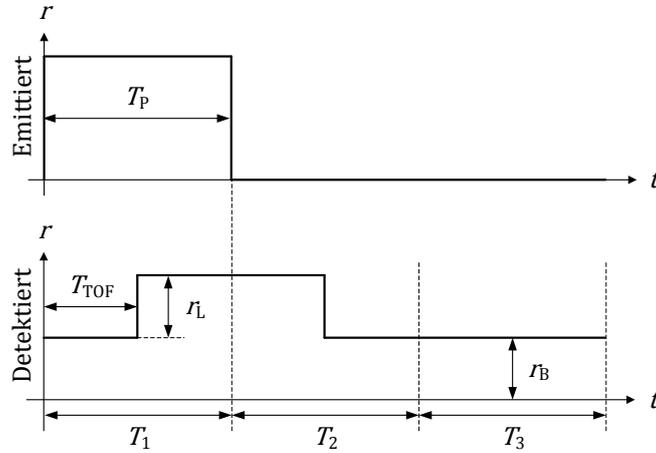


Abbildung 3.4: Indirekte Laufzeitmessung mit Pulsmodulation nach [25]

verbleibenden Teil des reflektierten Pulses. Das dritte Fenster dient der Erfassung der Hintergrundstrahlung [25]. Die Erwartungswerte N_i der Zeitfenster T_i ergeben sich durch Integration über die Ereignisraten des reflektierten Pulses r_L und des Hintergrundes r_B sowie mit der Laufzeit T_{TOF} zu

$$\begin{aligned} N_1 &= T_P r_B + (T_P - T_{\text{TOF}}) r_L, \\ N_2 &= T_P r_B + T_{\text{TOF}} r_L \\ \text{und } N_3 &= T_P r_B. \end{aligned} \quad (3.4)$$

Die maximal messbare Entfernung ist durch die Pulsdauer T_P entsprechend

$$d_{\text{max}} = \frac{c}{2} T_P \quad (3.5)$$

begrenzt [51]. Dies garantiert, dass der reflektierte Puls spätestens nach $2T_P$ wieder empfangen und vollständig von den ersten beiden Fenstern erfasst wird. Unter idealen Bedingungen, d. h. ohne Hintergrundstrahlung, lässt sich aus den in den ersten beiden Fenstern T_1 und T_2 integrierten Lichtmengen N_1 und N_2 die Entfernung nach

$$d = \frac{N_2}{N_1 + N_2} d_{\text{max}} \quad (3.6)$$

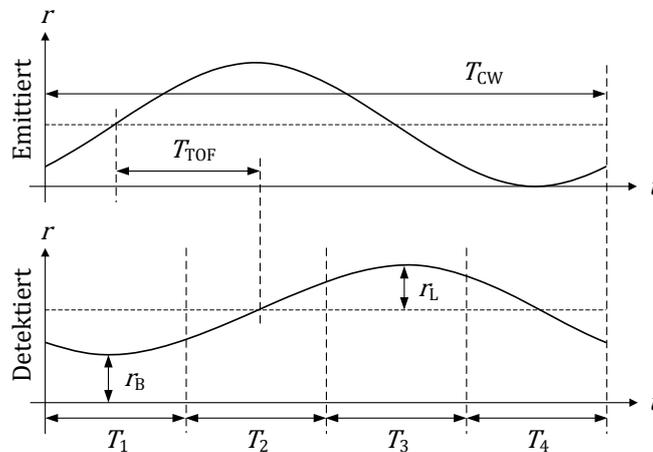


Abbildung 3.5: Indirekte Laufzeitmessung mit Sinusmodulation nach [25]

berechnen. Da in realen Anwendungsfällen in der Regel Hintergrundstrahlung vorhanden ist, muss die im dritten Fenster erfasste Hintergrundstrahlung zunächst von den ersten beiden Fenstern subtrahiert werden. Damit folgt

$$d = \frac{N_2 - N_3}{N_1 + N_2 - 2N_3} d_{\max} \quad (3.7)$$

für die Berechnung der Distanz [25]. Der Wert des dritten Fensters N_3 kann zudem zur Erzeugung eines Grauwertbildes verwendet werden.

3.3.2.2 Sinusmodulation

Bei dieser Art der indirekten Laufzeitmessung wird ein sinusmoduliertes Signal ausgesendet und das reflektierte Signal durch Integration der einfallenden Photonen in vier Zeitfenstern erfasst. Wie in Abbildung 3.5 dargestellt, sind diese gleichmäßig auf die Periode verteilt. Durch Integration des empfangenen Signals ergeben sich die Erwartungswerte N_i der Zeitfenster T_i zu

$$N_i = T_W(r_L + r_B) + \frac{T_{CW}r_L}{\pi} \sin\left(\frac{\pi T_W}{T_{CW}}\right) \sin\left(\frac{\pi(i-1)}{2} - 2\pi\frac{T_{TOF}}{T_{CW}}\right) \quad (3.8)$$

mit der Fensterbreite T_W , der Periodendauer der emittierten Welle T_{CW} und der Lichtlaufzeit T_{TOF} . Aus den gezählten Photonen der einzelnen Fenster N_i lassen

sich anschließend Mittelwert, Amplitude und Phase des empfangenen Signals rekonstruieren. Den Mittelwert des Signals erhält man durch einfache Mittelung der vier gemessenen Werte entsprechend [54]

$$r_L + r_B = \frac{1}{4T_W} (N_1 + N_2 + N_3 + N_4). \quad (3.9)$$

Für die Amplitude werden die um π verschobenen Werte voneinander subtrahiert, wodurch sich der Mittelwert aufhebt. Durch anschließendes Quadrieren der Differenzen und Summenbildung entfallen die Sinusfunktionen. Die Gleichung ergibt sich damit zu [54]

$$r_L = \frac{\pi}{2T_{CW}} \sin^{-1} \left(\frac{\pi T_W}{T_{CW}} \right) \sqrt{(N_4 - N_2)^2 + (N_1 - N_3)^2}. \quad (3.10)$$

Die Differenz aus Mittelwert und Amplitude liefert die Hintergrundintensität, welche zur Erzeugung eines Grauwertbildes herangezogen werden kann. Die Amplitude ist weiterhin ein Maß für die Güte des reflektierten Signals.

Zur Bestimmung der Laufzeit und damit der Entfernung eines Zielobjektes in der Szene ist die Phase des reflektierten Signals relevant, welche sich über

$$\varphi = 2\pi \frac{T_{TOF}}{T_{CW}} \quad (3.11)$$

direkt in die Laufzeit umrechnen lässt. Hierfür werden ebenfalls die Differenzen der um π verschobenen Werte gebildet sowie anschließend deren Quotient bestimmt. Dieser stellt den Tangens der Phase dar, aus welchem sich im Weiteren die Distanz entsprechend

$$d = \frac{c}{2} T_{TOF} = d_{\max} \frac{\varphi}{2\pi} = \frac{d_{\max}}{2\pi} \operatorname{atan} \left(\frac{N_4 - N_2}{N_1 - N_3} \right) \quad (3.12)$$

berechnen lässt [25], [54]. Die maximal messbare Entfernung ist dabei durch die halbe Modulationswellenlänge gegeben. Dies ist der Tatsache geschuldet, dass eine größere Entfernung eine nicht eindeutige Phasenverschiebung über 2π hervorruft.

4 Statistik laufzeitbasierter Messverfahren

In diesem Kapitel werden die im vorhergehenden Kapitel vorgestellten Verfahren zur laufzeitbasierten Distanzmessung genauer untersucht. Die direkte Laufzeitmessung beruht auf der Bestimmung der Zeit zwischen Emission und Detektion eines Laserpulses. Die Grenzen dieses Verfahrens werden durch die Photonenstatistik, die zeitlichen Jitter der einzelnen Komponenten des Systems sowie die Hintergrundstrahlung bestimmt. Dagegen beruht das indirekte Messverfahren auf der Zählung von Photonen, weshalb die Statistik solcher Zählprozesse unter dem Einfluss der SPAD-typischen Totzeit betrachtet wird. Diese Erkenntnisse werden im Weiteren zur Bestimmung der zu erwartenden Varianzen und eines Vergleichs der beiden indirekten Methoden zur Laufzeitmessung mit SPADs herangezogen.

4.1 Photonenstatistik

Die Bestimmung einer Distanz über die Laufzeitmessung mittels SPADs beruht auf der Erfassung einzelner Photonen. Aus diesem Grund liegt es nahe, zunächst die zeitliche Verteilung von Photonen im Licht zu betrachten. Es lassen sich drei Typen von Licht unterscheiden, welche jeweils eine andere statistische Verteilung der Photonen aufweisen und für die Laufzeitmessung relevant sind:

Kohärentes Licht, welchem ideales Laserlicht entspricht und als Strahlungsquelle im Messsystem eingesetzt wird, chaotisches Licht, wie es von bestimmten künstlichen Lichtquellen emittiert wird, und thermisches Licht, wie es thermische Lichtquellen wie die Sonne emittieren [49].

4.1.1 Kohärentes Licht

Im Falle von kohärentem Licht weist das elektromagnetische Feld eine konstante Phase, Frequenz und Amplitude auf. Da die Intensität als das Quadrat der Amplitude des elektrischen Feldes definiert ist, ist diese ebenfalls konstant. Aufgrund der Quantisierung des Lichtes ist die mittlere Anzahl an Photonen, die während eines Zeitfensters der Dauer T_W empfangen werden, gegeben durch

$$\mu = rT_W \quad (4.1)$$

mit dem Photonenfluss r , welcher sich aus der Intensität, der Fläche und der Photonenenergie berechnen lässt. Zur Bestimmung der Photonenstatistik wird das Zeitfenster in n Teilstücke der Dauer T_W/n zerlegt, sodass die Wahrscheinlichkeit zwei oder mehr Photonen in einem Teilzeitfenster zu empfangen vernachlässigt werden kann. Die Wahrscheinlichkeit bei n Messungen x Photonen zu empfangen lässt sich nun mittels der Binomialverteilung

$$P_B(x, n, p) = \frac{n!}{x! (n-x)!} p^x (1-p)^{n-x} \quad (4.2)$$

berechnen, wobei die Wahrscheinlichkeit durch $p = rT_W/n$ gegeben ist [49]. Die Analogie zum Binomialtheorem garantiert dabei, dass die Verteilung normiert ist [55]. Die Binomialverteilung gibt somit die Verteilung der Photonenzahl bei einer Messung über eine Dauer von T_W an. Der Erwartungswert μ beschreibt die mittlere Anzahl an erfolgreichen Versuchen x für unendlich viele Wiederholungen des Experiments mit je n Versuchen. Er berechnet sich aus dem Produkt der Wahrscheinlichkeit $P_B(x, n, p)$ mit der jeweiligen Zahl der erfolgreichen Versuche x integriert über alle möglichen x entsprechend

$$\mu_B = \sum_{x=0}^n x \frac{n!}{x! (n-x)!} p^x (1-p)^{n-x} = np. \quad (4.3)$$

Wie erwartet zeigt (4.3), dass der Erwartungswert von x dem Produkt aus der Erfolgswahrscheinlichkeit p und der Anzahl der Versuche n entspricht und mit (4.1) übereinstimmt. Die Varianz ist definiert als die quadrierte Abweichung des Erwartungswertes vom Messwert multipliziert mit dessen Wahrscheinlichkeit integriert über alle x und berechnet sich nach [55] zu

$$\sigma_B^2 = \sum_{x=0}^n (x - \mu)^2 \frac{n!}{x! (n-x)!} p^x (1-p)^{n-x} = np(1-p). \quad (4.4)$$

Die Varianz ist proportional zur Versuchsanzahl n und erreicht ihren Maximalwert bei einer Erfolgswahrscheinlichkeit p von 50 %. Damit, wie bisher vorausgesetzt, die Wahrscheinlichkeit zwei oder mehr Photonen innerhalb eines Fensters der Dauer T_W/n zu empfangen tatsächlich gegen null strebt, muss der Grenzwert $n \rightarrow \infty$ betrachtet werden. Damit folgt für die Wahrscheinlichkeit $p \rightarrow 0$ und für die Verteilung

$$\lim_{p \rightarrow 0} P_B(x, n, p) = P_P(x, \mu_P) = \frac{\mu_P^x}{x!} \exp(-\mu_P). \quad (4.5)$$

Für diesen Grenzfall geht die Binomialverteilung in die Poisson-Verteilung über. Die Poisson-Verteilung erlaubt eine Beschreibung der statistischen Verteilung von Szenarien, in welchen nur die Zahl erfolgreicher Versuche x bzw. deren Erwartungswert bekannt ist. Auch für Zählprozesse, bei denen die Zahl der erfassten Ereignisse klein gegenüber der Zahl maximal erfassbarer Ereignisse ist, kann die Verteilung angewendet werden [55]. Der Erwartungswert der Poisson-Verteilung ist bereits in (4.5) enthalten und entspricht μ_P . Die Varianz σ_P^2 entspricht ebenfalls dem Erwartungswert [49]. Abbildung 4.1 zeigt einen Vergleich zwischen der Poisson- und der Binomialverteilung für einen Erwartungswert von 10 sowie eine Erfolgswahrscheinlichkeit der Binomialverteilung von 10 % und 50 %. Hierbei zeigt sich, wie die Binomialverteilung mit sinkender Erfolgswahrscheinlichkeit in die Poisson-Verteilung übergeht.

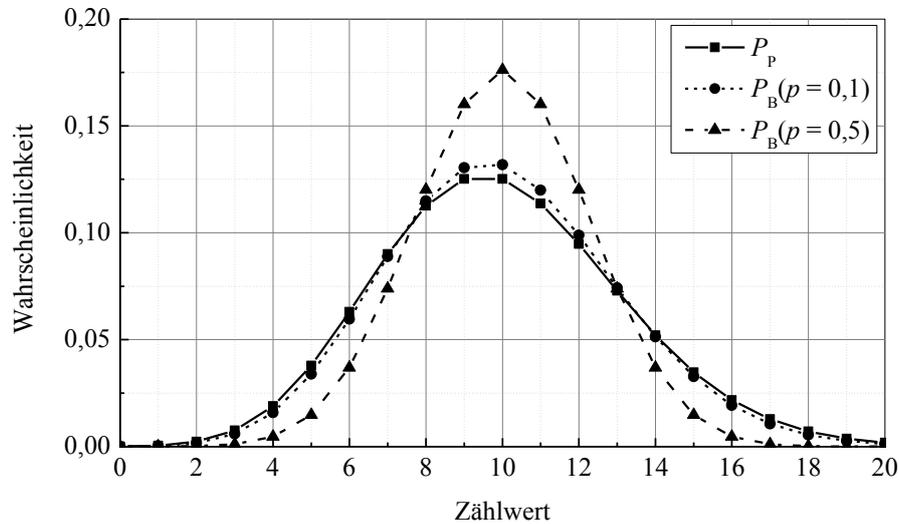


Abbildung 4.1: Poisson- und Binomialverteilung für $\mu = 10$, $p = 0,1$ und $p = 0,5$

Besonders für das direkte Verfahren zur Laufzeitmessung spielt neben der Verteilung der Zählwerte auch die zeitliche Ankunft der Photonen eine wichtige Rolle. Dazu zählen die Verteilung der Zwischenankunftszeiten (Interarrival Time, IAT) sowie der Zeit bis zur Ankunft des ersten Photons nach Beginn der Beobachtung zu einem beliebigen Zeitpunkt (First-Photon Arrival Time, FPAT). Unter der Annahme einer konstanten Intensität, ist die Wahrscheinlichkeit ein Photon zu empfangen ebenfalls eine konstante Funktion. Beginnt die Beobachtung des Photonenstromes zum Zeitpunkt $t = 0$, ist die Wahrscheinlichkeit, zu diesem Zeitpunkt ein Photon zu empfangen, gegeben durch den Photonenfluss r mal eines infinitesimalen Zeitintervalls dt . Zu einem späteren Zeitpunkt $t_1 > 0$ muss zusätzlich die Wahrscheinlichkeit berücksichtigt werden, dass im Zeitintervall $[0, t_1[$ kein Photon empfangen wurde. Für beliebiges $t \geq 0$ gilt

$$P_1(t) = r \left(1 - \int_0^t P_1(\tau) d\tau \right). \quad (4.6)$$

Diese Differentialgleichung hat die Lösung

$$P_1(t) = r \exp(-rt), \quad (4.7)$$

was der Wahrscheinlichkeitsdichtefunktion (Probability Density Function, PDF) der FPAT sowie der IAT für $t \geq 0$ entspricht. Dieser Zusammenhang lässt sich auch aus der Poisson-Verteilung ableiten: Nach dieser ist die Wahrscheinlichkeit, mindestens ein Photon in einem Zeitfenster der Dauer T_W zu empfangen, gegeben durch $1 - \exp(-rT_W)$ [56]. Dies entspricht dem Integral über die PDF der Photonenankunftszeit $P_1(t)$ über $t = [0, T_W]$. Damit gilt

$$P_P(x > 0, rT_W) = 1 - \exp(-rT_W) = \int_0^{T_W} P_1(t) dt. \quad (4.8)$$

Werden beide Seiten der Gleichung differenziert, ergibt sich für $P_1(t)$ die Verteilung nach (4.7). Auch die Erlang-Verteilung, welche später in Abschnitt 4.2.1.1 im Zusammenhang mit den indirekten Messverfahren behandelt wird, liefert dieses Resultat.

4.1.2 Thermisches Licht

Das Spektrum der emittierten elektromagnetischen Strahlung eines schwarzen Körpers mit der Temperatur T ist gegeben durch das Plancksche Strahlungsgesetz entsprechend

$$S(\nu, T) d\nu = \frac{8\pi h \nu^3}{c^3} \frac{1}{\exp\left(\frac{h\nu}{k_B T}\right) - 1} d\nu \quad (4.9)$$

mit der Frequenz ν , dem Planckschen Wirkungsquantum h , der Lichtgeschwindigkeit c und der Boltzmann-Konstanten k_B . In einer einzelnen Schwingungsmode mit konstanter Frequenz ν sind die Energiezustände entsprechend

$$E_n = \left(n + \frac{1}{2}\right) h\nu \text{ für } n \in \mathbb{N} \quad (4.10)$$

quantisiert. Mittels der Boltzmann-Statistik lässt sich die Wahrscheinlichkeit, dass genau n Photonen in der Mode enthalten sind, berechnen. Diese ergibt sich aus dem Verhältnis der Wahrscheinlichkeit des Energiezustandes n zu der Gesamtwahrscheinlichkeit aller Energiezustände zu

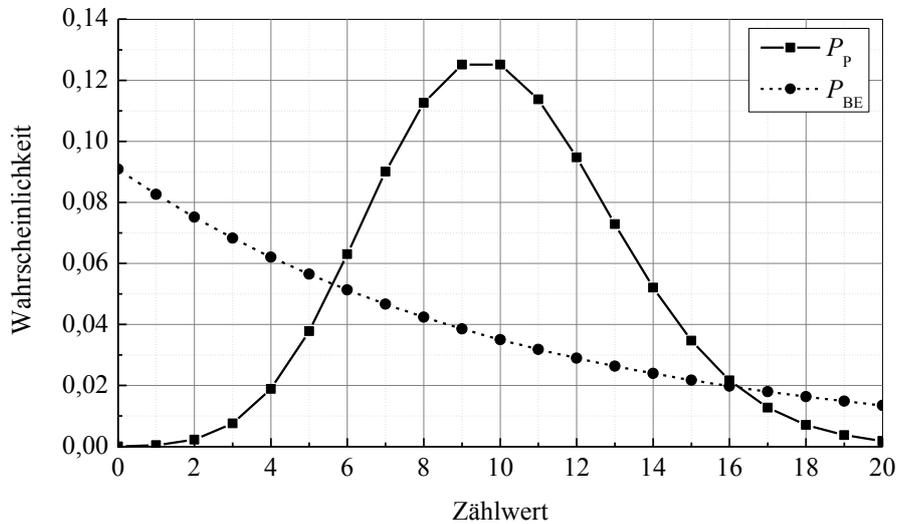


Abbildung 4.2: Poisson- und Bose-Einstein-Verteilung für $\mu = 10$

$$P_n = \frac{\exp\left(-\frac{E_n}{k_B T}\right)}{\sum_n \exp\left(-\frac{E_n}{k_B T}\right)} \text{ für } n \in \mathbb{N}. \quad (4.11)$$

Aus dieser Wahrscheinlichkeit lässt sich die Verteilung berechnen zu

$$P_{BE}(x, \mu) = \frac{1}{\mu + 1} \left(\frac{\mu}{\mu + 1}\right)^x. \quad (4.12)$$

Diese Verteilung ist die sogenannte Bose-Einstein-Verteilung [49], deren Verlauf eine exponentielle Form aufweist. Aufgrund der exponentiellen Verteilung der Besetzungswahrscheinlichkeiten der Energiezustände in (4.11) ist dieses Resultat wenig überraschend. Abbildung 4.2 zeigt den Vergleich zwischen Poisson- und Bose-Einstein-Verteilung für einen Erwartungswert von $\mu = 10$. Dabei ist die deutlich höhere Varianz der Bose-Einstein-Verteilung erkennbar. Es gilt hervorzuheben, dass dieses Resultat nur für eine einzelne Mode thermischen Lichtes gilt. In realen Anwendungen besteht das Licht jedoch aus einem Kontinuum an Moden, wodurch sich die Varianz auf μ reduziert und die Zahl der erfassten Photonen wieder der Poisson-Verteilung entspricht [49].

4.1.3 Chaotisches Licht

Chaotisches Licht, auch als teilkohärentes Licht bezeichnet, stellt ein Mittel zwischen kohärentem und thermischem Licht dar. In diesem Fall ist die Varianz abhängig von der Fluktuation der Intensität und liegt zwischen den beiden Extremen gegeben durch kohärentes und thermisches Licht. Die Fluktuation der Intensität spielt jedoch nur eine Rolle, wenn das Zeitfenster im Bereich der Kohärenzzeit liegt, welche gegeben ist durch

$$\tau_c = \frac{1}{\Delta\nu} \approx \frac{\lambda^2}{c\Delta\lambda} \quad (4.13)$$

wobei $\Delta\nu$ der Breite des Frequenzspektrums entspricht. Für eine Mittenwellenlänge von $\lambda = 905$ nm und eine Bandbreite von $\Delta\lambda = 20$ nm liegt diese Zeit bei 137 fs und damit weit unterhalb der in der betrachteten Anwendung verwendeten Zeitfenster. Aus diesem Grund kann chaotisches Licht ebenfalls als Poissonverteilt angenommen werden [49].

4.2 Indirekte Laufzeitmessung

Die indirekte Messung von Lichtlaufzeiten beruht auf der Integration des reflektierten Lichtes in mehreren Zeitfenstern. In herkömmlichen Photodioden erzeugt ein einfallendes Photon ein EHP, das durch das elektrische Feld der RLZ getrennt wird. Die so generierte Ladung wird in einer Kapazität gesammelt und auf diese Weise in ein zur Anzahl der eingefallenen Photonen proportionales Spannungssignal umgewandelt. Aufgrund der geringen Ladung bei schwachen Lichtverhältnissen ist das Spannungssignal ebenfalls niedrig. Das resultierende geringe SNR wirkt sich negativ auf die Präzision der Laufzeitmessung aus. Bei der Verwendung von SPADs wird für jedes einfallende Photon ein makroskopischer Strompuls erzeugt, wodurch die Weiterverarbeitung digital erfolgt und das Spannungs- bzw. Stromrauschen der nachfolgenden Steuer- und Auswerteelektronik vernachlässigt werden kann. Jedoch wirken sich im Falle von SPADs nicht durch Signalphotonen ausgelöste Lawinendurchbrüche negativ auf die Präzision aus. Bei geringen Belichtungsstärken spielt zudem die statistische

Verteilung der einfallenden Photonen eine wichtige Rolle. Im Folgenden wird die Auswirkung der Photonenstatistik auf die Präzision der indirekten Laufzeitmessung untersucht.

4.2.1 Statistik

Nach der Detektion eines Photons durch eine SPAD muss der Lawinenstrom gelöscht und die SPAD wieder auf ihren ursprünglichen Spannungswert aufgeladen werden, bevor ein weiteres Photon erfasst werden kann. Aufgrund dieser Tatsache ist eine SPAD nicht in der Lage alle einfallenden Photonen innerhalb eines Zeitfensters zu zählen. Um aus den gezählten Photonen die eingefallenen Photonen zu bestimmen, wird im Weiteren die Auswirkung der Totzeit auf die Poisson-Verteilung untersucht.

Entsprechend [56] kann ein Zählprozess unter den Gesichtspunkten der Erneuerungstheorie betrachtet werden. Sie besagt, dass ein System nach einem Ereignis sich wieder genau wie zu Beginn des Versuchs verhält. Für Zählprozesse mit Totzeit wird der Ursprungszustand definiert als der, in welchem der Detektor bereit zum Zählen eines Ereignisses ist [57]. Befindet sich der Detektor in diesem Zustand und ein Ereignis wird erfasst, vergeht unter der Annahme einer nicht verlängerbaren Totzeit die Totzeit T_T , bis sich das System wieder im Ursprungszustand befindet und damit wie zu Beginn verhält. Bei der Zählung von Photonen ohne Totzeit handelt es sich um einen Poisson-Prozess, d. h. die Varianz σ_p^2 entspricht dem Erwartungswert μ_p , welcher sich für eine konstante Photonendetektionsrate r nach (4.1) berechnet. Ist der Erneuerungsprozess persistent, d. h. das Ereignis tritt mit einer Wahrscheinlichkeit von 100 % auf, und die Erneuerungszeit hat einen endlichen Erwartungswert μ_ε und eine endliche Varianz σ_ε^2 , ist die Anzahl der Ereignisse in der Zeit T_W asymptotisch normalverteilt mit Erwartungswert und Varianz [56] entsprechend

$$\mu_T = \frac{T_W}{\mu_\varepsilon} \quad \text{und} \quad \sigma_T^2 = \frac{T_W \sigma_\varepsilon^2}{\mu_\varepsilon^3}. \quad (4.14)$$

Für eine konstante Rate r bestimmt sich der Erwartungswert der Erneuerungszeit ohne Betrachtung der Totzeit T_T zu $\mu_\varepsilon = 1/r$ und die Varianz zu $\sigma_\varepsilon^2 = 1/r$.

Zur Berücksichtigung einer konstanten, nicht verlängerbaren Totzeit wird diese zum Erwartungswert der Erneuerungszeit addiert [57]. Es folgt damit

$$\mu_{\varepsilon} = \frac{1}{r} + T_{\text{T}} = \frac{1 + rT_{\text{T}}}{r}. \quad (4.15)$$

Die Varianz bleibt dagegen unverändert. Eingesetzt in (4.14) folgt für Erwartungswert und Varianz der Ereignisse und damit der Zahl gezählter Photonen

$$\mu_{\text{T}} = \frac{rT_{\text{W}}}{1 + rT_{\text{T}}} \quad \text{und} \quad \sigma_{\text{T}}^2 = \frac{rT_{\text{W}}}{(1 + rT_{\text{T}})^3}. \quad (4.16)$$

Erwartungsgemäß strebt der Erwartungswert μ_{T} für $r \rightarrow \infty$ gegen $T_{\text{W}}/T_{\text{T}}$, da sobald der Detektor bereit für das Erfassen einen weiteren Photons ist, auch ein Photon einfällt. Entsprechend strebt für diesen Grenzfall die Varianz gegen null. Prinzipiell sinkt die Varianz mit steigender Totzeit, da diese für hohe Werte die statistische Verteilung der detektierten Photonen dominiert. Bildet man die relative Varianz aus dem Quotienten von Erwartungswert und Varianz nach

$$\frac{\sigma_{\text{T}}^2}{\mu_{\text{T}}} = \frac{1}{(1 + rT_{\text{T}})^2} \quad (4.17)$$

zeigt sich, dass die relative Varianz mit steigender Totzeit und Einfallrate abnimmt. Entspricht die Totzeit bspw. der durchschnittlichen Wartezeit zwischen der Detektion zweier Photonen $rT_{\text{T}} = 1$, reduziert sich die relative Varianz um den Faktor vier gegenüber der Poisson-Verteilung.

Der berechnete Erwartungswert und die Varianz nach (4.16) gelten nur, wenn die Photonendetektionsrate r bereits eine gewisse Zeit lang vor der Messung konstant war und während der Messung konstant bleibt. Dies ist in der betrachteten Anwendung jedoch nicht der Fall, sodass die PDFs der einzelnen Photonen betrachtet werden müssen. Zunächst stellt sich die Frage nach der Verteilung der FPAT. Anders als im totzeitfreien Fall ist diese nicht identisch mit der Verteilung der IAT, da sich die SPAD gerade in Totzeit oder nicht in Totzeit befinden kann. Die SPAD befindet sich zum Zeitpunkt $t = 0$ in Totzeit, wenn im Zeitintervall $t = [-T_{\text{T}}, 0]$ ein Photon detektiert wurde. In diesem Fall ist die Zeit bis zur erneuten Einsatzbereitschaft der SPAD relevant. Die Wahrscheinlichkeit,

dass die Totzeit zum Zeitpunkt t_1 endet und wieder ein Photon detektiert werden kann, entspricht der Wahrscheinlichkeit, dass zum Zeitpunkt $t_1 - T_T$ ein Photon detektiert wurde. Damit ergibt sich für die Verteilung der FPAT

$$P_1(t) = \int_0^{T_T} P_{\text{Ph}}(\tau - T_T) P_{\text{IA}}(t - \tau, \tau) d\tau + \left(1 - \int_{-T_T}^0 P_{\text{Ph}}(t) dt \right) P_{\text{IA}}(t, 0) \quad (4.18)$$

mit der Verteilung der Photonen IAT ohne Einfluss der Totzeit $P_{\text{IA}}(t, t_0)$ und der Photonenempfangswahrscheinlichkeit $P_{\text{Ph}}(t)$. Der erste Summand gilt für den Fall, dass die SPAD zu Beginn in Totzeit ist und entspricht der Faltung zwischen der Detektionsrate zum Zeitpunkt $t - T_T$ und der Verteilung der IAT. Da nur der Zeitraum $t = [-T_T, 0]$ von $P_{\text{Ph}}(t)$ relevant ist, wird das Faltungsintegral auf diesen Bereich begrenzt. Der zweite Summand deckt den Fall ab, die SPAD empfangsbereit vorzufinden. Hierbei ist die Wartezeit bis zum nächsten Photon durch die Verteilung der Photonenankunft in totzeitfreien Fall gegeben. Analog zu (4.6) ergibt sich diese durch Lösen von

$$P_{\text{IA}}(t, t_0) = r(t + t_0) \left(1 - \int_0^t P_{\text{IA}}(\tau, t_0) d\tau \right) \text{ für } t \geq 0. \quad (4.19)$$

Im Falle einer zeitabhängigen Photonenrate $r(t)$ ist die Verteilung zusätzlich eine Funktion des Startzeitpunktes t_0 , da die IAT eine Funktion der zukünftigen Photonenrate ist. Da die zeitliche Verschiebung im Faltungsintegral erfolgt, ist (4.19) stets definiert für $t \geq 0$. Ist die Verteilung des ersten Photons bekannt, lässt sich die Verteilung des k -ten Photons über

$$P_k(t) = \int_0^{\infty} P_{k-1}(\tau) P_{\text{IA}}(t - \tau - T_T, \tau + T_T) d\tau \quad (4.20)$$

für $k = 2, 3, \dots$ und mit der Verteilung der IAT unter dem Einfluss von Totzeit $P_{\text{IA}}(t - T_T, t_0 + T_T)$ berechnen.

4.2.1.1 Sonderfall $T_T = 0$ und $r(t) = r$

Ohne Totzeit und bei konstanter Photonendetektionsrate r geht (4.18) in (4.7) über, womit (4.20) in die Erlang-Verteilung

$$P_k(t) = P_E(t, k) = \frac{r^k t^{k-1}}{(k-1)!} \exp(-rt) \text{ für } t \geq 0 \quad (4.21)$$

übergeht. Eine Summation von P_E für $k \in \mathbb{N}$ liefert eine konstante Photonendetektionswahrscheinlichkeit von r .

4.2.1.2 Sonderfall $T_T \neq 0$ und $r(t) = r$

Bei vorhandener Totzeit muss diese nach jeder Detektion eines Photons zunächst verstreichen, bevor ein weiteres Photon detektiert werden kann. Folglich kann bei einer Photonendetektion zum Zeitpunkt $t = 0$ kein Photon im Intervall $t = [0, T_T]$ erfasst werden. Für die PDF der IAT nach (4.19) ergibt sich

$$P_{IA}(t, t_0) = P_{IA}(t) = r \exp(-rt) \text{ für } t \geq 0. \quad (4.22)$$

Unter der Annahme einer konstanten Photonendetektionsrate r , reduziert sich die effektive Detektionsrate unter dem Einfluss von Totzeit auf

$$r_T = \frac{r}{1 + rT_T}, \quad (4.23)$$

was sich aus dem reduzierten Erwartungswert in (4.16) ergibt. Setzt man diese konstante Detektionsrate in (4.18) für die Photonendetektionswahrscheinlichkeit P_{Ph} sowie die Verteilung der IAT nach (4.22) ein, ergibt sich für die Verteilung der Zeit bis zur ersten Photonendetektion

$$P_1(t) = \int_0^{\min(t, T_T)} r_T r \exp(-r(t - \tau)) d\tau \quad (4.24)$$

$$+ (1 - r_T T_T) r \exp(-rt) \text{ für } t \geq 0,$$

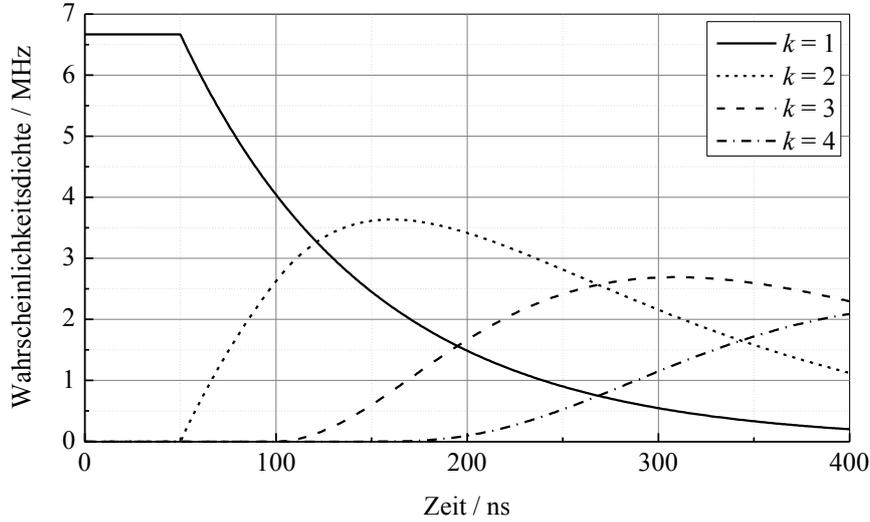


Abbildung 4.3: PDF der Photonenankunftszeit für $k = 1$ bis 4

wobei die Begrenzung des Faltungsintegrals zusätzlich die Beschränkung von $P_{1A}(t)$ auf $t \geq 0$ berücksichtigt. Eine Lösung der Gleichung liefert

$$P_1(t) = \begin{cases} r_T & \text{für } 0 \leq t < T_T \\ r_T \exp(-r(t - T_T)) & \text{für } t \geq T_T \end{cases} \quad (4.25)$$

für die PDF der ersten Photonendetektion. Mit diesem Resultat und (4.20) lässt sich die PDF des k -ten Photons zu

$$P_k = \begin{cases} r_T \left(1 - e^{-rt_k} \sum_{n=-1}^{k-2} \frac{r^n t_k^n}{n!} \right) & \text{für } 0 \leq t_k < T_T \\ r_T \left(e^{-rt_{k+1}} \sum_{n=0}^{k-1} \frac{r^n t_{k+1}^n}{n!} - e^{-rt_k} \sum_{n=-1}^{k-2} \frac{r^n t_k^n}{n!} \right) & \text{für } t_k \geq T_T \end{cases} \quad (4.26)$$

mit $t_k = t - (k - 1)T_T$ berechnen. Die Variable t_k berücksichtigt die zeitliche Verschiebung der IAT um die Totzeit. Weiterhin ist die Fakultät über die Gammafunktion entsprechend $\Gamma(n + 1) = n!$ definiert, damit strebt der Betrag der Fakultät einer negativen ganzen Zahl gegen unendlich und die entsprechenden Summanden entfallen. Die Summation von P_k für $k \in \mathbb{N}$ liefert wie erwartet eine konstante Wahrscheinlichkeitsdichte von r_T nach (4.23). Die nach

(4.26) berechneten PDFs der Photonenankunftszeiten für $k = 1$ bis 4 sind in Abbildung 4.3 für eine Totzeit von 50 ns und eine Photonendetektionsrate von 10 MHz dargestellt. Das Plateau für $k = 1$ im Intervall $t = [0, T_T]$ resultiert aus der reduzierten Empfangswahrscheinlichkeit durch die Möglichkeit die SPAD in Totzeit vorzufinden.

4.2.2 Varianz

Für die Berechnung der Varianzen der beiden indirekten Verfahren zur Laufzeitbestimmung werden die Abkürzungen $r_{LB} = r_L + r_B$ mit der Photonenrate des Lasersignals r_L und des Hintergrundes r_B sowie $r_{L,T}$, $r_{B,T}$ und $r_{LB,T}$ für die unter Totzeiteinfluss reduzierten Raten entsprechend (4.23) eingeführt. Dabei gilt zu beachten, dass $r_{LB,T} \neq r_{L,T} + r_{B,T}$ für $T_T > 0$ gilt. Da die gemessene Distanz in der indirekten Laufzeitmessung eine Funktion der Zählwerte ist, wird die Fehlerfortpflanzung der Zählwerte in den einzelnen Zeitfenstern auf die Distanz betrachtet. Ist $x = f(x_1, x_2, \dots, x_n)$ gilt für die Varianz von x im Falle nichtkorrelierter Schwankungen in den Argumenten nach [55]

$$\sigma_x^2 = \sum_{i=1}^n \sigma_i^2 \left(\frac{\partial x}{\partial x_i} \right)^2. \quad (4.27)$$

Zur Bestimmung der Varianzen der Distanz d wird die Fehlerfortpflanzung nach (4.27) auf die Distanzberechnung nach (3.7) bzw. (3.12) angewendet. Im ersten Schritt wird das Totzeitmodell nach (4.16) zur Berechnung der Erwartungswerte und Varianzen in den einzelnen Zeitfenstern verwendet, womit Sondereffekte durch Ratenänderungen vernachlässigt werden. Da die Varianz im betrachteten Modell keine lineare Funktion der Rate ist, muss zu jedem Zeitpunkt t stets die aktuelle Gesamtrate der einfallenden Photonen berücksichtigt werden. Für die Zahl der im Zeitfenster $t = [t_1, t_2]$ detektierten Photonen gilt allgemein

$$N_{\text{Ph}} = \int_{t_1}^{t_2} r(t) dt \quad (4.28)$$

mit der zeitabhängigen Detektionsrate $r(t)$. Im nächsten Schritt werden die Erwartungswerte und Varianzen der Zeitfenster unter Zuhilfenahme der PDFs der einzelnen Photonendetektionen und deren Auswirkungen auf die Distanzmessung bestimmt.

4.2.2.1 Pulsmodulation

Im indirekten Verfahren mit PM wird bei der Verwendung von SPADs die Distanz aus den innerhalb der drei Zeitfenster gezählten Ereignissen berechnet (siehe Abschnitt 3.3.2.1). Bei nichtidealer Pulsform muss die Anzahl der Photonen durch Integration bestimmt werden, welchen Einfluss dies auf die Distanzberechnung hat, wird in Anhang A.1 betrachtet. Bei Annahme eines konstanten Hintergrundes und eines idealen, rechteckigen Laserpulses ist die Detektionsrate der einfallenden Photonen dagegen abschnittsweise konstant, womit die Zahl der Ereignisse durch Multiplikation der Messdauer mit der entsprechenden Rate bestimmt werden kann. Mit den Raten $r_{B,T}$ und $r_{LB,T}$, der Pulsdauer T_P und der Laufzeit T_{TOF} gilt für die gezählten Photonen in den Zeitfenstern N_i

$$\begin{aligned} N_1 &= T_{TOF}r_{B,T} + (T_P - T_{TOF})r_{LB,T}, \\ N_2 &= T_{TOF}r_{LB,T} + (T_P - T_{TOF})r_{B,T} \\ \text{und } N_3 &= T_P r_{B,T}. \end{aligned} \tag{4.29}$$

Die Varianzen der Zählwerte in den einzelnen Messfenster ergeben sich durch Einsetzen der jeweiligen Raten und Messdauern in das statische Varianzmodell nach (4.16) zu

$$\begin{aligned} \sigma_1^2 &= \frac{T_{TOF}r_{B,T}}{(1 + r_B T_T)^2} + \frac{(T_P - T_{TOF})r_{LB,T}}{(1 + r_{LB} T_T)^2}, \\ \sigma_2^2 &= \frac{T_{TOF}r_{LB,T}}{(1 + r_{LB} T_T)^2} + \frac{(T_P - T_{TOF})r_{B,T}}{(1 + r_B T_T)^2} \\ \text{und } \sigma_3^2 &= \frac{T_P r_{B,T}}{(1 + r_B T_T)^2}. \end{aligned} \tag{4.30}$$

Zur Berechnung der Varianz der gemessenen Distanz werden nach (4.27) weiterhin die partiellen Ableitungen der Distanzberechnung (3.7) nach den einzelnen Zählwerten N_i benötigt. Diese ergeben sich zu

$$\begin{aligned}\frac{\partial d}{\partial N_1} &= d_{\max} \frac{N_3 - N_2}{(N_1 + N_2 - 2N_3)^2}, \\ \frac{\partial d}{\partial N_2} &= d_{\max} \frac{N_1 - N_3}{(N_1 + N_2 - 2N_3)^2} \\ \text{und } \frac{\partial d}{\partial N_3} &= d_{\max} \frac{N_2 - N_1}{(N_1 + N_2 - 2N_3)^2}\end{aligned}\tag{4.31}$$

mit d_{\max} nach (3.5). Einsetzen der Varianzen nach (4.30) und der partiellen Ableitungen nach (4.31) in die Fehlerfortpflanzung nach (4.27) liefert für die Varianz der gemessenen Distanz bei PM

$$\begin{aligned}\sigma_{\text{PM}}^2 &= \frac{c^2}{N_Z 4 T_P (r_{\text{LB,T}} - r_{\text{B,T}})^2} \left(\frac{(T_P T_{\text{TOF}} - T_{\text{TOF}}^2) r_{\text{LB,T}}}{(1 + r_{\text{LB}} T_T)^2} \right. \\ &\quad \left. + \frac{(7 T_{\text{TOF}}^2 - 7 T_P T_{\text{TOF}} + 2 T_P^2) r_{\text{B,T}}}{(1 + r_{\text{B}} T_T)^2} \right).\end{aligned}\tag{4.32}$$

Zur Reduktion der Varianz werden in realen Anwendungen die Photonen über N_Z Pulse akkumuliert, wodurch sich eine Reduktion um den Faktor N_Z ergibt. Der erste Faktor zeigt, dass die Varianz mit steigender Pulsdauer und Photonenrate des empfangenen Pulses abnimmt. Weiterhin zeigen die Nenner der beiden Summanden in der Klammer, dass die Varianz mit steigenden Raten und Totzeit abnimmt. Für $T_T = 0$ entspricht (4.32) dem in [25] publizierten Resultat, dort wurde die Berechnung der Varianz ohne Berücksichtigung der Totzeit durchgeführt. Die berechnete Varianz für $T_P = 200$ ns, $r_L = 10$ MHz, $r_B = 5$ MHz und $N_Z = 1$ mit und ohne Einfluss der Totzeit ist in Abbildung 4.4 dargestellt. Dabei zeigt sich die Parabelform der Varianz, welche sich in (4.32) erkennen lässt. Weiterhin erhöht sich die Varianz bei vorhandener Totzeit, was der reduzierten Photonenrate geschuldet ist. Dieses Varianzmodell gilt für konstante Raten, was in der betrachteten Anwendung jedoch nicht gegeben ist.

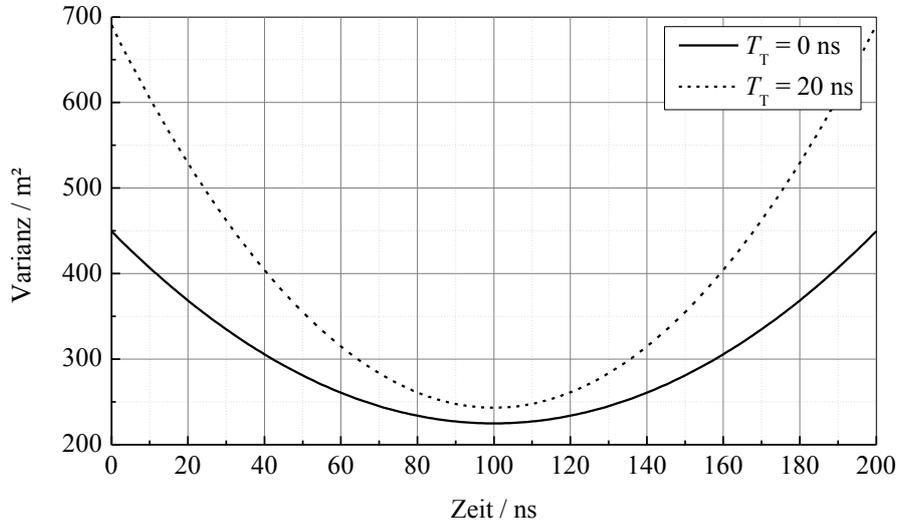


Abbildung 4.4: Varianz der indirekten Messung mit Pulsmodulation

Für eine detailliertere Betrachtung des Totzeiteinflusses auf die indirekte Laufzeitmessung mit PM, werden die PDFs der einzelnen Photonen während des Messzeitraums von $t = [0, 3T_P]$ betrachtet. Es wird angenommen, dass vor dem Messzeitraum die Photonennrate eine ausreichend lange Zeit konstant war, sodass $r(t < 0) = r_{B,T}$ gilt. Diese Annahme kann getroffen werden, da zwischen einzelnen Messungen in der Regel eine Pause zur Datenauslese oder aufgrund der begrenzten Wiederholungsrate des Lasers eingehalten werden muss. Damit folgt für die PDF der Photonendetektion unmittelbar

$$P_{Ph}(t) = r_{B,T} \text{ für } 0 \leq t < T_{TOF}, \quad (4.33)$$

da sich die Rate bis zur Ankunft des reflektierten Laserpulses nicht verändert. Dabei wird eine relativ langsame Änderung der Hintergrundrate $r_{B,T}$ vorausgesetzt, was aufgrund der Kürze der Messdauer für die meisten Anwendungen zutrifft. Es stellt sich nun die Frage nach der PDF des ersten Photons nach Anstieg der Rate aufgrund des Laserpulses. Der Einfachheit halber wird die Pulsdauer zunächst als unendlich angenommen, womit für die PDF der IAT

$$P_{IA}(t, t_0) = P_{IA}(t) = r_{LB} e^{-r_{LB}t} \text{ für } t \geq 0 \quad (4.34)$$

folgt. Damit lässt sich die PDF des ersten Photons nach (4.18) zu

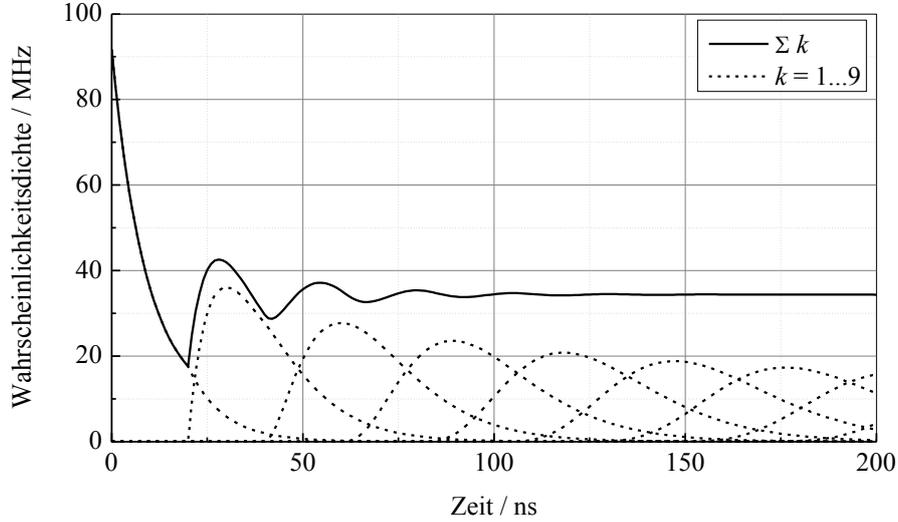


Abbildung 4.5: PDFs der einzelnen Photonen sowie Summe nach Anstieg der Photonenrate von 10 MHz auf 110 MHz

$$P_1(t) = \begin{cases} r_{B,T} \left(1 + \frac{r_L}{r_B} e^{-r_{LB}t} \right) & \text{für } 0 \leq t \leq T_T \\ r_{B,T} \left(e^{-r_{LB}(t-T_T)} + \frac{r_L}{r_B} e^{-r_{LB}t} \right) & \text{für } T_T < t \end{cases} \quad (4.35)$$

berechnen, wobei $T_{\text{TOF}} = 0$ angenommen wird. Analog zu (4.26) lässt sich auf Basis von (4.20) die PDF des k -ten Photons berechnen zu

$$P_k(t) = \begin{cases} r_{B,T} (1 + B(k) - A(t_k, k-1)) & \text{für } 0 \leq t_k < T_T \\ r_{B,T} (A(t_{k+1}, k) + B(k) - A(t_k, k-1)) & \text{für } t_k \geq T_T \end{cases} \quad (4.36)$$

für $k \in \mathbb{N}$ mit $t_k = t - (k-1)T_T$ sowie den Abkürzungen

$$A(x, m) = e^{-r_{LB}x} \sum_{n=-1}^{m-1} \frac{r_{LB}^n x^n}{n!} \quad \text{und} \quad B(x) = \frac{r_L r_{LB}^{x-1} t^{x-1}}{r_B (x-1)!} e^{-r_{LB}t}. \quad (4.37)$$

Verglichen mit (4.26) besteht der einzige Unterschied im Faktor $B(x)$, welcher bei konstanter Photonenrate, d. h. $r_L = 0$, entfällt. Abbildung 4.5 zeigt die PDF der Photonenankunftszeit nach (4.36) für $k = 1$ bis 9 und eine Totzeit von 20 ns nach einem Anstieg der Rate von $r_B = 10$ MHz auf $r_{LB} = 110$ MHz zum Zeit-

punkt $t = 0$ sowie deren Summe. Es zeigt sich dabei, dass sich die Summe erst nach einer gewissen Einschwingzeit auf einen statischen Wert, welcher durch (4.23) gegeben ist, einpendelt. Die Dauer dieses Vorgangs hängt von der Photonrate und der Totzeit ab.

Am Ende des Pulsempfangs fällt die Photonrate ab. Im Allgemeinen kann aufgrund der endlichen Pulsdauer die PDF der Photonendetektion im Intervall $t = [T_{\text{TOF}} + T_{\text{P}} - T_{\text{T}}, T_{\text{TOF}} + T_{\text{P}}]$ nicht als konstant angenommen werden. Stattdessen muss die Summe der nach (4.36) berechneten PDFs entsprechend

$$P_{\text{Ph}}(t) = \sum_{k=1}^{\infty} P_k(t) \quad (4.38)$$

betrachtet werden. Die PDF der ersten Photonendetektion nach dem Ende des Pulses ergibt sich nach (4.18) mit (4.38) zu

$$P_1(t) = \int_0^{\min(t, T_{\text{T}})} P_{\text{Ph}}(\tau + T_{\text{P}} - T_{\text{T}}) r_{\text{B}} \exp(-r_{\text{B}}(t - \tau)) d\tau + \left(1 - \int_{-T_{\text{T}}}^0 P_{\text{Ph}}(t + T_{\text{P}}) dt \right) r_{\text{B}} \exp(-r_{\text{B}}t). \quad (4.39)$$

Ist die Pulsdauer ausreichend lang, kann näherungsweise eine konstante Wahrscheinlichkeit im betrachteten Zeitintervall angenommen werden. Unter dieser Annahme ergeben sich die PDFs der Photonen entsprechend (4.36), wobei r_{LB} und r_{B} vertauscht werden, da die Rate vor der Änderung nun der Rate danach und umgekehrt entspricht, sowie r_{L} durch $-r_{\text{L}}$ ersetzt wird, da die Ratenänderung nun negativ ist. Abbildung 4.6 zeigt die auf diese Weise berechneten PDFs der einzelnen Photonen sowie deren Summe für eine Totzeit von 20 ns und Raten von $r_{\text{B}} = 10$ MHz und $r_{\text{LB}} = 110$ MHz. Hierbei zeigt sich eine reduzierte Detektionswahrscheinlichkeit nach dem Ende des Pulses zum Zeitpunkt $t = 0$, da sich die SPAD zu diesem Zeitpunkt mit höherer Wahrscheinlichkeit in Totzeit befindet. Weiterhin fällt aufgrund der geringeren Rate der Einschwingvorgang bis zum Erreichen eines statischen Zustandes kürzer aus.

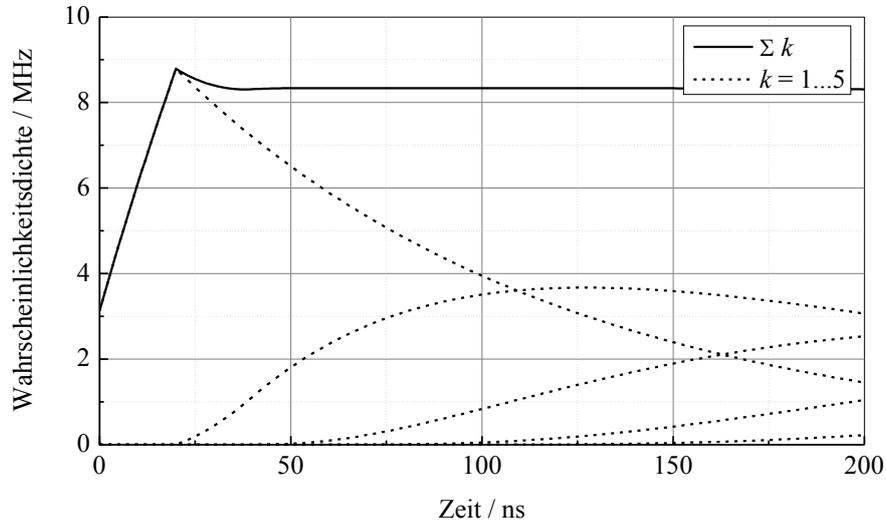


Abbildung 4.6: PDFs der einzelnen Photonen sowie Summe nach Abfall der Photonenrate von 110 MHz auf 10 MHz

Die beiden in Abbildung 4.5 und Abbildung 4.6 gezeigten Änderungen der Detektionswahrscheinlichkeit führen zu einem erhöhten Erwartungswert des Zählwertes im ersten sowie einer Reduktion im zweiten Zeitfenster. Dies resultiert in einem systematischen Fehler der Laufzeitmessung bei der Verwendung von (3.7) zur Berechnung der Distanz aus den Zählwerten. Die Erwartungswerte lassen sich durch eine Integration über die Summen der PDFs berechnen. Für das erste Zeitfenster gilt

$$\mu_1 = r_{B,T} T_{\text{TOF}} + \int_0^{T_P - T_{\text{TOF}}} P_{\text{Ph,L}}(t) dt \quad (4.40)$$

mit der Summe der PDFs der Photonen während des Laserpulses $P_{\text{Ph,L}}(t)$. Der erste Summand ergibt sich aus der konstanten Wahrscheinlichkeit vor der Ankunft des reflektierten Pulses auf dem Sensor. Für das zweite Zeitfenster gilt

$$\mu_2 = \int_{T_P - T_{\text{TOF}}}^{T_P} P_{\text{Ph,L}}(t) dt + \int_0^{T_P - T_{\text{TOF}}} P_{\text{Ph,B}}(t) dt \quad (4.41)$$

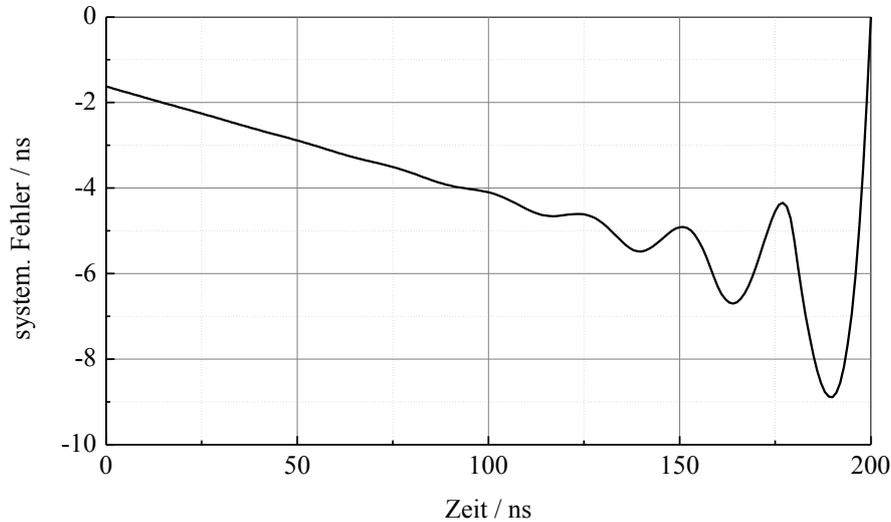


Abbildung 4.7: Systematischer Fehler der Laufzeitmessung

mit der Summe der PDFs der Photonen nach dem Laserpuls $P_{\text{Ph,B}}(t)$. In den beiden Summenfunktionen entspricht der Zeitpunkt $t = 0$ dem Anstieg bzw. Abfall der Rate. Folgt das dritte Zeitfenster, welches der Erfassung des Hintergrundes dient, unmittelbar dem zweiten, ergibt sich auch hier eine Abweichung vom idealen Erwartungswert. Da sich dieses Problem durch eine kurze Wartezeit vor dem Fenster lösen lässt, wird der Erwartungswert des dritten Fensters als ideal mit $\mu_3 = r_{\text{B,T}}T_{\text{P}}$ angenommen. Abbildung 4.7 zeigt den systematischen Fehler für Photonennraten von $r_{\text{B}} = 10$ MHz und $r_{\text{L}} = 100$ MHz. Negative Werte entsprechen dabei einer Verschiebung der Messwerte hin zu kürzeren Zeiten. Für höhere Laufzeiten spiegelt sich der Einschwingvorgang aus Abbildung 4.5 im Fehler wider, da zu diesem Zeitpunkt dieser Bereich vom ersten in das zweite Fenster übergeht. Prinzipiell steigt der Fehler mit der Ereignisrate während des Pulses. Da diese jedoch mit der Distanz quadratisch abnimmt, ist bei hohen Distanzen ein geringerer Fehler zu erwarten.

Zur Berechnung der Varianzen der Zählwerte werden die Fenster in Abschnitte konstanter Ereignisrate unterteilt. Grundsätzlich lässt sich die Wahrscheinlichkeit, genau k Photonen in einem Zeitintervall zu detektieren aus dem Integral über die Differenz zwischen der PDF des k -ten und $(k + 1)$ -ten Photons

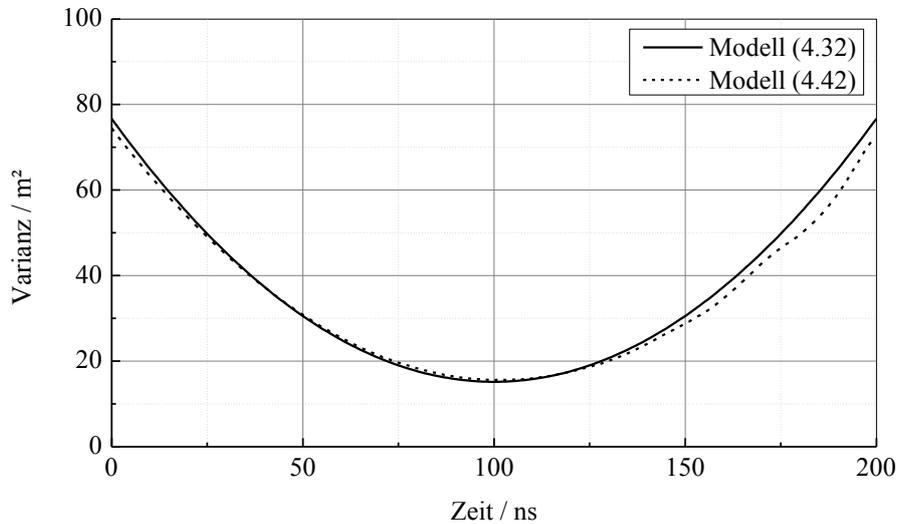


Abbildung 4.8: Vergleich des genäherten und exakten Varianzmodells

bestimmen. Aus den Wahrscheinlichkeiten für $k \in \mathbb{N}$ lässt sich im Weiteren die Varianz berechnen. Damit gilt allgemein

$$\sigma^2 = \sum_{k=0}^{\infty} \int_0^{t_1} P_k(t) - P_{k+1}(t) dt (k - \mu)^2 \quad (4.42)$$

mit $P_k(t)$ nach (4.26) bzw. (4.36) sowie $P_0(t) = \delta(t)$, mit der Diracschen Deltafunktion $\delta(t)$, für die Varianz der Ereigniszahl im Zeitintervall $t = [0, t_1]$. Abbildung 4.8 zeigt die mittels des Modells nach (4.32) sowie des exakten Modells auf Basis der Varianzen nach (4.42) berechnete Varianz der indirekten Laufzeitmessung mit PM für $T_p = 200$ ns, $r_L = 100$ MHz, $r_B = 10$ MHz und $N_Z = 1$. Dabei zeigen sich keine signifikanten Abweichungen zwischen beiden Modellen. Für geringe Photonennraten fällt der Unterschied geringer aus, da sich die PDF der Photonendetektion der idealisierten Annahme in (4.32) annähert. Der Vergleich zeigt, dass das Modell nach (4.32) für moderate Photonennraten eine gute Näherung der Varianz darstellt. Da für die in dieser Arbeit betrachteten Anwendungen die maximale Reichweite eines Systems von besonderem Interesse ist und hierbei geringe Photonennraten auftreten, wird für die weiteren Betrachtungen das Varianzmodell nach (4.32) herangezogen.

4.2.2.2 Sinusmodulation

Bei der Verwendung von CW wird die Distanz aus der Phasenverschiebung des emittierten Lichtsignals, welche wiederum aus den gezählten Ereignissen innerhalb der vier Zeitfenster bestimmt wird, berechnet (siehe Abschnitt 3.3.2.2). Ein durch Nichtlinearitäten verursachtes nichtideales Sinussignal beeinflusst dabei den systematischen und zufälligen Fehler der Distanzmessung. Eine Betrachtung dieser Effekte findet sich in Anhang A.2. Bei einem idealen Sinus gilt für die Detektionsrate über der Zeit nach dem statischen Totzeitmodell

$$r(t) = \frac{r_{LB} + r_L \sin\left(\frac{2\pi}{T_{CW}}(t - T_{TOF})\right)}{1 + T_T \left(r_{LB} + r_L \sin\left(\frac{2\pi}{T_{CW}}(t - T_{TOF})\right) \right)} \quad (4.43)$$

mit der Periodendauer T_{CW} und der Laufzeit T_{TOF} . Die gezählten Photonen bestimmen sich nach (4.28). Bei vier Zeitfenstern, die gleichmäßig über die Periode verteilt sind, ergibt sich für die gezählten Ereignisse der einzelnen Fenster

$$N_i = \frac{T_W}{T_T} - \frac{T_{CW}}{T_T \pi \sqrt{(1 + T_T r_B)^2 - (T_T r_L)^2}} \left(A\left(\frac{\pi T_W}{2 T_{CW}}\right) - A\left(-\frac{\pi T_W}{2 T_{CW}}\right) \right) \quad (4.44)$$

mit der Abkürzung

$$A(x) = \operatorname{atan} \left(\frac{(1 + T_T r_B) \tan\left(\frac{\pi(i-1)}{4} + x - \frac{\pi T_{TOF}}{T_{CW}}\right) + T_T r_L}{\sqrt{(1 + T_T r_B)^2 - (T_T r_L)^2}} \right) \quad (4.45)$$

für $i = 1$ bis 4 mit der Fensterbreite T_W . Da sich die Ereignisrate kontinuierlich über der Zeit ändern, wird die Varianz der Zählwerte über ein Integral nach

$$\sigma_i^2 = \int_{(i-1)T_{CW}/4 - T_W/2}^{(i-1)T_{CW}/4 + T_W/2} \frac{r_{LB} + r_L \sin\left(\frac{2\pi}{T_{CW}}(t - T_{TOF})\right)}{\left(1 + \left(r_{LB} + r_L \sin\left(\frac{2\pi}{T_{CW}}(t - T_{TOF})\right)\right) T_T\right)^3} dt \quad (4.46)$$

berechnet. Auf eine weitere Vereinfachung der Gleichung wird an dieser Stelle verzichtet. Die partiellen Ableitungen der Gleichung zur Berechnung der Distanz aus den Zählwerten (3.12) nach den Fenstern N_i ergibt

$$\frac{\partial d}{\partial N_1} = -\frac{\partial d}{\partial N_3} = \frac{d_{\max}(N_2 - N_4)}{2\pi((N_2 - N_4)^2 + (N_3 - N_1)^2)} \quad (4.47)$$

und

$$\frac{\partial d}{\partial N_2} = -\frac{\partial d}{\partial N_4} = \frac{d_{\max}(N_3 - N_1)}{2\pi((N_4 - N_2)^2 + (N_3 - N_1)^2)}.$$

Für die Varianz der Distanz ergibt sich mit (4.46) und (4.47)

$$\sigma_{\text{CW}}^2 = \frac{c^2 T_{\text{CW}}^2}{16\pi^2((N_4 - N_2)^2 + (N_3 - N_1)^2)^2} \sum_{i=1}^4 \left(\sigma_i^2 \left(\sin^2 \left(i \frac{\pi}{2} \right) (N_2 - N_4)^2 + \cos^2 \left(i \frac{\pi}{2} \right) \sin^2 (N_3 - N_1)^2 \right) \right). \quad (4.48)$$

Eine weitere Lösung der Gleichung führt zu keinem übersichtlicheren Resultat, weshalb stattdessen eine numerische Berechnung für beispielhafte Zahlwerte durchgeführt wird, welche in Abbildung 4.9 dargestellt ist.

Da (4.44) nur für $T_T > 0$ gilt, wird für den totzeitfreien Fall eine separate Berechnung durchgeführt. Die Erwartungswerte der Fenster sind nach (3.8) gegeben, was aufgrund der Poisson-Verteilung gleichzeitig den Varianzen der einzelnen Messfenster entspricht. Damit vereinfachen sich die partiellen Ableitungen zu

$$\frac{\partial d}{\partial N_1} = -\frac{\partial d}{\partial N_3} = -\frac{d_{\max} \cos \left(\frac{2\pi T_{\text{TOF}}}{T_{\text{CW}}} \right)}{4T_{\text{CW}} r_L \sin \left(\frac{\pi T_{\text{W}}}{T_{\text{CW}}} \right)} \quad (4.49)$$

und

$$\frac{\partial d}{\partial N_2} = -\frac{\partial d}{\partial N_4} = -\frac{d_{\max} \sin \left(\frac{2\pi T_{\text{TOF}}}{T_{\text{CW}}} \right)}{4T_{\text{CW}} r_L \sin \left(\frac{\pi T_{\text{W}}}{T_{\text{CW}}} \right)}.$$

Da die quadrierten partiellen Ableitungen von (3.12) nach N_1 und N_3 sowie N_2 und N_4 jeweils identisch sind, können die Varianzen jeweils addiert werden. Da

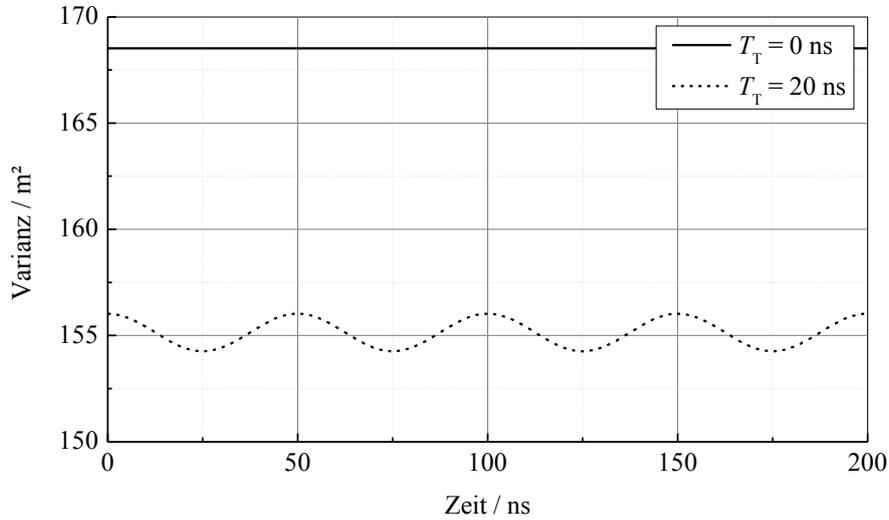


Abbildung 4.9: Varianz der indirekten Messung mit Sinusmodulation

zwischen beiden eine Phasenverschiebung von π liegt, eliminieren sich die Sinusterme und die Varianz der Laufzeit ergibt sich zu

$$\sigma_{\text{CW}}^2 = \frac{c^2 T_W r_{\text{LB}}}{32 r_{\text{L}}^2 \sin^2 \left(\frac{\pi T_W}{T_{\text{CW}}} \right)}. \quad (4.50)$$

In diesem Fall ist die Varianz laufzeitunabhängig. Diese Lösung stimmt mit der in [38] vorgestellten Lösung überein. Die Varianz nach (4.50) hängt auch von der gewählten Fensterdauer T_W ab. Für eine konstante Periodendauer T_{CW} liefert $T_W / \sin^2(\pi T_W / T_{\text{CW}})$ einen Minimalwert bei $T_W / T_{\text{CW}} = 0,37$. Ein naheliegender und einfach zu realisierender Wert liegt bei $T_W / T_{\text{CW}} = 0,25$. Bei diesem liegt der Faktor um etwa 10 % höher. Einsetzt in (4.50) folgt damit

$$\sigma_{\text{CW}}^2 = \frac{c^2 T_W r_{\text{LB}}}{N_z 64 r_{\text{L}}^2} \quad (4.51)$$

unter Berücksichtigung N_z wiederholter Messungen. In Abbildung 4.9 ist die berechnete Varianz mit und ohne Totzeit für die Parameter $T_{\text{CW}} = 200$ ns, $T_W = 50$ ns, $r_{\text{L}} = 5$ MHz, $r_{\text{B}} = 10$ MHz und $N_z = 1$ dargestellt. Dabei zeigen sich

eine leichte Reduktion der Varianz unter dem Einfluss von Totzeit sowie eine Abhängigkeit von der Laufzeit in der Form $A(T_T) \cos(8\pi T_{\text{TOF}}/T_{\text{CW}})$.

Auf eine detaillierte Analyse der Varianz über die Berechnung der PDFs der einzelnen Photonenankunftszeiten wird an dieser Stelle verzichtet, da eine entsprechende Berechnung keine signifikante Abweichung zum vorgestellten Varianzmodell bei geringen und moderaten Detektionsraten zeigt [58].

4.2.3 Vergleich der Modulationsarten

Für einen möglichst realitätsnahen Vergleich der Varianzen der beiden indirekten Methoden zur Laufzeitbestimmung nach (4.32) und (4.48), werden diese unter den folgenden Bedingungen betrachtet: Identische Ereignisrate der Hintergrundstrahlung, identische mittlere Ereignisrate der Lichtquelle, identische Integrationsdauer einer Messung sowie identische Reichweite. Die Fensterdauer bei der Messung mit CW wird zu $T_W/T_{\text{CW}} = 0,25$ gewählt. Eine identische Reichweite in beiden Methoden resultiert in der Bedingung

$$T_{\text{CW}} = T_{\text{P}}. \quad (4.52)$$

Für eine identische mittlere Ereignisrate der Anregung muss die Bedingung

$$\frac{1}{T_{\text{CW}}} \int_0^{T_{\text{CW}}} r_{\text{L,CW}} \left(1 + \sin\left(\frac{2\pi}{T_{\text{CW}}} t\right) \right) dt = \frac{1}{3T_{\text{P}}} \int_0^{T_{\text{P}}} r_{\text{L,PM}} dt \quad (4.53)$$

erfüllt sein. Eine Berechnung der Integrale und Einsetzen von (4.52) liefert

$$3r_{\text{L,CW}} = r_{\text{L,PM}} = 3r_{\text{L}}. \quad (4.54)$$

Die Anzahl der Messungen N_{Z} bestimmt sich aus dem Quotienten aus Integrationszeit und der Dauer einer Messung, welche sich im Falle von CW zu T_{CW} und für PM zu $3T_{\text{P}}$ ergibt. Umstellen beider Quotienten nach der Integrationszeit, Gleichsetzen und Einsetzen von (4.52) ergibt

$$N_{\text{Z,CW}} = 3N_{\text{Z,PM}}. \quad (4.55)$$

Mit (4.52), (4.54) und (4.55) werden nun die Varianzen verglichen.

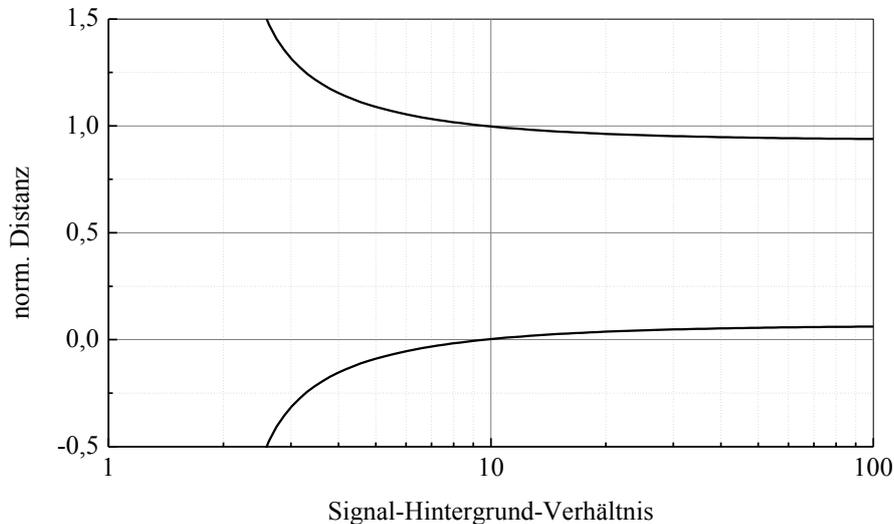


Abbildung 4.10: Vergleich der indirekten Messmethoden nach (4.56)

Ohne Berücksichtigung der Totzeit wird im CW-Verfahren statt (4.48) die Varianz nach (4.51) für den Vergleich mit dem PM-Verfahren (4.32) herangezogen. Welche Variante der indirekten Laufzeitmessung die geringe Varianz aufweist, ist vom Verhältnis zwischen r_L und r_B sowie zwischen T_{TOF} und T_P , was der normierten Distanz entspricht, abhängig. Der Quotient von r_L und r_B wird im Weiteren als Signal-Hintergrund-Verhältnis (Signal-to-Background-Ratio, SBR) bezeichnet. Gleichsetzen der beiden Varianzen ergibt

$$\frac{T_{\text{TOF}}}{T_P} = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{(29r_B - 3r_L)}{48(2r_B - r_L)}}. \quad (4.56)$$

Der Kurvenverlauf nach (4.56) ist in Abbildung 4.10 dargestellt. Der relevante Bereich mit einer Messdistanz zwischen der minimalen und maximalen Entfernung entspricht einer normierten Distanz im Intervall $[0,1]$. Zwischen beiden Kurven zeigt die Messung mit CW eine geringere Varianz. Der Vergleich zeigt, dass nur für Laufzeiten nahe null oder der maximalen Reichweite und geringer Hintergrundstrahlung die Variante mit PM eine geringere Varianz liefert. Da eine hohe Reichweite bei schwierigen Umweltbedingungen, d. h. bei niedrigem SBR, erzielt werden soll, stellt CW das geeignetere Messkonzept dar.

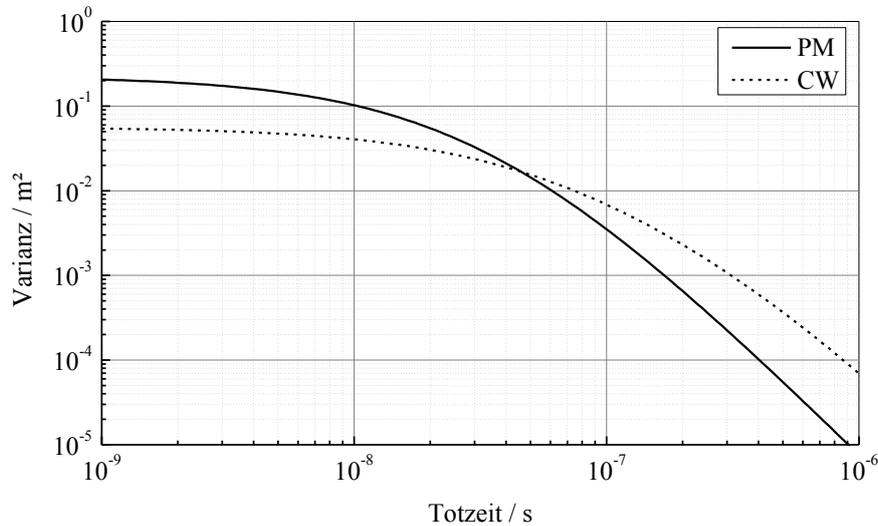


Abbildung 4.11: Varianz der indirekten Laufzeitmessungen über die Totzeit

Um die Totzeit im Vergleich zu berücksichtigen, wird der Einfluss dieser auf die Varianzen betrachtet. Die Form der Varianz nach (4.16) entspricht der eines Tiefpasses 3. Ordnung mit der Grenztotzeit

$$T_{T,3dB} = \frac{\sqrt[6]{2} - 1}{r} \approx \frac{1}{8r}, \quad (4.57)$$

wobei diese von der Photonenrate r abhängt. Analog lässt sich die Grenztotzeit für die Varianz der Laufzeitmessung mit PM nach (4.32) bestimmen. Hierbei weisen die beiden Summanden der Klammer unterschiedliche Grenztotzeiten auf. Um die Einflüsse der Totzeit vernachlässigen zu können, sollte die Totzeit unter der geringeren der beiden Zeiten liegen. Eine Abnahme der Varianz mit 60 dB pro Dekade zeigt sich jedoch erst ab einer Totzeit über der höheren der beiden Grenztotzeiten, da ab diesem Punkt beide Summanden mit demselben Faktor abfallen. Für die Laufzeitmessung mit CW ist die Überlegung analog. Hier treten aufgrund der kontinuierlichen Variation von r jedoch unendlich viele Grenztotzeiten auf. Die geringste Totzeit liegt in diesem Fall aufgrund unterschiedlicher Definition der Rate r_L über der PM, während die höchste Grenztotzeit identisch zur PM ist. Für $r_L \gg r_B$ und $T_T \gg T_{T,3dB}$ ist die Varianz der PM aufgrund der früher abfallenden Kennlinie geringer als bei der CW. Abbildung

4.11 zeigt den Verlauf der Varianz beider Messverfahren über der Totzeit. Für die Berechnung wurden Detektionsraten von $r_L = 10$ MHz und $r_B = 0$, eine Laufzeit von 100 ns, eine Puls- bzw. Periodendauer von 200 ns und eine Summation von $N_Z = 500$ Einzelmessungen angenommen. Da in realen Anwendungen Totzeiten zwischen 100 ns und 20 ns üblich sind [14], wobei geringere Werte bevorzugt werden, schneidet die indirekte Messung mit CW auch unter Berücksichtigung der Totzeit besser ab.

4.3 Direkte Laufzeitmessung

Bei der direkten Messung der Lichtlaufzeit zur Entfernungsbestimmung wird eine hochpräzise elektronische Stoppuhr mit dem Aussenden eines Laserpulses gestartet und mit dem Empfang des reflektierten Pulses wieder gestoppt [4]. Für die Genauigkeit der Messung sind die zeitlichen Schwankungen der beteiligten Komponenten relevant. Dazu zählen die SPAD, der TDC und die Synchronisation [59]. Der Jitter der SPAD ist vom physikalischen Aufbau der Diode abhängig und wurde bereits in Abschnitt 2.4.3 näher betrachtet. Ein weiterer Faktor ist die Ankunftszeit des für die Laufzeitmessung relevanten Photons, welche ebenfalls einer statistischen Schwankung unterliegt. Die Betrachtungen in diesem Abschnitt erfolgen anhand eines idealen Rechteckpulses, lassen sich aber auf beliebige Pulsformen übertragen.

4.3.1 Statistik

In der direkten Laufzeitmessung wird der Ankunftszeitpunkt des ersten detektierten Photons in jedem Messzyklus erfasst und in einem Histogramm vermerkt. Der Ankunftszeitpunkt unterliegt statistischen Schwankungen, welche von der Pulsform und der Intensität des Pulses abhängen. Entsprechend [60] ist die PDF der Ankunftszeit des ersten Photons eines Pulses gegeben zu

$$P_1^*(t, N_{\text{Ph}}) = P_{\text{Ph}}(t) N_{\text{Ph}} \left(1 - \int_0^t P_{\text{Ph}}(t) dt \right)^{N_{\text{Ph}}-1} \quad (4.58)$$

mit der Verteilung der Photonen innerhalb des Pulses $P_{\text{Ph}}(t)$, was der Pulsform entspricht, und der Anzahl der Photonen N_{Ph} pro Puls. (4.58) ergibt sich aus der Wahrscheinlichkeit, dass eines der Photonen zum Zeitpunkt t eintrifft multipliziert mit der Wahrscheinlichkeit, dass $N_{\text{Ph}} - 1$ Photonen erst nach dem Zeitpunkt t eintreffen. Dabei wird von der Annahme ausgegangen, dass exakt N_{Ph} Photonen innerhalb der Pulsdauer einfallen. Eine realitätsnähere Annahme stellt die Poisson-Verteilung nach (4.5) für die Anzahl der Photonen dar. Bezieht man diese mit ein, ergibt sich die PDF des ersten Photons durch die mit der Poisson-Verteilung gewichtete Addition von (4.58) zu

$$P_1(t, \mu) = \sum_{x=0}^{\infty} P_{\text{P}}(x, \mu) P_1^*(t, x) \quad (4.59)$$

mit dem Erwartungswert der Photonenzahl μ . Für die Vergleichbarkeit mit der indirekten Methode wird im Folgenden die Varianz des direkten Verfahrens mit der Erfassung des jeweils ersten Photons je Zyklus bestimmt. Bei konstanter Photonenrate des reflektierten Pulses r_{L} und Pulsdauer T_{P} ergibt sich für den Erwartungswert der detektierten Photonen

$$\mu = r_{\text{L}} T_{\text{P}}. \quad (4.60)$$

Weiterhin gilt nach [60] für die PDF des ersten Photons bei Gleichverteilung

$$P_1^*(t, N_{\text{Ph}}) = \frac{N_{\text{Ph}}}{T_{\text{P}}} \left(1 - \frac{t}{T_{\text{P}}}\right)^{N_{\text{Ph}}-1}. \quad (4.61)$$

Damit folgt für die Gesamtverteilung nach (4.59)

$$P_1(t, \mu) = \sum_{x=0}^{\infty} \frac{(r_{\text{L}} T_{\text{P}})^x}{x!} e^{-r_{\text{L}} T_{\text{P}}} \frac{x}{T_{\text{P}}} \left(1 - \frac{t}{T_{\text{P}}}\right)^{x-1} = r_{\text{L}} e^{-r_{\text{L}} t} \quad (4.62)$$

was der erwarteten Exponentialverteilung für die Zwischenankunftszeiten bei einer konstanten Photonenrate r_{L} entspricht. Der Erwartungswert dieser Verteilung ist zu $\mu = r_{\text{L}}^{-1}$ und die Varianz zu $\sigma^2 = r_{\text{L}}^{-2}$ gegeben [61]. Tatsächlich ergibt sich die genannte Varianz nur für den Fall einer unendlichen Pulsdauer.

Bei endlicher Pulsdauer fällt die Exponentialverteilung nach (4.62) mit dem Ende des Pulses auf null entsprechend

$$P_1(t, T_P) = (\Theta(t) - \Theta(t - T_P))r_L e^{-r_L t}, \quad (4.63)$$

da ab diesem Zeitpunkt keine Photonen mehr empfangen werden können. Aus dieser Wahrscheinlichkeitsverteilung lässt sich der Erwartungswert zu

$$\mu = \frac{1}{r_L} - \frac{r_L T_P + 1}{r_L} e^{-r_L T_P} \quad (4.64)$$

sowie die Varianz zu

$$\begin{aligned} \sigma^2 = \frac{1}{r_L^2} & (1 - (1 + r_L^2 T_P^2) e^{-r_L T_P} + (1 - r_L^2 T_P^2) e^{-2r_L T_P} \\ & - (1 + r_L T_P)^2 e^{-3r_L T_P}) \end{aligned} \quad (4.65)$$

berechnen. Den Erwartungen entsprechend zeigen (4.64) und (4.65) eine Reduktion des Erwartungswertes und der Varianz gegenüber dem Fall unendlicher Pulsdauer, welche mit sinkender Pulsdauer größer ausfällt. Analog zur indirekten Methode skaliert die Varianz mit dem Faktor N_Z^{-1} bei der Mittelung von N_Z Messwerten. Da bei endlicher Pulsdauer nicht bei jeder Einzelmessung ein Photon empfangen wird, muss die Anzahl der Laserpulse zur Bestimmung von N_Z mit $1 - \exp(-r_L T_P)$ multipliziert werden. Dieser Faktor stellt die Wahrscheinlichkeit für eine Photonendetektion innerhalb der Ankunft des reflektierten Laserpulses dar.

4.3.2 Hintergrundstrahlung

Berücksichtigt man zusätzlich vorhandene Hintergrundstrahlung, gibt es eine Wahrscheinlichkeit dafür, dass das erste registrierte Photon aus der Hintergrundstrahlung und nicht aus dem reflektierten Laserpuls resultiert. Unter der Annahme eines konstanten Hintergrundlichtes, gilt für die Wahrscheinlichkeit für das Eintreffen eines Hintergrundphotons vor der Ankunft des reflektierten Signalpulses

$$P_{\text{BG}}(T_{\text{TOF}}) = \int_0^{T_{\text{TOF}}} r_{\text{B}} \exp(-r_{\text{B}}t) dt = 1 - \exp(-r_{\text{B}}T_{\text{TOF}}). \quad (4.66)$$

Dabei zeigt sich, dass für hohe Hintergrundraten, wie sie in realen Anwendungen auftreten, und für große Messdistanzen die Wahrscheinlichkeit für das Erfassen eines Hintergrundphotons anstelle des reflektierten Pulses hoch ist. Für solche Anwendungen ist das direkte Messverfahren über die Erfassung des ersten Photons jedes Zyklus ohne Hintergrundlichtunterdrückung ungeeignet.

4.3.3 Histogramm

Um trotz Hintergrundstrahlung die Laufzeit über das direkte Verfahren zu messen, werden mehrere Messwerte in einem Histogramm akkumuliert, aus welchem anschließend mittels eines geeigneten Algorithmus die Laufzeit bestimmt wird. Da die Güte des Algorithmus einen großen Einfluss auf die Varianz des direkten Verfahrens hat, kann keine allgemeingültige Aussage getroffen werden. Um zumindest eine Abschätzung der Varianz zu erhalten, wird im Folgenden die einfachste denkbare Auswertung des Histogramms beispielhaft betrachtet. Bei dieser wird die Position des Maximalwertes als Laufzeit des empfangenden Pulses interpretiert. Zunächst wird die PDF der ersten Photonendetektion bestimmt. Hierfür wird eine zeitabhängige Photonenerate $r(t)$ angenommen. Die Verteilung des ersten Photons ergibt sich durch die jeweilige Photonenerate multipliziert mit der verbleibenden Wahrscheinlichkeit gemäß

$$P_1(t) = r(t) \left(1 - \int_0^t P_1(\tau) d\tau \right). \quad (4.67)$$

Für beliebige Photoneneraten kann (4.67) durch schrittweise Berechnung der Gleichung in Intervallen von dt beginnend bei $t = 0$ gelöst werden. Im Folgenden wird die Verteilung für eine konstante Rate des reflektierten Pulses sowie des Hintergrundes bestimmt. Für die Photonenerate gilt

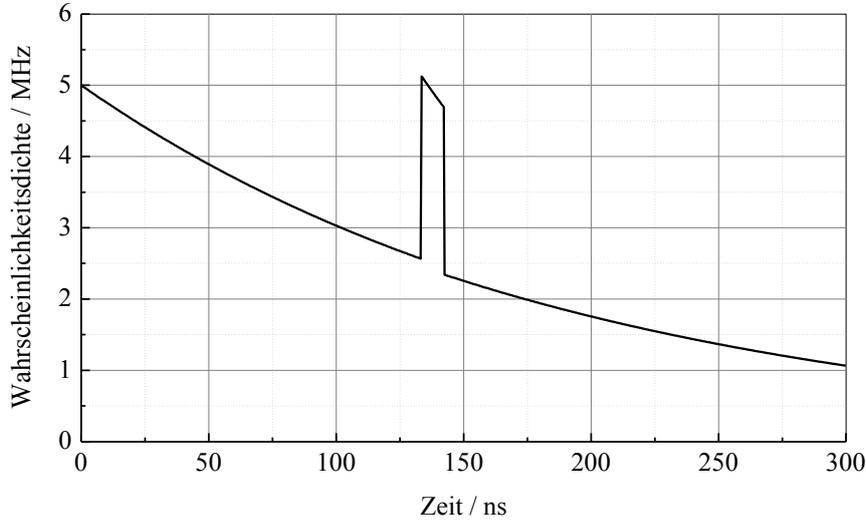


Abbildung 4.12: PDF der ersten Photonendetektion bei idealem Rechteckpuls

$$r(t) = r_B + r_L \left(\Theta(t - T_{\text{TOF}}) - \Theta(t - (T_{\text{TOF}} + T_P)) \right). \quad (4.68)$$

Die Photonenrate lässt sich in drei Abschnitte mit jeweils konstanter Rate aufteilen. Damit folgt für die Verteilung

$$P_1(t) = \begin{cases} r_B e^{-r_B t} & \text{für } 0 \leq t < T_{\text{TOF}} \\ r_{\text{LB}} e^{-r_{\text{LB}}(t-T_{\text{TOF}})} e^{-r_B T_{\text{TOF}}} & \text{für } T_{\text{TOF}} \leq t < T_{\text{TOF}} + T_P \\ r_B e^{-r_B t} e^{-r_L T_P} & \text{für } T_{\text{TOF}} + T_P \leq t \end{cases} \quad (4.69)$$

Abbildung 4.12 zeigt die PDF der ersten Photonendetektion nach (4.69) für eine Detektionsrate von jeweils 5 MHz für den reflektierten Laserpuls sowie das Hintergrundlicht. Weiterhin wurde eine Reichweite von 45 m, eine Messdistanz von 20 m sowie eine Pulsbreite von 9 ns angenommen. Aus dieser PDF kann die Wahrscheinlichkeit jeder Klasse, ein Photon zu empfangen durch Integration über die Breite der Klasse T_{Bin} gemäß

$$p_i = \int_{(i-1)T_{\text{Bin}}}^{iT_{\text{Bin}}} P_1(t) dt \quad (4.70)$$

berechnet werden. Die Zählwerte der einzelnen Klassen folgen der Binomialverteilung, folglich ist der Erwartungswert nach (4.3) gegeben zu

$$\mu_i = N_Z p_i, \quad (4.71)$$

mit der Anzahl akkumulierter Einzelmessungen N_Z . Für die Varianzen der Zählwerte der Klassen folgt gemäß (4.4)

$$\sigma_i^2 = N_Z p_i (1 - p_i). \quad (4.72)$$

Die Wahrscheinlichkeit, dass eine Klasse den Maximalwert enthält, ergibt sich aus dem Produkt der Wahrscheinlichkeit, dass diese Klasse einen Wert über oder gleich ζ enthält, mit der Wahrscheinlichkeit, dass die restlichen Klassen einen Wert kleiner ζ enthalten, aufsummiert über alle ζ . Für Klasse i ergibt sich

$$P_i = \sum_{\zeta=0}^{\infty} \left(P_B(\zeta, N_Z, p_i) \prod_{\substack{n=1 \\ n \neq i}}^{N_{\text{Bin}}} \sum_{k=0}^{\zeta-1} P_B(k, N_Z, p_n) \right) \quad (4.73)$$

mit der Binomialverteilung P_B nach (4.2) und der Anzahl der Klassen N_{Bin} . Abbildung 4.13 zeigt die berechneten Wahrscheinlichkeiten der Klassen nach (4.73). Die angenommene Reichweite von 45 m entspricht 300 ns, was bei einer Breite der Klassen von 1 ns in 300 Klassen resultiert. Das Beispiel zeigt, wie eine hohe Raten des Hintergrundlichtes zu Fehlinterpretationen der hier betrachteten Auswertung zur Distanzbestimmung führen. Da (4.73) den Fall, dass ein mehrfaches Maximum auftritt nicht berücksichtigt, da dies zu mehrdeutigen Resultaten führt, liegt die über alle Klassen kumulierte Wahrscheinlichkeit bei lediglich 54,2 %. Aus den berechneten Wahrscheinlichkeiten lässt sich der Erwartungswert und die Standardabweichung der direkten Messmethode berechnen. Für den in Abbildung 4.13 betrachteten Fall ergeben sich diese zu $\mu = 8,325$ m und $\sigma = 7,483$ m. Eine präzise und zuverlässige Messung der Distanz ist im betrachteten Beispiel somit nicht möglich.

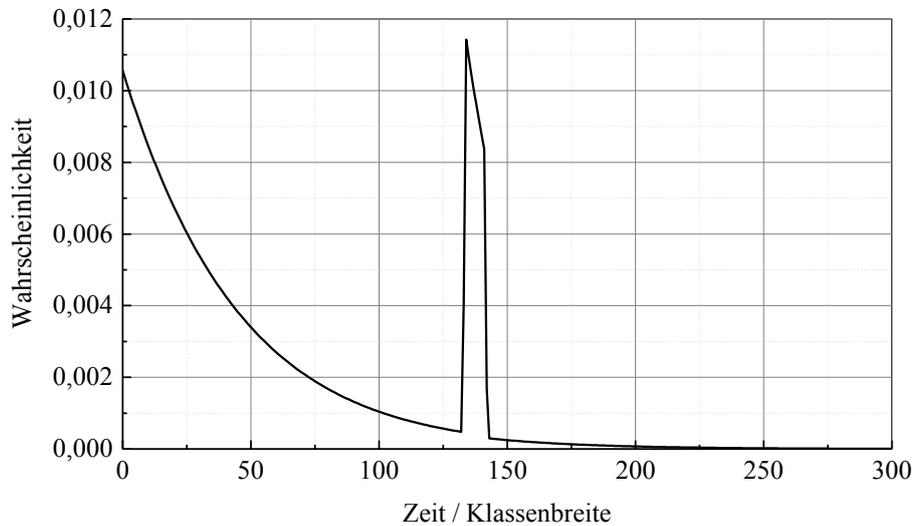


Abbildung 4.13: Wahrscheinlichkeiten der Klassen den Maximalwert zu enthalten

Für eine vom Algorithmus unabhängige Beurteilung des direkten Messverfahrens wird im Folgenden eine Kenngröße direkt aus dem resultierenden Histogramm mehrerer einzelner Zeitmessungen abgeleitet. Die Erwartungswerte der Klassen werden nach (4.71) berechnet. Damit ein Algorithmus die Ankunftszeit des reflektierten Pulses im Histogramm ermitteln kann, muss der Wert der entsprechenden Klasse über den benachbarten Klassen liegen. Aus dieser Überlegung wird ein SNR für das Histogramm entsprechend

$$a_{\text{Hist}} = \frac{N_L}{\sqrt{\sigma_L^2 + \sigma_B^2}} \quad (4.74)$$

definiert. Hierbei ist N_L die Anzahl der während dem Empfang des reflektierten Pulses erfassten Photonen, N_B die durch den Hintergrund gezählten Photonen und σ_L^2 bzw. σ_B^2 die Varianzen der Zählwerte. Die Definition des SNR ist in Abbildung 4.14 illustriert. Zur Berechnung des SNR wird der Zählwert der Klasse unmittelbar nach der Ankunft des reflektierten Pulses betrachtet, da die Position dieser Klasse die gesuchte Laufzeit repräsentiert. Da in der realen Anwendung die Breite der Klassen durch die zeitliche Auflösung des TDCs bestimmt ist, welche entsprechend Tabelle 1.1 deutlich unter 1 ns liegt, und geringe Photonennraten bei der Ermittlung der Reichweite von Interesse sind, liegen die

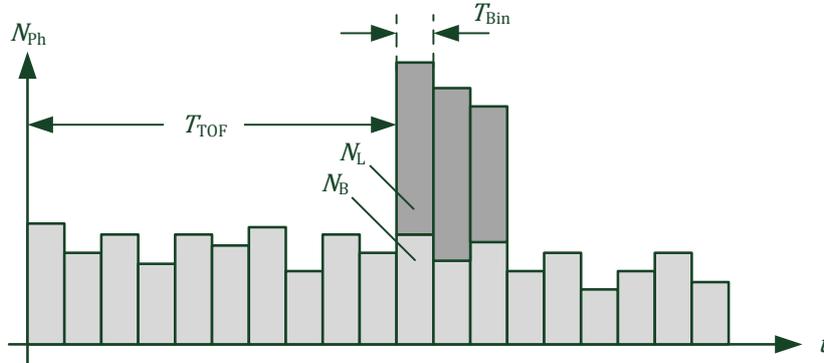


Abbildung 4.14: Definition des SNR über das Histogramm

Wahrscheinlichkeiten der Klassen typischerweise deutlich unter 5 %, womit die Binomialverteilung durch die Poisson-Verteilung genähert werden kann. Dementsprechend erfolgt die weitere Berechnung unter der Prämisse Poissonverteilter Zählwerte der Klassen. Damit folgt

$$a_{\text{Hist}} = \frac{N_L}{\sqrt{N_L + N_B}} \quad (4.75)$$

für das SNR, da in der Poisson-Verteilung die Varianz gleich dem Erwartungswert ist. Für die Anzahl der detektierten Photonen durch den Hintergrund gilt

$$N_B = N_Z \int_{T_{\text{TOF}}}^{T_{\text{TOF}}+T_{\text{Bin}}} P_1(t)|_{r_L=0} dt \approx N_Z T_{\text{Bin}} r_B e^{-r_B T_{\text{TOF}}} \quad (4.76)$$

mit der Wahrscheinlichkeitsdichte $P_1(t)$ nach (4.69) wobei $r_L = 0$. Die Näherung ergibt sich unter der Annahme konstanter Wahrscheinlichkeit über der Klasse. Für N_L gilt

$$N_L = N_Z \int_{T_{\text{TOF}}}^{T_{\text{TOF}}+T_{\text{Bin}}} P_1(t) dt - N_B \approx N_Z T_{\text{Bin}} r_L e^{-r_B T_{\text{TOF}}}. \quad (4.77)$$

Einsetzen von (4.76) und (4.77) in (4.75) ergibt

$$a_{\text{Hist}} = \sqrt{N_Z T_{\text{Bin}} e^{-r_B T_{\text{TOF}}}} \frac{r_L}{\sqrt{r_L + r_B}}. \quad (4.78)$$

Wie erwartet nimmt a_{Hist} mit steigender Hintergrundstrahlung und Distanz ab und steigt mit der Pulszahl, Binbreite und Pulsintensität an. Eine Ableitung der zu erwartenden Messgenauigkeit aus a_{Hist} gestaltet sich schwierig, da sobald der Algorithmus nicht mehr in der Lage ist die Ankunftszeit des Pulses im Histogramm zu ermitteln, eine Messung unmöglich ist. Bei gegebenem Algorithmus kann aber ein minimales SNR bestimmt werden, bis zu welchem der reflektierte Puls im Histogramm erkannt wird. Aus diesem SNR lässt sich die maximale Reichweite des Systems abschätzen.

4.3.4 Varianz

Die Varianz des direkten Messverfahrens setzt sich aus verschiedenen Komponenten zusammen. Im Falle einer Messung ohne Hintergrundstrahlung und unter der Annahme einer idealen SPAD, ist die Varianz nur durch die Ankunft des ersten Photons des reflektierten Laserpulses nach (4.65) bestimmt. Durch Baugruppen eingebrachte Jitter werden dabei vernachlässigt. Kommt Hintergrundstrahlung ins Spiel, müssen die gemessenen Laufzeiten in einem Histogramm akkumuliert werden, aus welchem anschließend die tatsächliche Laufzeit des Laserpulses ermittelt wird. Die in Abschnitt 4.3.3 für eine beispielhafte Auswertung berechneten Parameter Erwartungswert und Varianz beinhalten sowohl die Varianz des ersten Photons als auch die Auswirkungen des Histogramms. Eine generelle Abschätzung der Leistungsfähigkeit des direkten Messverfahrens gestaltet sich aufgrund der Abhängigkeit von Algorithmus schwierig, für eine erste Abschätzung kann das in (4.78) hergeleitete SNR dienen.

5 Modellbildung und Simulation

Zur Bewertung der vorgestellten und neuen Verfahren sowie zur Verifikation der Berechnungen wird im Folgenden ein Modell zur Simulation der Verfahren erstellt. Da insbesondere für das direkte Verfahren eine geschlossene Berechnung der zu erwartenden Leistung aufgrund der Abhängigkeit vom Auswertalgorithmus des Histogramms nur schwer möglich ist, dient die Simulation auch dem Vergleich beider Verfahren. Im ersten Schritt werden hierzu die Detektionsraten der einfallenden Photonen des reflektierten Signals sowie der Hintergrundstrahlung aus Sonnenlicht oder künstlicher Beleuchtung bestimmt. Diese Raten spielen eine entscheidende Rolle für die Varianz der SPAD-basierten Laufzeitmessung. Im Weiteren werden darauf aufbauend Simulationen durchgeführt. Abbildung 5.1 zeigt ein Übersichtsdiagramm des entworfenen Modells. Im oberen Bereich sind die Eingangsparameter geordnet nach den fünf Kategorien Strahlungsquelle, Szene, Optik, Sensor und System dargestellt. Aus den Parametern der Strahlungsquelle wird zunächst unter Berücksichtigung der Augensicherheit die maximal zulässige optische Leistung bestimmt. Anschließend wird daraus und mit den Parametern der Szene, Optik und des Sensors die Rate detektierter Photonen berechnet. Basierend auf diesen und den Systemparametern erfolgt die Simulation, welche statistische Eigenschaften der Messung wie Erwartungswert und Standardabweichung liefert. Mit Hilfe dieser Werte kann die Leistungsfähigkeit der Messung in Hinblick auf Reichweite und Zuverlässigkeit beurteilt werden.

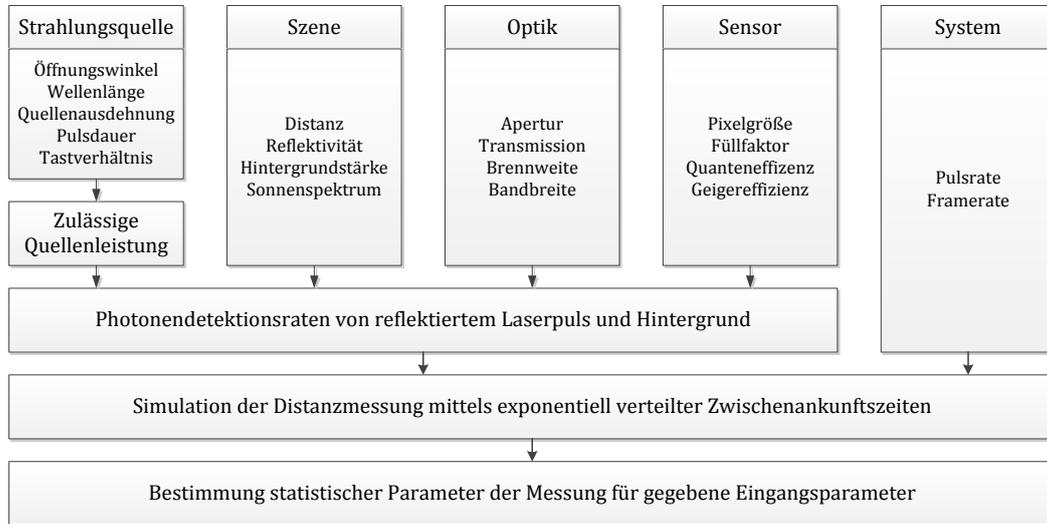


Abbildung 5.1: Modellbildung und Simulation der Laufzeitmessung

5.1 Lasergenerierte Photonenerate

In diesem Abschnitt soll berechnet werden, welche Photonendetektionsrate in einem Pixel bei gegebener optischer Quellenleistung P_L zu erwarten ist. Die Bestrahlungsstärke der Szene im Abstand d zur Quelle ist unter Annahme homogener Bestrahlung mit einer punktförmigen Laserquelle mit dem horizontalen und vertikalen Öffnungswinkel α_H und α_V gegeben zu

$$S_L = \frac{P_L}{d^2 \Omega} = \frac{P_L}{d^2 4 \arcsin\left(\sin\left(\frac{\alpha_H}{2}\right) \sin\left(\frac{\alpha_V}{2}\right)\right)}, \quad (5.1)$$

wobei die Öffnungswinkel in Radiant angegeben werden. Die Szene wird als Lambertscher Reflektor, dessen Reflexionscharakteristik mit dem Faktor η_L berücksichtigt wird, angenommen. Dieser entspricht dem Verhältnis zwischen ausgesendeter und empfangener Leistung und ist gegeben zu

$$\eta_L = \left(\frac{D}{2d}\right)^2 \quad (5.2)$$

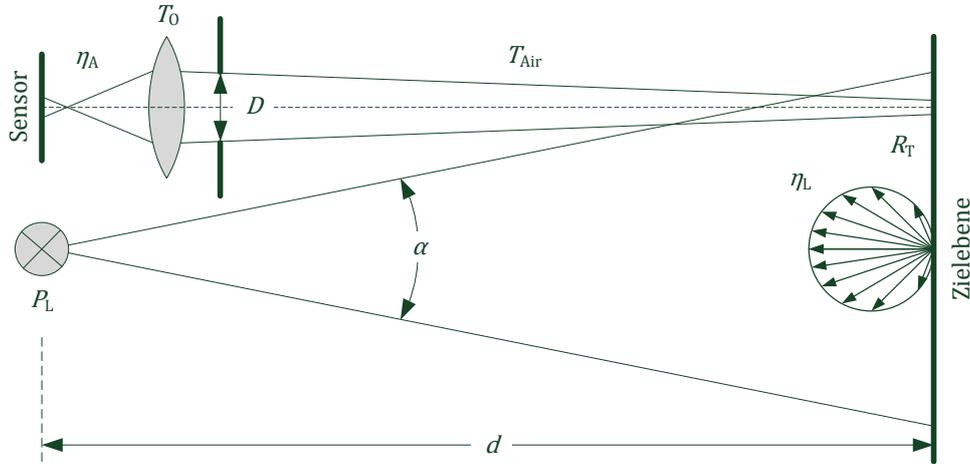


Abbildung 5.2: Modell zur Berechnung der lasergenerierten Photonenrate in Anlehnung an [32]

mit dem Durchmesser der Empfangsapertur D [38]. Zur Bestimmung der Bestrahlungsstärke auf dem Sensor wird das Abbildungsverhältnis von der Szene auf den Sensor berücksichtigt. Dieses ist gegeben zu

$$\eta_A = \left(\frac{d}{f}\right)^2 \quad (5.3)$$

mit der Brennweite der Optik f . Weiterhin werden die Reflektivität des Zielobjektes R_T , die Transmission des Objektivs T_0 und die Transmission in der Luft T_{Air} berücksichtigt. Für die Bestrahlungsstärke auf dem Sensor gilt damit

$$S_S = S_L \eta_L \eta_A R_T T_0 T_{Air} = \frac{S_L}{4f_{\#}^2} R_T T_0 T_{Air} \quad (5.4)$$

mit der Blendenzahl $f_{\#} = f/D$. Um aus der Bestrahlungsstärke des Sensors die theoretisch erfasste Photonenrate zu ermitteln, wird S_S mit der Pixelfläche A_{Pix} , dem Füllfaktor η_{FF} und der wellenlängenabhängigen PDE $\eta_{PDE}(\lambda)$ multipliziert und das Resultat durch die Photonenergie hc/λ dividiert. Es folgt

$$r_L = S_S A_{Pix} \eta_{FF} \eta_{PDE}(\lambda) \frac{\lambda}{hc} \quad (5.5)$$

wobei h dem Planckschen Wirkungsquantum, c der Lichtgeschwindigkeit und λ der Wellenlänge der verwendeten Strahlung entspricht.

5.2 Augensicherheit

Für den Einsatz in frei zugänglichen Systemen wie dem Automobil, ist es notwendig, dass die Strahlungsquelle die Richtlinien zur Augensicherheit einhält. Der Internationale Standard IEC 60825-1 beschreibt Grenzwerte und Messmethoden der maximal zulässigen Bestrahlung zur Klassifizierung von Lasersystemen. Für das hier betrachtete System wird die Laserklasse 1M betrachtet, welche Systeme auszeichnet, die ohne die Verwendung optischer Hilfsmittel keine Gefahr für den Menschen darstellen [62]. Zur Ermittlung der maximalen Photonenerate wird im Folgenden die erlaubte Laserleistung bestimmt. Die gegebenen Grenzwerte sind auf die zugängliche Emission (engl. Accessible Emission Limit, AEL) bezogen, diese berücksichtigt für divergierende Strahlung die Apertur des Auges sowie den Mindestabstand für Fokussierung. Bei gepulstem oder intensitätsmoduliertem Laser, wie er in Laufzeitmesssystemen zum Einsatz kommt, müssen verschiedene Grenzwerte berücksichtigt werden: Für einen einzelnen Puls, für einen mit der Anzahl der Pulse innerhalb einer Puls-kette gewichteten Einzelpuls und für die mittlere Leistung einer Puls-kette. Ein einzelner Puls innerhalb einer Puls-kette darf den Grenzwert für einzelne Pulse nicht überschreiten. Für eine Pulsdauer von 1 ns bis 18 μ s und eine Wellenlänge im Bereich von 700 nm bis 1050 nm gilt eine maximale Pulsenergie [62]

$$E_p = 2 \cdot 10^{-7} C_4 C_6 \text{ J}. \quad (5.6)$$

Der Korrekturfaktor C_4 berücksichtigt die Wellenlänge und ist gegeben zu

$$C_4 = 10^{0,002(\lambda-700)}, \quad (5.7)$$

wobei die Wellenlänge λ in Nanometer angegeben wird [62]. Die Ausdehnung der Laserquelle wird durch den Faktor C_6 mit

$$C_6 = \beta / \beta_{\min} \text{ für } \beta_{\min} < \beta < \beta_{\max} \quad (5.8)$$

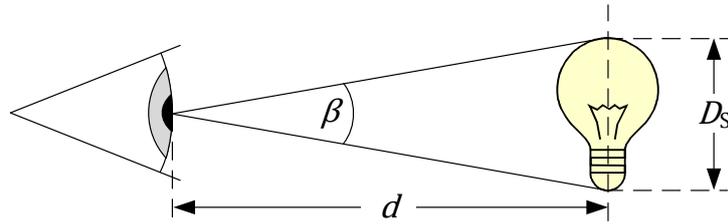


Abbildung 5.3: Winkelausdehnung der Lichtquelle in Anlehnung an [62]

mit $\beta_{\min} = 1,5 \text{ mrad}$, $\beta_{\max} = 100 \text{ mrad}$ und dem Winkel β , welcher mit

$$\beta = 2 \operatorname{atan}\left(\frac{D_s}{2d}\right) \quad (5.9)$$

aus dem Durchmesser der Quelle D_s und dem Abstand zur Quelle d bestimmt werden kann, berücksichtigt [62]. Die durch (5.6) festgelegte Energie darf maximal während eines Pulses durch eine Apertur mit 7 mm Durchmesser in 10 cm Entfernung gelangen. Zur Bestimmung der maximalen Laserleistung wird die Leistungsdichte im Abstand der Messapertur betrachtet. Da sich, wie in Abbildung 5.4 dargestellt, je nach Öffnungswinkel der Laserquelle nur eine teilweise Durchflutung ergibt, wird nur die tatsächlich durchflutete Fläche der Apertur betrachtet. Mit der durchfluteten Fläche der Apertur A_F gilt für die maximal zulässige Energiedichte des Lasers in der Entfernung der Messapertur

$$\omega_P = \frac{E_P}{A_F}. \quad (5.10)$$

Aus der Pulsdauer und dem Öffnungswinkel der Abstrahlung kann im Weiteren die maximale Laserleistung berechnet werden. Für die Laserleistung gilt

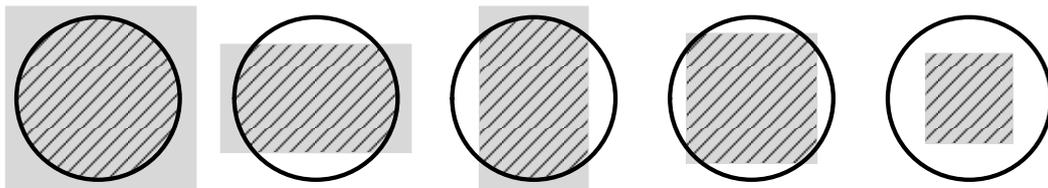


Abbildung 5.4: Mögliche Durchflutungsstufen der Apertur

$$P_L \leq \omega_P \frac{d^2 4 \arcsin\left(\sin\left(\frac{\alpha_H}{2}\right) \sin\left(\frac{\alpha_V}{2}\right)\right)}{T_P} \quad (5.11)$$

mit der Pulsdauer T_P . Weiterhin darf die mittlere Leistung einer Folge von Pulsen der Gesamtdauer T_{PT} den Grenzwert für einen Puls der Dauer T_{PT} nicht überschreiten. Bei Annahme einer kontinuierlichen Messung geht T_{PT} gegen unendlich. Für die Berechnung wird der höchste Zeitwert in [62] gewählt. Für die AEL gilt

$$P_{PT} = 7 \cdot 10^{-4} C_4 C_6 C_7 T_2^{-\frac{1}{4}} \text{ W} . \quad (5.12)$$

mit C_4 nach (5.7), C_6 nach (5.8), $C_7 = 1$ [62] und

$$T_2 = 10 \cdot 10^{\frac{\beta - \beta_{\min}}{98,5}} \text{ für } \beta_{\min} < \beta < \beta_{\max} . \quad (5.13)$$

Bei hohen Zeitdauern wird die AEL nicht in Joule sondern in Watt angegeben, womit die Dauer der Pulschette irrelevant ist. Analog zu (5.10) lässt sich durch Division der Leistung durch die durchflutete Fläche der Apertur eine Leistungsdichte bestimmen, aus der sich die maximale Laserleistung zu

$$P_L \leq P_{PT} \frac{d^2 4 \arcsin\left(\sin\left(\frac{\alpha_H}{2}\right) \sin\left(\frac{\alpha_V}{2}\right)\right)}{A_F \Gamma_{DC}} \quad (5.14)$$

mit dem Tastverhältnis Γ_{DC} ergibt. Als dritte Bedingung darf die Energie eines Pulses nicht den Grenzwert für einen Puls unter Berücksichtigung eines Korrekturfaktors, welcher von der Anzahl der Pulse innerhalb der Pulsnettendauer abhängt, überschreiten. Nach [62] kann für die Kettendauer ein maximaler Wert von T_2 nach (5.13) angenommen werden. Für die AEL gilt nach [62]

$$E_{PC} = E_P N_P^{-\frac{1}{4}} \quad (5.15)$$

mit der Anzahl der Pulse N_P , welche gegeben ist durch

$$N_P = \frac{T_2 \Gamma_{DC}}{T_P}. \quad (5.16)$$

Für $N_P > 1$ gilt $E_{PC} < E_P$ und die maximale Laserleistung kann nach (5.11) mit E_{PC} statt E_P bestimmt werden. Weiterhin stellt sich die Frage, ob die mittlere Leistung einer Pulschette oder die gewichtete Energie eines Einzelpulses die Grenze für die maximale Laserleistung definiert. Die mittlere Leistung ist begrenzend, wenn

$$P_{PT} < E_{PC} \frac{\Gamma_{DC}}{T_P} = E_P \left(\frac{\Gamma_{DC}^3}{T_P^3 T_2} \right)^{\frac{1}{4}} \quad (5.17)$$

gilt. Mit (5.6) und (5.12) folgt

$$\frac{\Gamma_{DC}}{T_P} = f_L > 53,14 \text{ kHz} \quad (5.18)$$

mit der Pulsemissionsrate des Lasers f_L .

5.3 Hintergrundstrahlung

Bei einer Vielzahl von Anwendungen ist der Sensor zusätzlicher Strahlung ausgesetzt, die einen negativen Einfluss auf die Leistungsfähigkeit des Sensorsystems ausübt. Im Automotivbereich ist die Sonne die größte Quelle von Hintergrundstrahlung aufgrund ihrer hohen Leistungsdichte im relevanten Bereich der Wellenlänge. Zur Abschätzung der durch Sonneneinstrahlung generierten Ereignisrate wird das in Abbildung 5.5 dargestellte elektromagnetische Leistungsdichtespektrum der Sonne auf der Erdoberfläche AM1.5G [63] betrachtet. Diese Leistungsdichte tritt bei wolkenlosem Himmel und einer senkrechten Einstrahlung auf und entspricht einer Beleuchtungsstärke von 100 klx. Für die Leistungsdichte in der Szene gilt

$$S_B(\lambda) = S_{AM1.5G}(\lambda) \eta_B \quad (5.19)$$

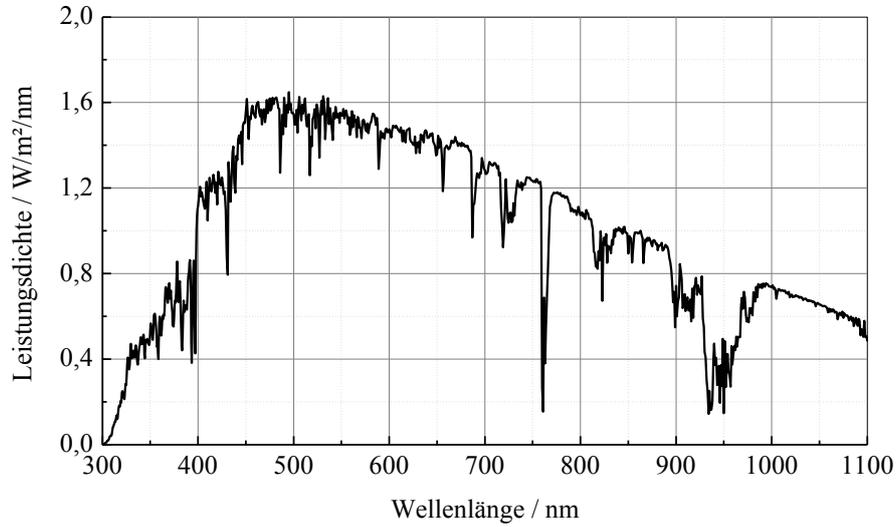


Abbildung 5.5: Leistungsdichtespektrum AM1.5G der Sonne nach [63]

mit dem Faktor η_B zur Anpassung der Lichtstärke an verschiedene Szenarien. Die weitere Berechnung folgt Abschnitt 5.1 mit $S_B(\lambda)$ anstelle von S_L . Zur Reduktion der Hintergrundstrahlung werden optische Bandpassfilter mit einer Mittenwellenlänge entsprechend des verwendeten Lasers eingesetzt. Um diese in der Berechnung zu berücksichtigen, wird die spektrale Leistungsdichte $S_B(\lambda)$ mit der Transmissionskennlinie des Filters $T_F(\lambda)$ multipliziert. Da sich im Gegensatz zu einem idealen Laser das Hintergrundlicht über einen breiten Wellenlängenbereich erstreckt, ergibt sich die erfasste Photonenerate durch Integration über die Wellenlänge entsprechend

$$r_B = \frac{R_T T_O T_{Air} A_{Pix} \eta_{FF}}{4f_{\#}^2 hc} \int_0^{\infty} S_B(\lambda) T_F(\lambda) \eta_{PDE}(\lambda) \lambda d\lambda. \quad (5.20)$$

Unter der Annahme eines idealen Bandpassfilters mit der Bandbreite $2\Delta\lambda$ und der Mittenwellenlänge λ_C folgt

$$r_B = \frac{R_T T_O T_{Air} A_{Pix} \eta_{FF}}{4f_{\#}^2 hc} \int_{\lambda_C - \Delta\lambda}^{\lambda_C + \Delta\lambda} S_B(\lambda) \eta_{PDE}(\lambda) \lambda d\lambda. \quad (5.21)$$

Um das Modell weiter zu präzisieren kann noch die Wellenlängenabhängigkeit der anderen Größen berücksichtigt werden. Der Einfachheit halber wird an dieser Stelle aber darauf verzichtet.

5.4 Monte-Carlo-Simulation

Die Simulation erfolgt in drei Schritten: Simulation einer einzelnen Distanzmessung über eine gegebene Anzahl an Laserpulsen auf der Basis von computergenerierten Zufallszahlen, Bestimmung statistischer Parameter wie Erwartungswert und Varianz aus einer Vielzahl von Distanzmessungen sowie Wiederholung des Ganzen für verschiedene Distanzen zwischen Sensor und Zielobjekt. Für eine einzelne Distanzmessung werden im Fall des direkten Verfahrens die Zeiten zwischen der Emission des Laserpulses und der Detektion des ersten Ereignisses über mehrere Laserpulse in einem Histogramm akkumuliert. Aus diesem wird anschließend mittels eines Algorithmus die Distanzinformation ermittelt. Analog werden beim indirekten Verfahren die einfallenden Photonen in den einzelnen Zeitfenstern über viele Laserpulse aufsummiert und anschließend die Distanz aus den Zählwerten der Fenster mit der entsprechenden Formel bestimmt. In beiden Fällen werden die Zwischenankunftszeiten der Photonen mittels computergenerierter exponentialverteilter Zufallszahlen erzeugt. Der Erwartungswert entspricht dabei den zuvor berechneten Ereignisraten für den Hintergrund bzw. den reflektierten Laserpuls. Die Anzahl der Laserpulse einer einzelnen Distanzmessung ergibt sich aus der Pulsrate des Lasers f_L und der gewünschten Bildrate des Messsystems f_F zu

$$N_P = N_Z = \frac{f_L}{f_F}. \quad (5.22)$$

Entsprechend den vorherigen Berechnungen liefert eine höhere Pulszahl bessere Messergebnisse. Hierfür ist allerdings eine Reduktion der Bildrate, was in zeitkritischen Anwendungen zu Problemen führen kann, oder eine Erhöhung der Laserpulsrate, die in der Regel durch den Laser selbst, die verwendete Treiberstufe oder die Augensicherheit begrenzt ist, notwendig.

5.5 Modellverifikation

Da sowohl die Berechnungen in Kapitel 4 als auch die Monte-Carlo-Simulation auf exponentialverteilten IATs der einfallenden Photonen beruhen, kann durch eine Gegenüberstellung beider Resultate eine erste Verifikation der theoretischen Modelle durchgeführt werden. Für diesen Zweck werden entsprechend Tabelle 5.1 beispielhafte Parameter gewählt. Mit diesen Werten ergeben sich aus (5.5) bzw. (5.21) Photonendetektionsraten von $r_L(d) = 38,441 \text{ GHz m}^2/d^2$ und $r_B = 9,37 \text{ MHz}$. Mit diesen Raten erfolgt im Weiteren die Bestimmung der Standardabweichung des indirekten Verfahrens mit PM über das Varianzmodell nach (4.32) sowie mittels Monte-Carlo-Simulation. Dafür werden eine Pulsdauer von 667 ns, was gleichzeitig der Breite der Messfenster sowie einer Reichweite von 100 m entspricht, und eine Akkumulation über $N_Z = 400$ Einzelmessungen angenommen. Abbildung 5.6 zeigt die relative Standardabweichung der Distanzmessung als Funktion der Distanz. Dabei zeigt sich eine gute Übereinstimmung zwischen Berechnung und Monte-Carlo-Simulation über den dargestellten Distanzbereich.

Tabelle 5.1: Parameter zur Verifikation der theoretischen Berechnungen

Symbol	Wert	Einheit	Beschreibung
P_L	75/25	W	Optische Quellenleistung
α_H	20	Grad	Horizontaler Quellenöffnungswinkel
α_V	0,25	Grad	Vertikaler Quellenöffnungswinkel
$f_{\#}$	7,6/23	-	Blendenzahl
R_T	50	%	Reflektanz der Szene
T_O	50	%	Transmissionsfaktor der Optik
T_{Air}	100	%	Transmissionsfaktor der Luft
A_{Pix}	0,01	mm ²	Pixelfläche
η_{FF}	80	%	Pixelfüllfaktor
η_{PDE}	2	%	Photonendetektionseffizienz
λ	905	nm	Wellenlänge der Laserstrahlung
η_B	0,4/1	-	Faktor der Hintergrundstrahlung
$\Delta\lambda$	40	nm	Filterbandbreite

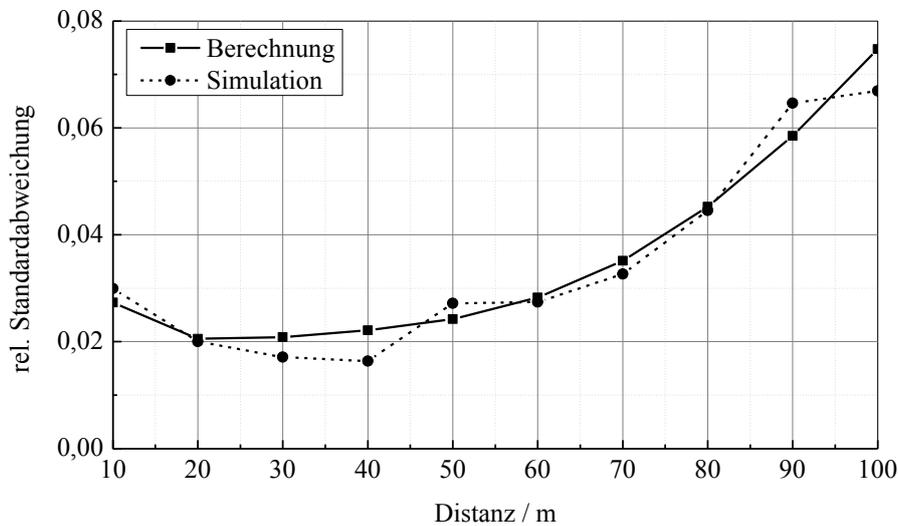


Abbildung 5.6: Berechnete und simulierte Standardabweichung des indirekten Messverfahrens mit Pulsmodulation

Für die Berechnung bzw. Simulation des direkten Verfahrens wird die Blendenzahl auf 23 erhöht, der Hintergrund auf 100 klx angehoben und die optische Leistung auf $P_L = 25 \text{ W}$ abgesenkt, womit Raten von $r_L(d) = 1,414 \text{ GHz m}^2/d^2$ und $r_B = 2,58 \text{ MHz}$ erreicht werden. Die maximale Distanz wird mit 25 m bei einer Breite der Klassen von 1 ns angenommen, wodurch sich die Zahl der Klassen zu 176 ergibt. Weiterhin werden eine Pulsdauer von 10 ns und eine Akkumulation über $N_Z = 400$ Laserpulse gewählt. Die berechneten Werte des Erwartungswertes und der Standardabweichung werden mittels des in Abschnitt 4.3.3 für eine beispielhafte Datenauswertung beschriebenen Verfahrens aus den Wahrscheinlichkeiten der einzelnen Klassen des Histogramms den Maximalwert zu enthalten ermittelt. Aus Gründen der Vergleichbarkeit erfolgt die Monte-Carlo-Simulation mit identischer Auswertung, dabei folgt analog zur Berechnung im Falle eines mehrfachen Maximums keine Distanzbestimmung, stattdessen wird die Messung als ungültig gewertet. Die Anzahl der ungültigen Messungen liegt in der Monte-Carlo-Simulation je nach Distanz zwischen 12 % und 51 %. Die in Abbildung 5.7 dargestellten Kurven zeigen eine gute Übereinstimmung über dem gesamten Distanzbereich, sodass das beschriebene Simulationsmodell ein geeignetes Werkzeug zur Abschätzung der Leistungsfähigkeit verschiedener Messverfahren sowie Konzepten zu deren Optimierung darstellt.

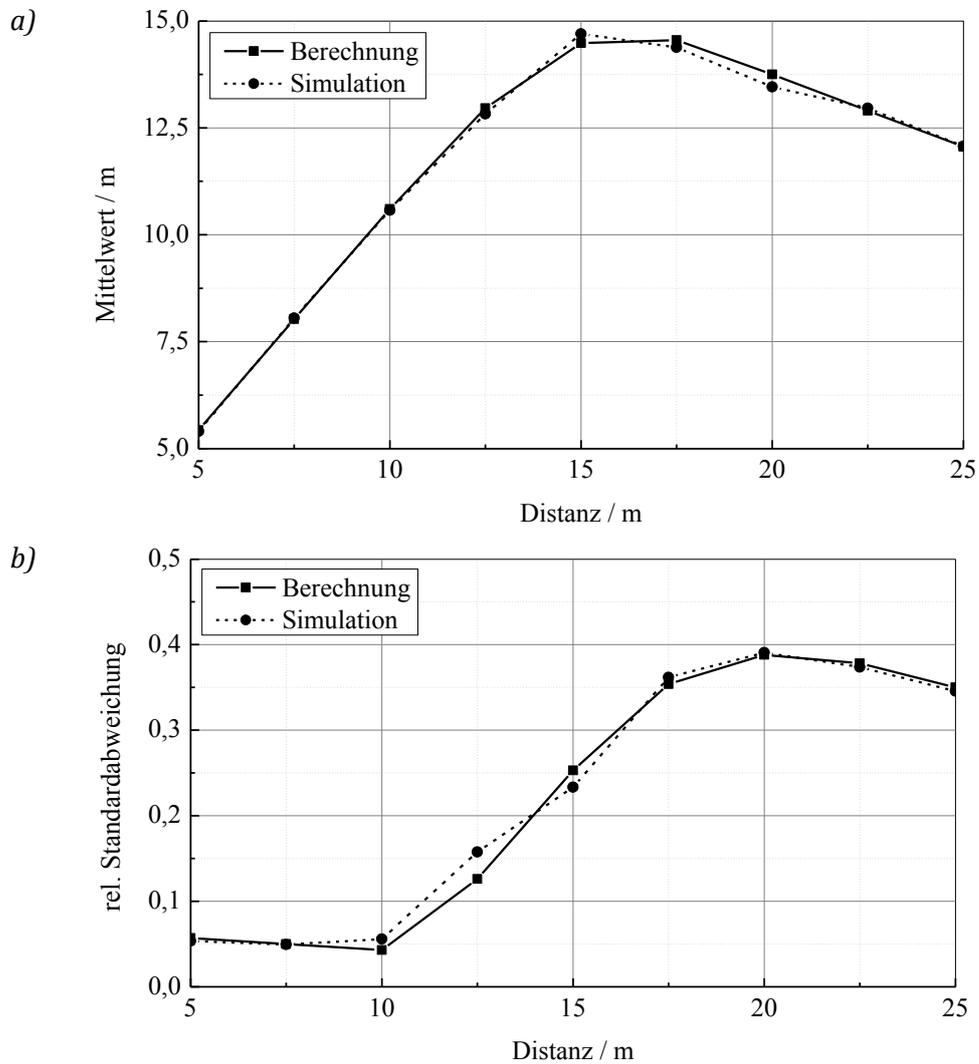


Abbildung 5.7: Berechneter und simulierter Mittelwert (a) und Standardabweichung (b) des direkten Messverfahrens

Im Allgemeinen kann die Bestimmung der statistischen Parameter auch für komplexere Verfahren oder Algorithmen zur Datenauswertung analytisch erfolgen, jedoch erlaubt die Simulation eine Einschätzung dieser ohne eine zeitintensive mathematische Herleitung der Bestimmungsgleichungen.

5.6 Vergleich direktes und indirektes Verfahren

Auf Basis des vorgestellten Simulationsmodells der Verfahren zur laufzeitbasierten Distanzmessung wird im Folgenden ein Vergleich zwischen dem direkten und indirekten Verfahren durchgeführt. Für diesen Zweck wird zunächst die zulässige optische Emissionsleistung der Laserquelle bestimmt. Zu deren Berechnung werden die Quellenöffnungswinkel aus Tabelle 5.1, eine Laserpulsrate von 100 kHz bei gepulsten Verfahren, eine punktförmige Quelle, d. h. $\beta = \beta_{\min}$, sowie eine Reichweite von 25 m, was 167 ns entspricht, angenommen.

Im indirekten Verfahren mit CW ist der Laser dauerhaft aktiv. In Anlehnung an [38] wird ein Tastverhältnis von 50 % bei einer Periodendauer von 167 ns zur Berechnung verwendet. Erwartungsgemäß ergibt sich in diesem Fall eine sehr geringe zulässige Leistung von 10,2 mW. Im indirekten Verfahren mit PM ist die Pulsdauer durch die Reichweite gegeben, weshalb sich durch die lange Pulsdauer von 167 ns für dieses Verfahren ebenfalls eine geringe Leistung von 305,4 mW ergibt. Im direkten Verfahren ist die Pulsdauer unabhängig von der Reichweite und kann folglich sehr kurz gewählt werden. In Anlehnung an diverse Veröffentlichungen [11], [12] wird ein realitätsnaher Wert von 10 ns gewählt, womit sich eine zulässige Leistung von 5,1 W ergibt, was um den Faktor 17 bzw. 500 über den indirekten Verfahren liegt.

Mit den zulässigen Leistungen der Verfahren erfolgt die Bestimmung der Standardabweichung. Entsprechend den Betrachtungen in Abschnitt 4.2 ist für das indirekte Verfahren unter Vernachlässigung der Totzeit eine möglichst hohe Photonenrate zu bevorzugen. Unter Totzeiteinfluss hingegen sättigt die Zählrate der SPADs, was zur Abnahme des SBR führt. Wird analog zu Abschnitt 4.3.3 das SNR im indirekten Verfahren mit PM berechnet, ergibt sich

$$a_{\text{PM}} = \frac{r_L \sqrt{T_W N_Z (1 + r_{\text{LB}} T_T)^3}}{(1 + (r_B + r_{\text{LB}}) T_T + r_B r_{\text{LB}} T_T^2) \sqrt{r_{\text{LB}}}}, \quad (5.23)$$

was unter der Annahme eines konstanten SBR eine monoton steigende Funktion von r_B bzw. r_L ist, welche mit steigenden Photonenraten ebenfalls sättigt. Zur Festlegung der Photonenraten und damit der Blendenzahl wird der Punkt des SNR gewählt, bei welchem es um 50 % gegenüber dem Maximalwert abgefallen

ist und dementsprechend eine weitere Erhöhung der Raten keine signifikanten Vorteile mehr bietet. Dieser Punkt liegt näherungsweise bei der Hälfte der maximalen Zählrate und soll im Falle maximalen Hintergrundes, welcher bei einer Beleuchtungsstärke von 100 klx und einer Reflektanz von 100 % auftritt, nicht überschritten werden. Bei einer angenommenen Totzeit von 20 ns ergibt sich diese zu 25 MHz, was zusammen mit den in Tabelle 5.1 gegebenen Sensorparametern eine Blendenzahl von 10,47 und eine distanz- und leistungsabhängige Photonenerate des Lasers von $r_L(d, P_L) = P_L 546,2 \text{ MHz m}^2/d^2\text{W}$ ergibt.

Im direkten Verfahren muss die Photonenerate durch Anpassung der Objektivblende so eingestellt werden, dass die Wahrscheinlichkeit von Fehldetektionen gemäß (4.66) einen gewissen Grenzwert nicht überschreitet. Allerdings ist diese Wahrscheinlichkeit eine Funktion des Hintergrundes, der Reflektanz und der Distanz. Soll die Wahrscheinlichkeit einen Wert von 80 % bei maximaler Hintergrundintensität und Distanz nicht überschreiten, ergibt sich eine erlaubte Photonenerate von 9,63 MHz, eine Blendenzahl von 16,9 und eine Photonenerate des Lasers von $r_L(d, P_L) = P_L 209,5 \text{ MHz m}^2/d^2\text{W}$.

Für die indirekten Verfahren erfolgt die Berechnung der Abweichung auf Basis der theoretischen Modelle (4.32) bzw. (4.48), während für das direkte Verfahren eine Simulation durchgeführt wird. Die Bestimmung der Distanz aus dem Histogramm erfolgt durch eine Mittelwertfilterung mit anschließender Bestimmung der Position des Maximums. Die Breite des Filters wird entsprechend der Pulsbreite auf 10 ns festgelegt und dient der Reduktion der Varianz der einzelnen Klassen, welche mit einer Breite von 500 ps angenommen werden. Die Anzahl der Einzelmessungen N_z ergibt sich aus der Bildrate von 25 Hz mittels (5.22) zu 4000 bei den gepulsten Verfahren. Für die indirekte Messung mit CW ergibt sich aus der Periodendauer von 167 ns und der Bildrate von 25 Hz ein Wert von 239.520 für die Anzahl an Einzelmessungen.

Abbildung 5.8 zeigt die berechnete bzw. simulierte Standardabweichung der indirekten sowie des direkten Verfahrens für eine Beleuchtungsstärke des Hintergrundlichtes von 50 klx. Dabei zeigt sich für das direkte Verfahren eine deutlich geringere Abweichung als für die beiden indirekten. Ab einer Distanz von 5 m zeigen die indirekten Verfahren näherungsweise identische Werte, da die höhere Anzahl der Messungen die geringe zulässige Emissionsleistung im

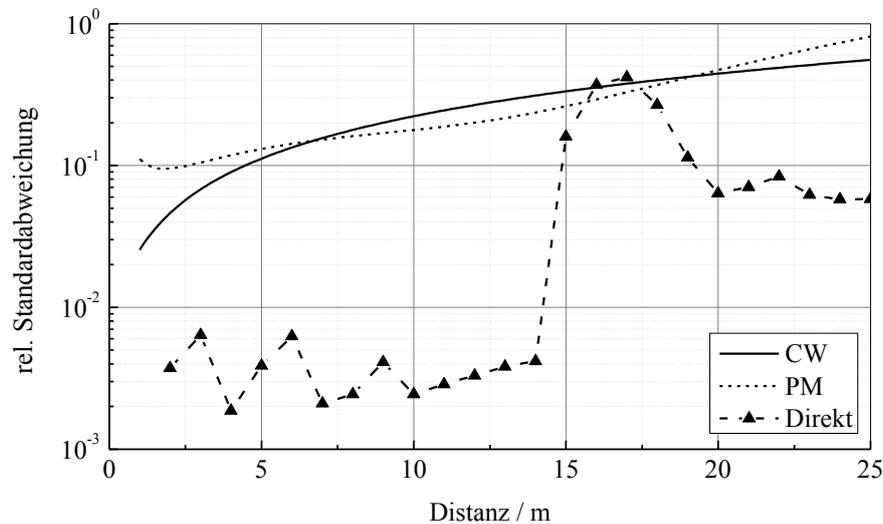


Abbildung 5.8: Standardabweichung des indirekten Verfahrens mit CW und PM sowie des direkten Verfahrens

CW-Verfahren kompensiert. Das direkte Verfahren zeigt bis zu einer Distanz von 14 m eine Standardabweichung von unter 1 % und steigt anschließend stark an. Dieser Anstieg lässt sich durch die PDF des ersten Ereignissen gemäß (4.69) erklären: Mit zunehmender Distanz nimmt die Höhe des Laserpulses im Histogramm ab, während die Werte der vorderen Klassen aufgrund des Hintergrundlichtes konstant bleiben. Sobald sich die Höhen annähern, steigt die Wahrscheinlichkeit, dass sich das Maximum des Histogramms an dessen Beginn befindet. Folglich steigt ab diesem Punkt die Varianz der Distanzmessung. Mit weiter steigender Messdistanz sinkt die Höhe des Pulses weiter, sodass häufiger geringe Distanzwerte aufgrund des Hintergrundlichtes gemessen werden und die Varianz wieder abnimmt. Aus (4.69) bestimmt sich die Distanz, bei welcher beide Höhen identisch sind, zu 17,35 m. Bei diesem Wert zeigt die Abweichung in Abbildung 5.8 den höchsten Wert.

Unter Berücksichtigung der Augensicherheit zeigt das direkte Messverfahren den geringsten Messfehler. Dieser Vorteil nimmt für höhere Reichweiten zu, da in den indirekten Verfahren längere Pulsdauern benötigt werden, welche in geringeren zulässigen Emissionsleistungen resultieren, was insbesondere bei hohem Hintergrundlicht die Messvarianz erhöht.

6 Konzepte zur Unterdrückung von Hintergrundlicht

In diesem Kapitel werden auf Grundlage der Betrachtungen in den vorhergehenden Kapiteln neue Konzepte der SPAD-basierten Laufzeitmessung entworfen. Das in der Einleitung erwähnte Problem des geringen Füllfaktors aktueller 2D-Arrays spielt zunächst keine Rolle, da ein Zeilensensor entworfen wird. Bei diesem ist es möglich, die In-Pixel-Elektronik außerhalb der Zeile zu platzieren, was den Füllfaktor innerhalb der Zeile erhöht. Das Hauptaugenmerk für den Entwurf neuer Konzepte liegt in der Reduktion des Einflusses der Hintergrundstrahlung. Diese kann die durch aktive Beleuchtung zu erwartenden Photonenraten im Pixel deutlich überschreiten und die Messung mit bestehenden Methoden verschlechtern oder gar unmöglich machen. Ziel der neuen Konzepte ist es, durch intelligente Ausnutzung von Photonenstatistik oder Signalverarbeitung das SNR zu erhöhen und die Güte der Messung zu verbessern.

6.1 Photonenkorrelation

In [11] wird die Detektion korrelierter Photonen in einem Scanner-basierten direkten Laufzeitsensor zur Reduktion des Hintergrundeinflusses genutzt. Dabei findet ein Ereignis nur dann statt, wenn nach der Detektion eines ersten Photons innerhalb einer vorgegebenen Korrelationszeit ein weiteres Photon

registriert wird. Unter Ausnutzung der räumlichen und zeitlichen Korrelation zwischen Photonen, der sogenannten Koinzidenz, kann eine verbesserte Unterdrückung der Hintergrundstrahlung erreicht werden, da deren Photonen, im Gegensatz zum Lasersignal, nicht zeitlich konzentriert sind [11]. Um Korrelationszeiten unterhalb der Totzeit der SPAD zu realisieren, werden im Sensor zwei oder mehr SPADs je Pixel verwendet. Ein Ereignis wird nur gezählt, wenn innerhalb einer festgelegten Zeit mindestens zwei SPADs ein einfallendes Photon registrieren. Im Allgemeinen kann die Zahl der innerhalb der Korrelations- bzw. Koinzidenzzeit erforderlichen Photonen, im Weiteren als Koinzidenztiefe bezeichnet, sowie die Zeit selbst beliebig gewählt werden. Im Folgenden wird die Rate solcher Ereignisse als Funktion der Rate detektierter Einzelphotonen berechnet. Auf dieser Basis wird untersucht, wie sich die Koinzidenz zur Unterdrückung des Hintergrundlichtes und so zur Steigerung der Reichweite und Zuverlässigkeit eines lichtlaufzeitbasierten Messsystems einsetzen lässt.

6.1.1 Modell

Die Statistik von Koinzidenzereignissen wurde in [64] untersucht. Dabei wurde angenommen, dass zwei durch parametrische Abwärts wandlung erzeugte Photonenstrahlen von je einer SPAD erfasst werden. Nur wenn nach einer Detektion im primären Strahl innerhalb einer festen Zeit auch ein Photon im sekundären Strahl erkannt wird, wird das Ereignis gezählt. Zur Beschreibung der Korrelation zwischen beiden Strahlen wurde ein Parameter eingeführt, welcher den Grad der Korrelation angibt. Nach analoger Vorgehensweise wird im Folgenden die Rate der Ereignisse bei der Verwendung von mehreren SPADs als Subpixel bestimmt. Für diesen Zweck wird angenommen, dass das gesamte Pixel eine Photonenrate von r_{Pixel} erfährt. Die Ereignisrate in jedem Subpixel ergibt sich aus der PDE η_{PDE} , der Anzahl der Subpixel N_S und dem Füllfaktor η_{FF} zu

$$r = \frac{r_{\text{Pixel}} \eta_{\text{PDE}} \eta_{\text{FF}}}{N_S}. \quad (6.1)$$

Die Photonenraten der einzelnen Subpixel werden als unabhängig angenommen und definieren den Erwartungswert der exponentialverteilten IATs. Zur Bestimmung der Ereignisrate werden die PDFs der auf dem Array detektierten

Photonen betrachtet und daraus der zeitliche Erwartungswert bestimmt. Die Verteilungen der Photonen IAT und der FPAT einer einzelnen totzeitbehafteten SPAD wurden bereits in Abschnitt 4.2.1.2 hergeleitet und werden an dieser Stelle der Vollständigkeit halber wiederholt. Die IAT der Photonen ist durch eine um die Totzeit verschobene Exponentialverteilung entsprechend

$$P_{IA,1}(t) = r \exp(-r(t - T_T)) \text{ für } t \geq T_T \quad (6.2)$$

gegeben. Für die FPAT gilt

$$P_{1,1}(t) = \begin{cases} r_T & \text{für } 0 \leq t < T_T \\ r_T \exp(-r(t - T_T)) & \text{für } t \geq T_T \end{cases} \quad (6.3)$$

mit der reduzierten Ereignisrate r_T und der Totzeit T_T . Mittels dieser Verteilungen wird im Folgenden die IAT sowie die FPAT eines Arrays bestehend aus N_S SPADs bestimmt. Die FPAT ergibt sich aus dem Produkt der Wahrscheinlichkeit, dass eine SPAD ein Photon detektiert, während alle anderen keines erfassen zu

$$P_{1,N_S}(t) = N_S P_{1,1}(t) \left(1 - \int_0^t P_{1,1}(\tau) d\tau \right)^{N_S-1}. \quad (6.4)$$

Einsetzen von (6.3) liefert

$$P_{1,N_S}(t) = \begin{cases} N_S r_T (1 - r_T t)^{N_S-1} & \text{für } 0 \leq t < T_T \\ N_S r_T e^{-r(t-T_T)} \beta & \text{für } t \geq T_T \end{cases} \quad (6.5)$$

mit

$$\beta = \left(1 - r_T T_T - \frac{r_T}{r} (1 - e^{-r(t-T_T)}) \right)^{N_S-1}. \quad (6.6)$$

Zur Bestimmung der Koinzidenzereignisrate muss weiterhin die erwartete Zeit zwischen der ersten Photonendetektion und der n -ten Detektion bekannt sein, wobei n die Koinzidenztiefe repräsentiert. Die PDF der weiteren Photonen lässt sich mittels Faltung der vorhergehenden Detektion mit der Verteilung der IAT berechnen. Aus diesem Grund wird zunächst die Verteilung der IAT bestimmt.

Da ein Array aus N_S SPADs betrachtet wird, müssen zwei Möglichkeiten einer nachfolgenden Photonendetektion betrachtet werden: Im ersten Fall wird das nächste Photon von derselben SPAD detektiert, welche auch das erste Photon registriert hat. In diesem Fall ist die IAT der empfangenden SPAD durch (6.2) gegeben, während die anderen SPADs eine Verteilung gemäß (6.3) aufweisen, da für diese der Zeitpunkt der hier betrachteten Photonendetektion zufällig ist. Im zweiten Fall wird das nachfolgende Photon von einer anderen SPAD, als der welche das erste Photon detektiert hat, erfasst. Die Gesamtverteilung der IAT ergibt sich durch die Summe beider Fälle zu

$$P_{IA,N_S}(t) = P_{IA,1}(t) \left(1 - \int_0^t P_{1,1}(\tau) d\tau \right)^{N_S-1} + (N_S - 1) P_{1,1}(t) \left(1 - \int_0^t P_{IA,1}(\tau) d\tau \right) \left(1 - \int_0^t P_{1,1}(\tau) d\tau \right)^{N_S-2}. \quad (6.7)$$

Durch das Einsetzen der Verteilungen nach (6.2) und (6.3) ergibt sich

$$P_{IA,N_S}(t) = \begin{cases} (N_S - 1) r_T (1 - r_T t)^{N_S-2} & \text{für } 0 \leq t < T_T \\ N_S r_T e^{-N_S r (t - T_T)} \left(\frac{r_T}{r} \right)^{N_S-2} & \text{für } t \geq T_T \end{cases} \quad (6.8)$$

Mit der Verteilung der IAT wird im Folgenden die PDF der n -ten Photonendetektion bestimmt. Ein Koinzidenzereignis findet immer dann statt, wenn die Zeit zwischen dem ersten und dem n -ten detektierten Photonen maximal der festgelegten Koinzidenzzeit entspricht. Für die Berechnung folgt daraus, dass bei einer Koinzidenztiefe von n die Verteilung des k -ten Photons mit $k = n - 1$ relevant ist. Diese lässt sich mittels (6.8) berechnen zu

$$P_{k,N_S}(t) = \int_0^{\infty} P_{k-1,N_S}(\tau) P_{IA,N_S}(t - \tau) d\tau \quad (6.9)$$

für $k = 2, 3, \dots$ und $P_{1,N_S}(t) = P_{IA,N_S}(t)$. Entsprechend [64] werden zur Berechnung der Koinzidenzrate zwei Fälle betrachtet: Im ersten Fall überschreitet die

Zeit zwischen der ersten und n -ten Detektion die vorgegebene Koinzidenzzeit, folglich wird kein Koinzidenzereignis registriert. Trotzdem muss die vergangene Zeit in die Berechnung der Ereignisrate miteinbezogen werden. Diese ist gegeben durch die Wartezeit bis zur ersten Photonendetektion plus der Koinzidenzzeit. Die Wartezeit bis zur ersten Detektion lässt sich aus der Verteilung des ersten Photons gemäß (6.5) bestimmen. Für deren Erwartungswert gilt

$$T_1 = E(t_1) = \int_0^{\infty} P_{1,N_S}(t)t dt. \quad (6.10)$$

Im zweiten Fall findet ein Koinzidenzereignis statt. Hier vergeht die Zeit bis zur ersten Detektion nach (6.10) zuzüglich der Zeit bis zur k -ten Detektion, welche unterhalb der Koinzidenzzeit ϑ liegt. Der Erwartungswert dieser Zeit wird aus der PDF der k -ten Photonendetektion entsprechend

$$T_2 = E(t_2) = \frac{1}{\int_0^{\vartheta} P_{k,N_S}(\tau) d\tau} \int_0^{\vartheta} P_{k,N_S}(\tau)\tau d\tau \quad (6.11)$$

bestimmt. Da in diesem Fall nicht über die komplette Verteilung integriert wird, ist eine Normierung des Erwartungswertes erforderlich. Die Koinzidenzrate ergibt sich aus der Summe der verstrichenen Zeiten in beiden Fällen gewichtet mit der jeweiligen Wahrscheinlichkeit zu

$$r_{\text{Coin}}^{-1} = T_1 + (\vartheta + r_{\text{Coin}}^{-1}) \int_{\vartheta}^{\infty} P_{k,N_S}(\tau) d\tau + T_2 \int_0^{\vartheta} P_{k,N_S}(\tau) d\tau. \quad (6.12)$$

Falls kein Koinzidenzereignis stattfindet, muss die Zeit zwischen zwei Ereignissen erneut vergehen, weshalb der Kehrwert der Rate auch in der rechten Seite der Gleichung auftaucht. Umgeformt nach der Rate ergibt sich

$$r_{\text{Coin}} = \frac{(1 - \int_{\vartheta}^{\infty} P_{k,N_S}(\tau) d\tau)}{T_1 + \vartheta \int_{\vartheta}^{\infty} P_{k,N_S}(\tau) d\tau + T_2 \int_0^{\vartheta} P_{k,N_S}(\tau) d\tau}. \quad (6.13)$$

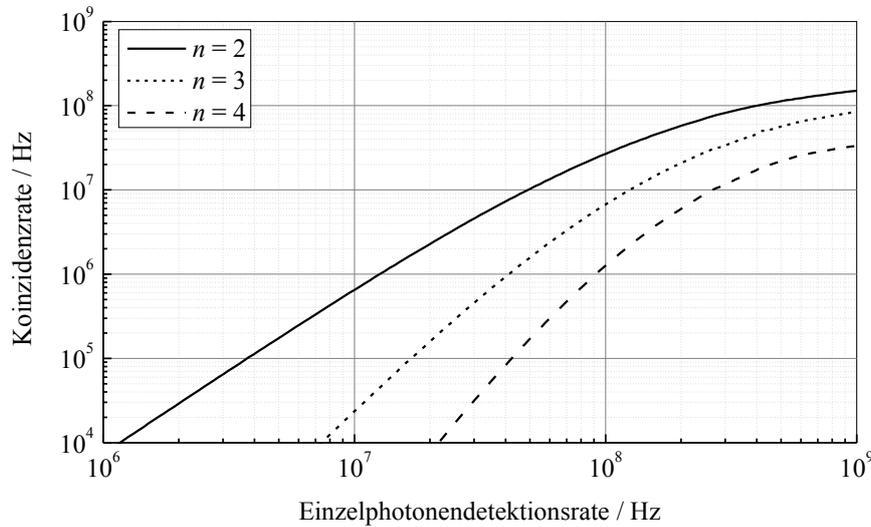


Abbildung 6.1: Berechnung der Koinzidenzrate als Funktion der Einzelphotonendetectionsrate des Pixels für verschiedene Koinzidenztiefen

Den Erwartungen entsprechend steigt die Koinzidenzrate mit höherer Koinzidenzwahrscheinlichkeit im Zähler sowie geringeren Wartezeiten im Nenner.

Im doppeltlogarithmischen Maßstab zeigt $r_{\text{Coin}} = f(N_S r)$ eine Steigung von n für $r\vartheta \ll 1$ was einer Proportionalität $r_{\text{Coin}} \propto (N_S r)^n$ entspricht. Abbildung 6.1 zeigt beispielhaft die Koinzidenzrate eines Arrays bestehend aus 4 SPADs mit einer Totzeit von 10 ns für eine Koinzidenzzeit von 10 ns und eine Koinzidenztiefe von 2 bis 4. Je nach software- oder hardwaretechnischer Umsetzung der Koinzidenzdetektion kann diese einen Einfluss auf die resultierende Ereignisrate haben, welche hier jedoch nicht weiter betrachtet wird.

6.1.2 Bewertung

Die Nutzung von Photonenkorrelation beeinflusst die Rate der erfassten Ereignisse. In diesem Abschnitt werden die Vor- und Nachteile analysiert.

6.1.2.1 Signal-Hintergrund-Verhältnis

Ein Vorteil der Photonenkoinzidenz ist die Verbesserung des SBR aufgrund der exponentiellen Zunahme der Wahrscheinlichkeit eines Koinzidenzereignisses. Dies soll an einem Zahlenbeispiel basierend auf der in Abbildung 6.1 gezeigten

Kurve für eine Koinzidenztiefe von 2 veranschaulicht werden. Ohne den Einsatz von Koinzidenz ergibt sich mit Detektionsraten von je 10 MHz für den Laserpuls und das Hintergrundlicht eine SBR von 1. Wird auf dieses Szenario Koinzidenz angewendet, ergibt sich für den Hintergrund eine Ereignisrate von 650 kHz. Während des Laserpulses liegt eine Rate von 20 MHz am Sensor an, was wiederum in einer Ereignisrate von 2,28 MHz resultiert. Nach der Subtraktion der Hintergrundrate bleiben davon 1,63 MHz, was ein SBR von 2,51 ergibt. Allgemein ergibt sich ohne Sättigungseffekte ein Zusammenhang gemäß

$$\frac{r_{\text{Coin,L}}}{r_{\text{Coin,B}}} = \left(\frac{r_{\text{L}}}{r_{\text{B}}} + 1 \right)^n - 1 \quad (6.14)$$

mit der Koinzidenztiefe n . Im gezeigten Beispiel ergäbe dies ein SBR nach der Koinzidenz von 3, was über dem berechneten Wert von 2,51 liegt. Die Zunahme des SBR wirkt sich positiv auf die Laufzeitmessung aus, da sich das reflektierte Lasersignal deutlicher von Hintergrund abhebt, was insbesondere im direkten Verfahren eine erhöhte Reichweite erwarten lässt. Im indirekten Verfahren führt die reduzierte Zählrate zu einer erhöhten relativen Varianz der einzelnen Zeitfenster, was dem verbesserten SBR entgegenwirkt.

6.1.2.2 Dynamikbereich

Aufgrund der exponentiellen Zunahme der Koinzidenzwahrscheinlichkeit erhöht sich der Dynamikbereich der resultierenden Ereignisrate. Werden die Sättigungseffekte vernachlässigt, gilt für den Dynamikbereich nach Koinzidenz

$$\frac{r_{\text{Coin,max}}}{r_{\text{Coin,min}}} = \left(\frac{r_{\text{max}}}{r_{\text{min}}} \right)^n . \quad (6.15)$$

Im direkten Verfahren kann dies dazu führen, dass bereits eine leichte Änderung der Hintergrundintensität eine starke Änderung der Ereignisrate mit sich bringt und die Distanzmessung deutlich verschlechtert. Um dem vorzubeugen müssen die Parameter der Koinzidenz so gewählt werden, dass auch bei maximal zu erwartender Hintergrundintensität dessen Ereignisrate ein tolerierbares Niveau nicht überschreitet. Dies hat jedoch den Nachteil, dass im Fall niedriger Reflektanz des Zielobjektes der reflektierte Laserpuls ebenfalls stark gedämpft

wird, was der Güte der Messung abträglich ist. Zur Lösung dieses Konfliktes können die Parameter der Koinzidenz im laufenden Betrieb angepasst werden. Das Prinzip der adaptiven Koinzidenz wird im folgenden Abschnitt genauer betrachtet. Im indirekten Verfahren sorgt der erhöhte Dynamikbereich ebenfalls für eine Verstärkung der Differenz bei unterschiedlicher Reflektanz. Generell ist für das indirekte Verfahren eine hohe Zählrate erstrebenswert, welche jedoch durch die Totzeit der SPADs limitiert wird.

6.1.3 Adaptive Koinzidenz

Um die Koinzidenz in der direkten Laufzeitmessung vorteilbringend einsetzen zu können, müssen die Parameter Koinzidenzzeit und -tiefe sinnvoll gewählt werden. Hierfür muss zunächst bestimmt werden, bei welcher Ereignisrate des Hintergrundlichtes die Messung ein optimales Ergebnis liefert. Grundsätzlich müssen zwei Bedingungen erfüllt sein, damit eine Messung überhaupt möglich ist: Zum einen muss der am Zielobjekt reflektierte Laserpuls detektiert werden. Dazu ist es notwendig, dass innerhalb der Pulsdauer die Wahrscheinlichkeit einer Ereignisdetektion größer null ist. Zur Berechnung wird die Verteilung der FPAT mit der Laufzeit T_{TOF} und der Pulsbreite T_{P} entsprechend

$$P_{\text{Puls}} = \int_{T_{\text{TOF}}}^{T_{\text{TOF}}+T_{\text{P}}} P_1(t) dt > 0 \quad (6.16)$$

integriert. Zweitens muss die Wahrscheinlichkeit, dass bereits vor der Ankunft des reflektierten Pulses ein Ereignis durch Hintergrundlicht ausgelöst wird, kleiner eins sein. Die Wahrscheinlichkeit ergibt sich analog durch Integration der FPAT von $t = 0$ bis zur Ankunftszeit des Pulses T_{TOF} .

Da die einfallenden Raten einer gewissen Dynamik aufgrund wechselnder Reflexionscharakteristika der Zielobjekte, Distanz und Hintergrundbedingungen unterliegen, können feste Koinzidenzparametern zu Problemen führen: Bei zu geringer Dämpfung der Raten sorgt der hohe Hintergrund für ein Ereignis noch vor der Ankunft des reflektierten Laserpulses, wodurch dieser nicht mehr detektiert werden kann. Im umgekehrten Fall kann eine zu starke Dämpfung zu

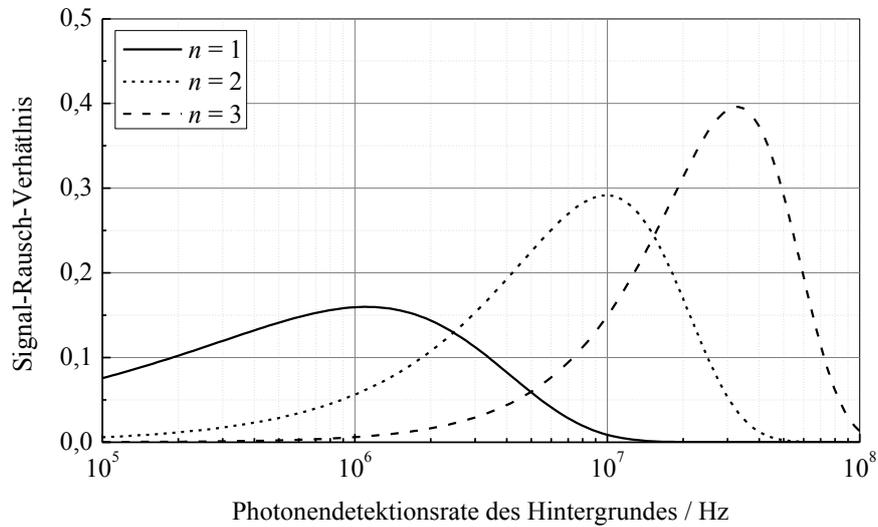


Abbildung 6.2: SNR des Histogramms nach (4.78)

einer Unterdrückung des Pulses führen, was insbesondere bei größeren Distanzen problematisch ist. Das Ziel der adaptiven Koinzidenz ist es, durch Anpassung der Koinzidenzparameter die Hintergrundrate möglichst konstant auf dem festgelegten Sollwert zu halten. Als Regelgröße dient die Hintergrundereignisrate, welche aus dem Histogramm oder durch Messung ermittelt wird.

Betrachtet man das in Abbildung 6.2 dargestellte SNR des Histogramms nach (4.78) für konstantes SBR ohne Koinzidenz ($n = 1$) für beispielhafte Werte, zeigt sich, dass es einer Normalverteilung ähnelt. Dies resultiert in einer Veränderung der Messleistung über dem Reflexionskoeffizienten des Zielobjektes. In Idealfall sollte das SNR über dem gesamten Dynamikbereich konstant sein. Durch die Verwendung von Koinzidenz reduzieren sich die Ereignisraten, wodurch eine Verbesserung der Messung bei höheren Hintergrundereignisraten erreicht wird. Dies zeigt sich durch eine Verschiebung des Maximums im SNR wie in Abbildung 6.2 für eine Koinzidenztiefe von $n = 2$ und 3 dargestellt. Durch die Zunahme des SBR zeigt sich zudem eine Erhöhung des Maximums. Die mit der Totzeit einhergehende Sättigung der Detektionsraten der SPADs wurde dabei vernachlässigt, da bei der direkten Messmethode nur das erste erfasste Ereignis relevant ist. Durch geeignete Umschaltung zwischen verschiedenen Koinzidenztiefen lässt sich der Dynamikbereich deutlich erweitern. Setzt

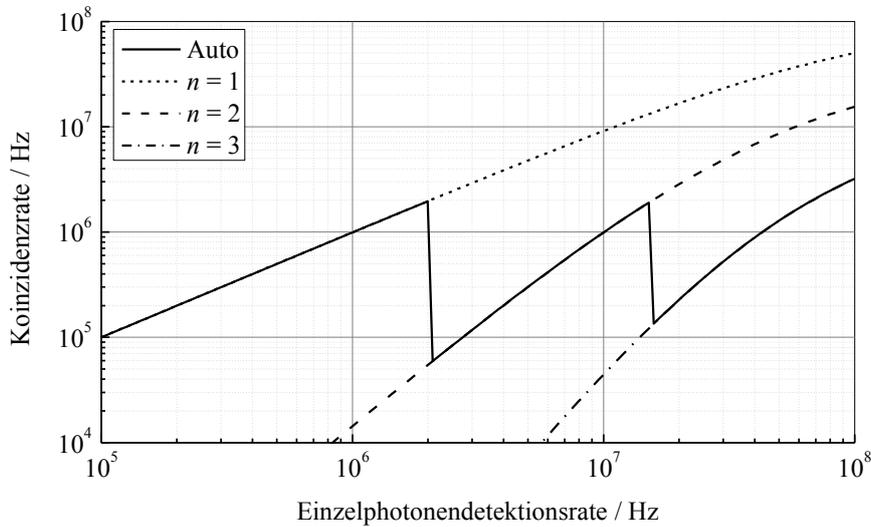


Abbildung 6.3: Anpassung der Ereignisraten durch Änderung der Koinzidenztiefe

man beispielsweise ein minimales SNR von 0,1 für eine erfolgreiche Messung voraus, erhöht sich der Dynamikbereich durch die beiden zusätzlichen Koinzidenztiefen von 26 dB auf 51 dB.

Eine andere Darstellungsart der adaptiven Koinzidenz ist in Abbildung 6.3 gezeigt. Hierbei zeigen die Kurven der einzelnen Koinzidenztiefen die resultierenden Raten von Ereignisdetektionen. Die vierte Kurve stellt die Rate dar, wie sie bei einer entsprechenden Umschaltung zwischen den Tiefen zu erwarten ist. Der Grenzwert für den Parameterwechsel wurde auf eine Hintergrundrate von 2 MHz festgelegt, da so stets ein maximales SNR gemäß Abbildung 6.2 erreicht wird. Zur weiteren Untersuchung der adaptiven Koinzidenz werden Simulationen mittels des in Abschnitt 5 beschriebenen Modells durchgeführt.

Zur Simulation von Koinzidenz werden zunächst die Photonankunftszeiten für jede SPAD eines Pixels unter Berücksichtigung der Totzeit generiert. Anschließend werden in diesen Zeiten Koinzidenzereignisse gesucht. Dafür muss die Zeit zwischen n aufeinanderfolgenden Detektionen in den SPADs kleiner als die gewählte Koinzidenzzeit sein. Da eine Schaltung zur Detektion von Koinzidenz nach [65] angenommen wird, muss die Totzeit der SPADs mindestens der Koinzidenzzeit entsprechen. D. h. die notwendigen Photonendetections zur Generation eines Koinzidenzereignisses müssen von verschiedenen SPADs

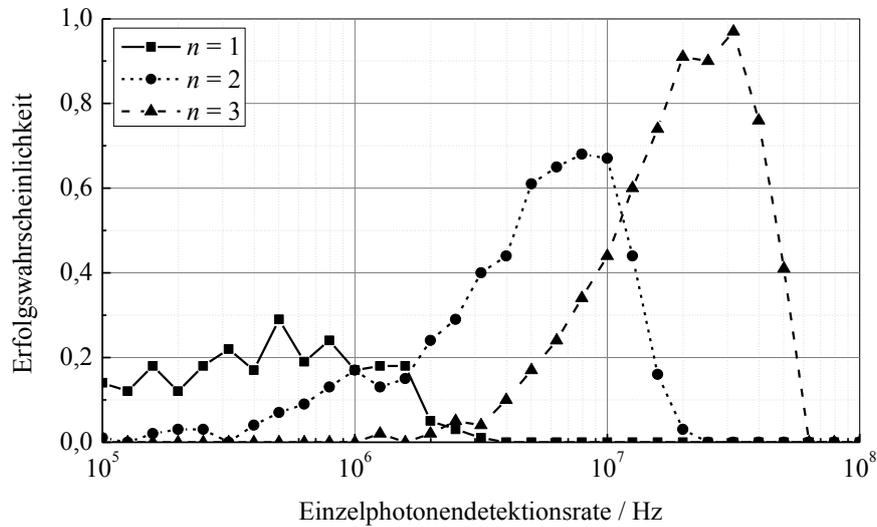


Abbildung 6.4: Simulation der Erfolgswahrscheinlichkeit der Distanzmessung

stammen. Abbildung 6.4 zeigt in Analogie zum SNR in Abbildung 6.2 die Erfolgswahrscheinlichkeit der Distanzmessung als Funktion der Hintergrundereignisrate für verschiedene Koinzidenztiefen. Die Erfolgswahrscheinlichkeit wird dabei aus einer vielfachen Wiederholung der Simulation einer Distanzmessung gewonnen, wobei eine Messung als erfolgreich gilt, wenn die gemessene Distanz eine maximale Abweichung zum tatsächlichen Wert nicht überschreitet. Die Simulation zeigt eine gute qualitative Übereinstimmung mit dem berechneten SNR. Ein Vergleich der beiden Graphen zeigt, dass bei dem verwendeten Auswertalgorithmus ein SNR von etwa 0,4 notwendig ist, um die festgelegte Messtoleranz in 90 % der Messungen einzuhalten. Das erforderliche SNR ist vom Algorithmus abhängig und wird daher an dieser Stelle nicht weiter betrachtet. Die Simulation bestätigt aber das Konzept einer adaptiven Anpassung der Koinzidenzparameter zur Vergrößerung des Dynamikbereichs.

Um den Übergang zwischen den einzelnen Koinzidenztiefen zu glätten und Einbrüche in der Messleistung zu verhindern, können neben der Koinzidenztiefe weitere Parameter variiert werden. Dazu zählen die Koinzidenzzeit sowie die Anzahl der SPADs, welche für die Detektion von Koinzidenzereignissen verknüpft werden. Wie sich eine Änderung der jeweiligen Größe auswirkt kann anhand des Modells in Abschnitt 6.1.1 ermittelt werden.

6.2 Mehrfache Ereignisdetektion

Dieses Konzept basiert auf der direkten Methode zur Laufzeitmessung. Bei der klassischen Methode wird die SPAD mit dem Aussenden des Laserpulses aktiviert und nur die Ankunftszeit des ersten detektierten Photons jedes Zyklus erfasst [13], [66]. Die Zeit zwischen Senden des Pulses und Erkennen des Photons wird in ein Histogramm eingetragen, woraus sich nach mehrfacher Wiederholung die Laufzeit bestimmen lässt. Mit steigender Hintergrundstrahlung erhöht sich jedoch die Wahrscheinlichkeit, dass das erste detektierte Photon nicht Bestandteil des am Zielobjekt reflektierten Pulses ist, sondern aus dem Hintergrundlicht resultiert. Die Wahrscheinlichkeit dafür ist gegeben zu

$$P_{\text{BG}}(T_{\text{TOF}}) = 1 - \exp(-r_{\text{B}}T_{\text{TOF}}) \quad (6.17)$$

mit der Photonenrate des Hintergrundes r_{B} und der Lichtlaufzeit T_{TOF} . Um die beschriebene Problematik der Falschdetektionen zu entschärfen, soll nicht nur die Ankunftszeit des ersten Photons, sondern möglichst aller detektierten Photonen erfasst werden. Für diesen Zweck wird die SPAD im Freilaufmodus betrieben, d. h. die SPAD wird zwischen den einzelnen Messzyklen nicht deaktiviert und nach der Detektion eines Ereignisses erfolgt unmittelbar das Zurücksetzen der Diode, sodass ein weiteres Ereignis erfasst werden kann. Damit können zum einen, sofern die Zyklusdauer größer als die Totzeit ist, mehrere Photonen je Zyklus erkannt werden und zum anderen variiert der Zeitpunkt, zu welchem die SPAD aktiv ist. Über viele Zyklen betrachtet erfolgt auf diese Weise eine homogene Erfassung des empfangenen Signals über der Zykluszeit. Im Weiteren erfolgt eine theoretische Analyse des Konzeptes der mehrfachen Ereignisdetektion zur Leistungsabschätzung.

Gemäß den Betrachtungen in Abschnitt 4.2.1 liefert die Summation der PDFs der detektierten Photonen eine zeitinvariante Photonendetektionswahrscheinlichkeit, welche der effektiven Photonenrate unter dem Einfluss von Totzeit entspricht. Die direkte Messung der Laufzeit mittels einer SPAD im Freilaufmodus und mehrfacher Ereignisdetektion scheint damit möglich.

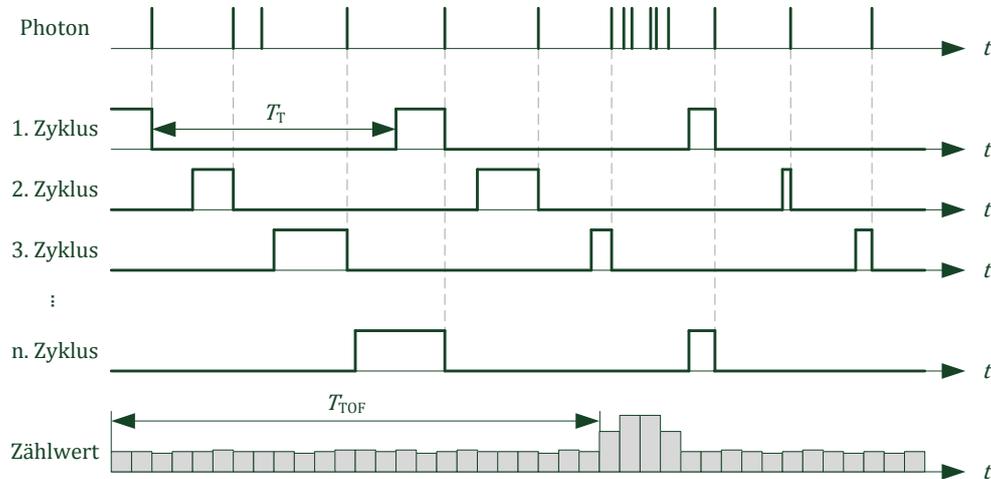


Abbildung 6.5: Prinzip der direkten Laufzeitmessung mit freilaufender SPAD und mehrfacher Ereignisdetektion

Die Erfassung der Ankunftszeit aller im Messfenster detektierten Photonen erfordert viele Speicherelemente, die in Falle eines Arrays den Füllfaktor des Sensors reduzieren sowie die Datenmenge erhöhen. Die Anzahl der benötigten Speicherbänke ist durch die Dauer des Messfensters, d. h. durch die Reichweite des Systems, und die Totzeit, welche die Rate möglicher Detektionen begrenzt, bestimmt. Eine alternative Methode stellt die Erfassung einer begrenzten Anzahl von Photonen innerhalb eines Messfensters dar.

6.2.1 Signal-Rausch-Verhältnis

Zur Abschätzung der Leistungsfähigkeit des Verfahrens, wird im Weiteren das SNR nach (4.74) betrachtet. Dieses beruht auf der Anzahl der Ereignisse der Klasse zum Zeitpunkt der Pulsankunftszeit und lässt sich aus der PDF bestimmen. Die PDF ist bei alleiniger Betrachtung des Hintergrundes und unter Vernachlässigung der Totzeit gegeben durch die Summe der Erlang-Verteilung nach (4.21) für $k = 1, 2 \dots N_{ph}$, wobei N_{ph} die Anzahl der möglichen Ereignisdetektionen im Messfenster darstellt, entsprechend

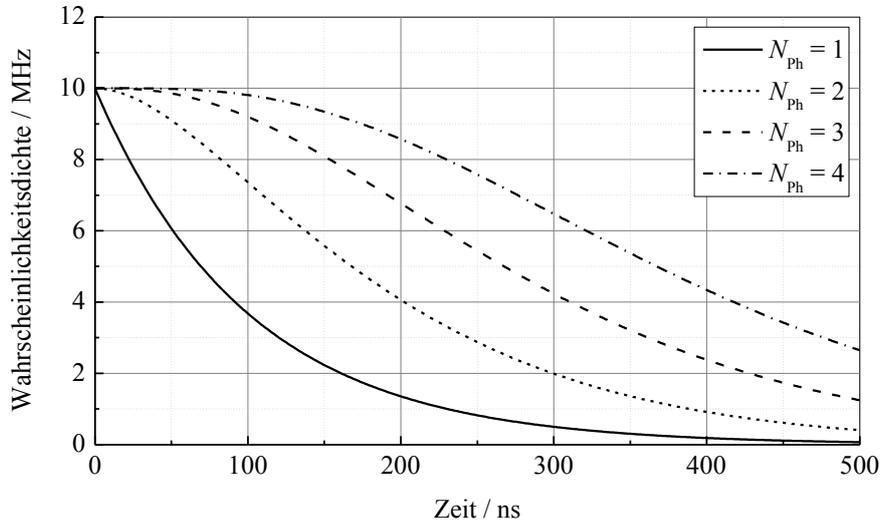


Abbildung 6.6: Summe der Erlang-Verteilung nach (6.18)

$$P_{Ph}(t) = \sum_{k=1}^{N_{Ph}} \frac{r_B^k t^{k-1}}{(k-1)!} \exp(-r_B t). \quad (6.18)$$

Die PDF nach (6.18) als Funktion der Zeit für eine Hintergrundereignisrate r_B von 10 MHz ist in Abbildung 6.6 für verschieden Anzahlen möglicher Detektionen dargestellt. Für eine einzelne Detektion ergibt sich die bekannte Exponentialverteilung. Mit steigender Anzahl erhöht sich der Bereich einer konstanten Dichte zu Beginn des Messfensters und gleichzeitig fällt die Kurve langsamer ab. Da der Wert der PDF zum Zeitpunkt der Pulsankunft eine direkte Auswirkung auf die Güte der Messung hat, lässt eine höhere Anzahl eine Verbesserung der Messung erwarten. Die PDF bei zusätzlicher Berücksichtigung der Photonennrate des reflektierten Laserpulses ist zur Bestimmung des SNR nicht erforderlich. Stattdessen kann das SNR allgemein ausgedrückt werden durch

$$a_{Hist} = \sqrt{\frac{N_Z T_{Bin} P_{Ph}(T_{TOF})}{r_B}} \frac{r_L}{\sqrt{r_L + r_B}}. \quad (6.19)$$

Der Faktor $P_{Ph}(T_{TOF})/r_B$ trägt der Reduktion der Detektionswahrscheinlichkeit eines Photons mit zunehmender Laufzeit Rechnung. Für $N_{Ph} = 1$ ist der Faktor

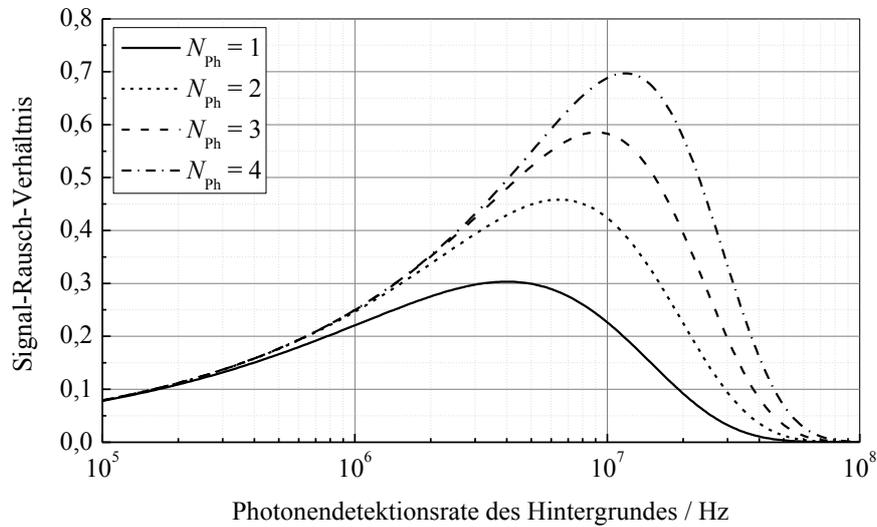


Abbildung 6.7: SNR des Histogramms nach (6.19)

entsprechend (4.78) gegeben durch $\exp(-r_B T_{\text{TOF}})$. Der Faktor $P_{\text{Ph}}(T_{\text{TOF}})$ hat keine einfache analytische Lösung, daher wird das SNR numerisch bestimmt und in einem Graphen veranschaulicht. Abbildung 6.7 zeigt das berechnete SNR für verschiedene N_{Ph} für eine Laufzeit von 250 ns. Hierbei zeigt sich für eine Ereignisrate von 10 MHz eine signifikante Zunahme des SNR, was den Erwartungen aus Abbildung 6.6 entspricht. Bei geringeren oder höheren Rate nähern sich die Kurven einander an, dies entspricht einer Verschiebung der Laufzeit in Abbildung 6.6 und deckt sich somit ebenfalls mit den Erwartungen durch die sich annähernden PDFs. Weiterhin verschiebt sich das Maximum des SNR zu höheren Raten mit steigender Anzahl der möglichen Detektionen. Dementsprechend ist das betrachtete Konzept besonders für höhere Hintergrundintensitäten interessant. Ein weiterer Vorteil ist eine verbesserte Detektion von mehreren Signalpulsen, welche durch mehrfache oder teilweise Reflexion des Lasersignals entstehen können. Bei einer starken Reflexion an einem Objekt, welches sich weiter vorne in der Szene befindet und den Detektor auslöst, ist somit die Detektion eines zweiten, weiter entfernten Objektes möglich.

Zur Verifikation der theoretischen Berechnungen wird eine Simulation der Distanzmessung für verschiedene N_{Ph} durchgeführt. Abbildung 6.8 zeigt die Erfolgswahrscheinlichkeit der Distanzmessung als Funktion der Ereignisrate

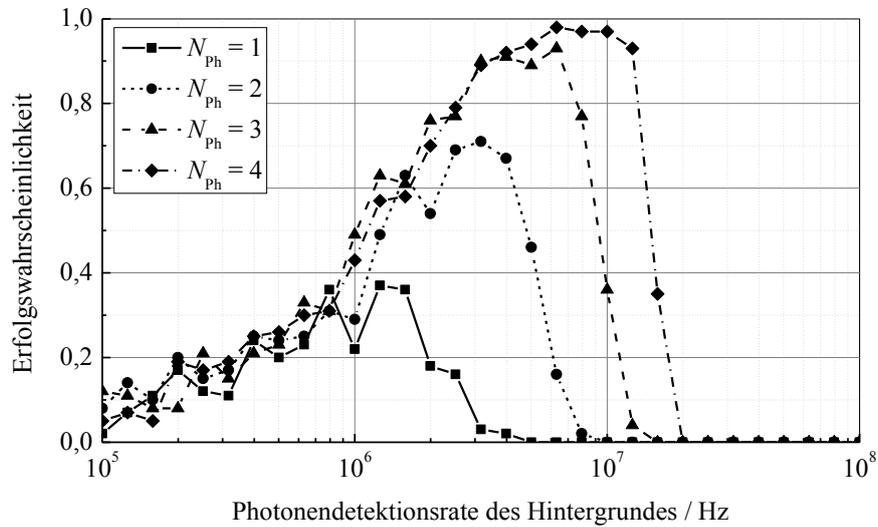


Abbildung 6.8: Simulation der Erfolgswahrscheinlichkeit für verschiedene N_{Ph}

des Hintergrundes. Die Erfolgswahrscheinlichkeit ist die Wahrscheinlichkeit, die Distanz mit einer maximalen Abweichung von 3 % zur wahren Distanz zu messen und wurde für die Darstellung aus 100 Distanzmessungen, wobei für jede die Ankunftszeit der ersten N_{Ph} Photonen aus 400 Messzyklen akkumuliert wurde, berechnet. Das Simulationsergebnis zeigt eine gute Übereinstimmung mit dem berechneten SNR in Abbildung 6.7 und bestätigt somit das Potential der Erfassung mehrerer Ereignisse innerhalb des Messfensters.

6.2.2 Bewertung

Das Verfahren beruht auf der direkten Laufzeitmessung, wobei nicht nur das erste sondern N_{Ph} Ereignisse erfasst und deren Ankunftszeiten im Histogramm akkumuliert werden. Durch das Verfahren ist es möglich, auch dann den reflektierten Laserpuls zu detektieren, wenn aufgrund von hohem Hintergrundlicht $N_{\text{Ph}} - 1$ Ereignisse vor dessen Ankunft am Detektor aufgetreten sind. Damit ist das Verfahren besonders bei hohem Hintergrund vorteilhaft und verbessert die Güte der Distanzmessung. Zur Erfassung von N_{Ph} Ereignissen ist es allerdings erforderlich, N_{Ph} Speicherbänke, wobei jede die Breite des Zeitstempels des TDCs hat, zu integrieren. Für zweidimensionale Arrays ergibt sich daraus eine Reduktion des Füllfaktors. Für Zeilensensoren, bei welchen die Pixelelektronik

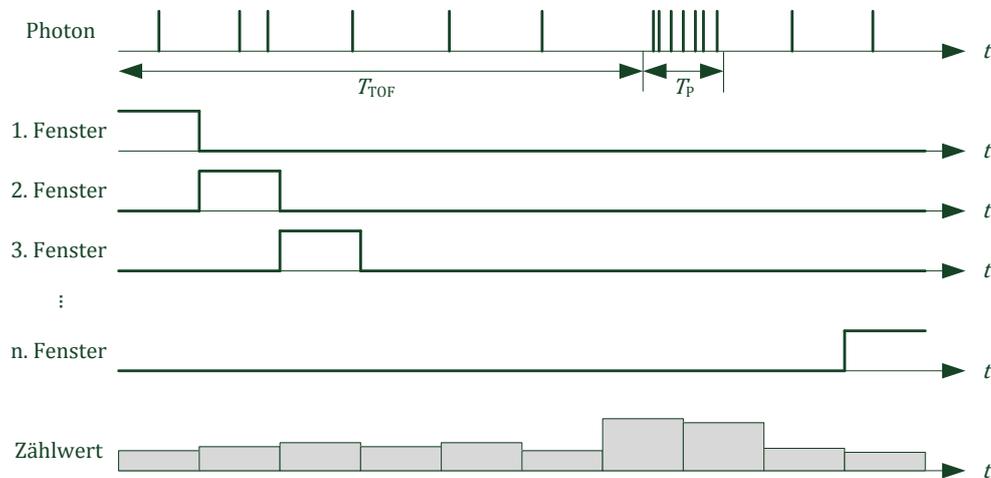


Abbildung 6.9: Prinzip der indirekten Laufzeitmessung mit mehreren Fenstern

außerhalb des photosensitiven Bereichs platziert wird, stellt das Verfahren hingegen eine Optimierungsmöglichkeit dar. Unabhängig von der Bauweise des Sensors sorgt die Erfassung mehrere Ereignisse zu einem erhöhten Datenaufkommen, welches bei der Auslegung der Ausleseschaltung berücksichtigt werden muss.

6.3 Erhöhte Fensteranzahl

Eines der Hauptprobleme der indirekten Laufzeitmessung ist die hohe Dauer der einzelnen Pulse, welche durch die gewünschte Reichweite des Messsystems bestimmt wird, und die damit einhergehende geringe zulässige Laserleistung. Eine kürzere Pulsdauer, wie sie im direkten Verfahren zum Einsatz kommt, erlaubt dagegen eine höhere Laserleistung, was zu einer besseren Unterdrückung der Hintergrundstrahlung führt. Um die Pulsdauer im indirekten Verfahren mit PM zu reduzieren, kann die Anzahl der Fenster erhöht werden. Die notwendige Anzahl der Fenster ergibt sich aus der maximal zu messenden Laufzeit, welche sich aus der Reichweite des Systems ergibt, und der Fensterdauer, welche der Pulsdauer entsprechen muss. Entsprechend Abbildung 6.9 wird die gesamte Messdauer durch die Fenster sukzessive abgedeckt. Am Ende eines Messzyklus ergibt sich ein Histogramm, wobei jede Klasse die Anzahl der innerhalb eines

Fensters gezählten Photonen enthält. Analog zum klassischen Verfahren mit drei Fenstern liegt der Puls in bis zu zwei Fenstern. Zur Bestimmung der Laufzeit werden die zwei aufeinanderfolgenden Fenster, in welchen der reflektierte Puls empfangen wurde, ermittelt. Aus diesen wird die Laufzeit nach

$$T_{\text{TOF}} = \left(\frac{N_{i+1} - \bar{N}}{N_i + N_{i+1} - 2\bar{N}} + i - 1 \right) T_P \quad (6.20)$$

mit den Fenstern N_i und N_{i+1} in welchen der Puls liegt, der Lichtlaufzeit T_{TOF} und der Pulsdauer T_P berechnet. Der zusätzliche Term $(i - 1)T_P$ trägt der Verschiebung um $i - 1$ Fenster Rechnung. Da hier nicht nur ein Fenster die Hintergrundstrahlung erfasst, wird der Mittelwert aller nicht im Rückkehrzeitraum des Pulses befindlichen Fenster zur Kompensation der Hintergrundstrahlung entsprechend

$$\bar{N} = \frac{1}{N_W - 2} \sum_{\substack{k=1 \\ k \neq i, i+1}}^{N_W} N_k \quad (6.21)$$

mit der Anzahl der Fenster N_W verwendet.

6.3.1 Varianz

Die Varianz dieses Verfahren berechnet sich zunächst analog zum klassischen indirekten Verfahren mit PM und drei Fenstern. Dabei werden die Zählwerte in den Fenstern N_i gemäß der Poisson-Verteilung angenommen. Durch die Nutzung des Mittelwertes zur Subtraktion des Hintergrundes reduziert sich die Varianz des Hintergrundwertes um den Faktor $N_W - 2$, womit sich aus (6.20) nach Anwendung der Fehlerfortpflanzung eine Varianz von

$$\sigma_\delta^2 = \frac{c^2}{4T_P(r_{\text{LB,T}} - r_{\text{B,T}})^2} \left(\frac{(T_P^2 - 3T_P\delta + 3\delta^2)r_{\text{B,T}}}{(1 + r_{\text{B}}T_{\text{T}})^2} + \frac{(T_P^2 - 4\delta T_P + 4\delta^2)r_{\text{B,T}}}{(N_W - 2)(1 + r_{\text{B}}T_{\text{T}})^2} + \frac{(T_P\delta - \delta^2)r_{\text{LB,T}}}{(1 + r_{\text{LB}}T_{\text{T}})^2} \right) \quad (6.22)$$

mit $\delta = T_{\text{TOF}}$ modulo T_P ergibt. Verglichen mit der Varianz des klassischen Verfahrens nach (4.32) entspricht dies einer Verbesserung um

$$\Delta\sigma_\delta^2 = -\frac{c^2(4\delta^2 - 4T_P\delta + T_P^2)r_{B,T}(N_W - 3)}{4T_P(r_{LB,T} - r_{B,T})^2(1 + r_{B,T}T_T)^2(N_W - 2)}. \quad (6.23)$$

Da die Varianz zudem proportional zum Kehrwert der Pulsdauer T_P ist, ergibt sich durch Senkung der Pulsdauer eine weitere Reduktion der Varianz. Im Gegensatz zur klassischen Variante der indirekten Laufzeitmessung ergibt sich durch die Bestimmung der Fenster, in welchen der Puls empfangen wurde, eine zusätzliche Fehlerquelle. Die Wahrscheinlichkeit, dass diese korrekt ermittelt werden, hängt vom verwendeten Algorithmus für die Ermittlung der Distanz aus dem Histogramm der Zählwerte ab. Beispielfhaft wird diese hier für einen einfachen Algorithmus, welcher den Maximalwert im Histogramm als Ankunfts-fenster des reflektierten Pulses interpretiert, bestimmt. Die Wahrscheinlichkeit, dass von $N_W - 2$ Fenstern mit dem Erwartungswert $\mu_B = N_Z r_B T_W$ mindestens eines einen Wert größer ζ aufweist ist gegeben zu

$$P_3(k > \zeta) = 1 - P_B\left(0, N_W - 2, \sum_{i=\zeta+1}^{\infty} P_P(i, \mu_B)\right) \quad (6.24)$$

mit der Binomialverteilung P_B und der Poisson-Verteilung P_P . Hierbei gibt die Binomialverteilung die Wahrscheinlichkeit an, dass von $N_W - 2$ Versuchen mit einer Erfolgswahrscheinlichkeit von

$$p = \sum_{i=\zeta+1}^{\infty} P_P(i, \mu_B) \quad (6.25)$$

keiner zum Erfolg führt. Die Erfolgswahrscheinlichkeit ist gegeben durch die Summe aller Wahrscheinlichkeiten für einen Zählwert größer ζ bei einem Erwartungswert von μ_B . Im Umkehrschluss liefert $1 - P_B$ die Wahrscheinlichkeit dafür, dass mindestens ein Versuch einen Wert über ζ aufweist. Weiterhin ist die Wahrscheinlichkeit, dass ein Fenster in dem zumindest ein Teil des Pulses

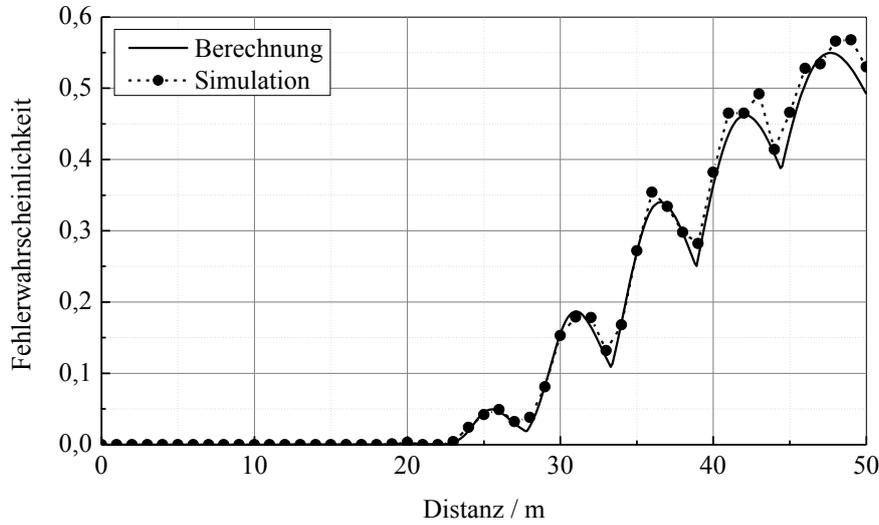


Abbildung 6.10: Fehlerwahrscheinlichkeit als Funktion der Distanz

empfangen wurde, d. h. mit Erwartungswert $\mu_{A,1} = N_Z(r_B T_W + r_L(T_P - \delta))$ bzw. $\mu_{A,2} = N_Z(r_B T_W + r_L \delta)$, einen Wert kleiner ζ aufweist gegeben zu

$$P_i(k \leq \zeta) = \sum_{i=0}^{\zeta} P_P(i, \mu_{A,i}). \quad (6.26)$$

Damit der reflektierte Puls im Histogramm nicht erkannt wird, müssen beide Fenster, in welchen der Puls empfangen wurde, einen Wert kleiner ζ aufweisen. Die Wahrscheinlichkeit hierfür ergibt sich durch Multiplikation beider Einzelwahrscheinlichkeiten nach (6.26). Die Fehlerwahrscheinlichkeit berechnet sich durch diskrete Differenzierung des Produkts multipliziert mit der Wahrscheinlichkeit nach (6.24) und Summenbildung über alle ζ zu

$$P_{\text{Fail}} = \sum_{\zeta=0}^{\infty} dP_{12}(k \leq \zeta) P_3(k > \zeta) \quad (6.27)$$

mit $dP_{12}(k \leq \zeta) = P_1(k \leq \zeta + 1)P_2(k \leq \zeta + 1) - P_1(k \leq \zeta)P_2(k \leq \zeta)$. Mit Hilfe der Fehlerwahrscheinlichkeit kann der Erwartungswert sowie die Varianz der Distanzmessung abgeschätzt werden. Unter der Annahme, dass der gemessene Distanzwert bei einer Fehlinterpretation über dem gesamten Empfangsbereich

gleichverteilt ist, ergibt sich der Erwartungswert aus der Addition der wahren Distanz gewichtet mit der Wahrscheinlichkeit den Puls korrekt zu lokalisieren und dem Erwartungswert der Gleichverteilung bei einer fehlerhaften Lokalisierung zu

$$\mu_d = d(1 - P_{\text{Fail}}) + \frac{d_{\text{max}}}{2} P_{\text{Fail}}. \quad (6.28)$$

Der Erwartungswert der Gleichverteilung ist durch die Hälfte des Empfangsintervalls gegeben. Dementsprechend ergibt sich die Varianz durch die Summe der gewichteten Varianz nach (6.22) sowie der Varianz der Gleichverteilung zu

$$\sigma_d^2 = \sigma_8^2(1 - P_{\text{Fail}}) + \frac{d_{\text{max}}^2}{12} P_{\text{Fail}} + (d - \mu_d)^2. \quad (6.29)$$

Der Summand $(d - \mu_d)^2$ ergibt sich aus der Verschiebung des Mittelwertes gegenüber der wahren Distanz und addiert sich zur Gesamtvarianz. Der mittels (6.28) bzw. (6.29) berechnete und simulierte Erwartungswert sowie die Standardabweichung des Verfahrens sind in Abbildung 6.11 dargestellt. Bei einer angenommenen Reichweite von 50 m und der Verwendung von 10 Messfenstern ergibt sich eine Fensterbreite von 5,56 m bzw. 37 ns, da der durch die Fenster abzudeckende Bereich den Messbereich um die Pulslänge überschreiten muss, um auch bei maximaler Distanz den Puls vollständig zu erfassen. Die angenommenen Ereignisraten betragen 1 MHz für den Hintergrund sowie $r_L(d) = 1 \text{ GHz m}^2/d^2$ für den Laserpuls. Die Totzeit wurde für diese Betrachtung vernachlässigt. Die Simulation zeigt eine gute Übereinstimmung mit der Berechnung. Lediglich für höhere Distanzen zeigt sich eine zunehmende Abweichung beider Kurven. Dennoch erlaubt das theoretische Modell eine Abschätzung der Leistungsfähigkeit des betrachteten Verfahrens.

Unter Berücksichtigung der Totzeit zeigt die Simulation des Verfahrens eine Verschiebung der ermittelten Laufzeit hin zu niedrigeren Werten für hohe Ereignisraten während des Pulses. Dieses Verhalten begründet sich in der Tatsache, dass durch die hohe Rate die SPAD im ersten Fenster auslöst und somit im zweiten Fenster durch die Totzeit teilweise inaktiv ist. Dieser Effekt wurde bereits in Abschnitt 4.2.2.1 untersucht. Im betrachteten Konzept verstärkt sich

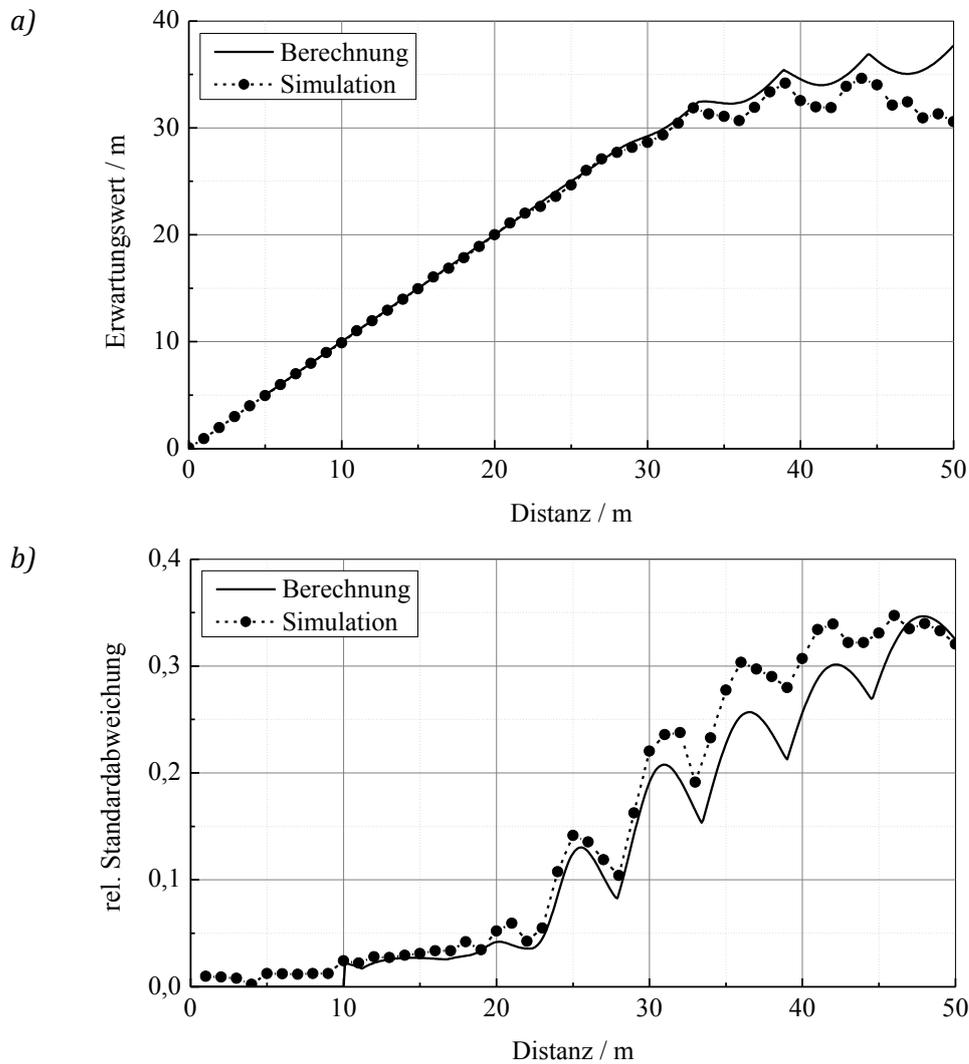


Abbildung 6.11: Berechneter und simulierter Mittelwert (a) und Standardabweichung (b) bei Verwendung mehrerer Messfenster

dieser Effekt durch die Reduktion der Fensterbreite zusätzlich, wodurch eine software- oder schaltungstechnische Korrektur erforderlich wird.

6.3.2 Bewertung

Die Erhöhung der Fensterzahl im indirekten Verfahren mit PM reduziert die Varianz und erlaubt gleichzeitig eine erhöhte Laserleistung. Abbildung 6.12 zeigt die berechnete Standardabweichung für das klassische Verfahren mit drei

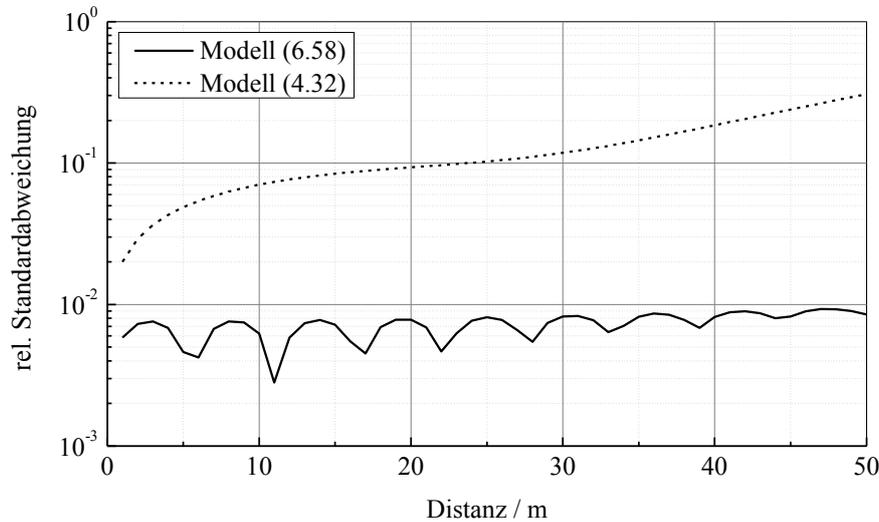


Abbildung 6.12: Standardabweichung des klassischen und vorgestellten Verfahrens

sowie das vorgestellte Verfahren mit 10 Fenstern. Hierbei zeigt sich für den gesamten Distanzbereich eine deutliche Verbesserung der Abweichung. In der Berechnung wurde eine 9-fach höhere Ereignisrate des Laserpulses für das neue Verfahren angenommen, da die Pulsdauer um diesen Faktor geringer ausfällt und sich so eine identische mittlere Laserleistung ergibt. Unter dieser Annahme wird die Fehlerwahrscheinlichkeit in Abbildung 6.10 um den Faktor drei gestreckt, wodurch diese über dem gesamten Messbereich einen vernachlässigbaren Wert aufweist und sich die Varianz nach (6.22) ergibt. Durch die zusätzlichen Zeitfenster ergibt sich allerdings ein erhöhter Hardwareaufwand, so wird für eine maximale Effizienz des Verfahrens für jedes Fenster ein eigener Zähler benötigt, was den Füllfaktor reduziert. Gleichzeitig erhöht sich das Datenaufkommen sowie die Komplexität des Algorithmus zur Datenauswertung, da nun zusätzlich die Position der Fenster, in welchen der Puls empfangen wurde, gefunden werden muss. Für Zeilensensoren, bei welchen der zusätzliche Platzbedarf weniger in Gewicht fällt, stellt das vorgestellte Verfahren eine gute Möglichkeit zur Steigerung der Reichweite dar. Für Flächensensoren muss dagegen ein Kompromiss zwischen Varianz und Platzbedarf gefunden werden.

6.4 Verfahrensauswahl

Nach der Betrachtung der verschiedenen Verfahren und Messkonzepte soll im Folgenden das vielversprechendste für die gegebene Zielanwendung ausgewählt werden. Ziel ist der Entwurf eines Zeilensensors für die Anwendung im Automobil. Damit spielt der Platzbedarf der Schaltung zunächst eine untergeordnete Rolle. Die größte Herausforderung stellt die Erzielung einer hohen Reichweite unter hoher Umgebungslichtintensität dar. Um eine Messung zu ermöglichen, muss die Intensität des reflektierten Laserpulses sich von der Hintergrundstrahlung abheben, was eine hohe optische Leistung des emittierten Laserpulses erfordert. Unter Berücksichtigung von Kriterien der Augensicherheit, setzt eine hohe Laserleistung möglichst kurze Pulse voraus. Dies schließt das indirekte Messverfahren aus, da selbst bei einer Verwendung mehrerer Messfenster entsprechend Abschnitt 6.3 der Puls deutlich länger als im direkten Messverfahren sein muss. Das in Abschnitt 6.1 beschriebene Verfahren der Photonenkorrelation sowie die Detektion mehrerer Ereignisse in Abschnitt 6.2 basieren auf dem direkten Verfahren und sind somit potentielle Verfahren für den Sensorentwurf. Photonenkorrelation mit variablen Parametern erlaubt die Messung bei hohem Hintergrundlicht über eine hohe Dynamik der Reflektanz des Zielobjektes und soll daher in den Sensorentwurf einfließen. Die Erfassung mehrerer Ereignisse erhöht die Schaltungskomplexität sowie die Datenmenge beträchtlich, weshalb dieses Konzept nicht im Sensor verwendet werden soll. Dennoch ist dieses Konzept vielversprechend und kann als mögliche Optimierung des Sensors für zukünftige Versionen angesehen werden.

7 SPAD-basierter Zeilensensor für automobile Anwendungen

In diesem Abschnitt erfolgt unter der Berücksichtigung gegebener Rahmenbedingungen und sinnvoll gewählter Annahmen ein Entwurf für ein Sensorsystem samt Laserquelle. Da das System auf die Verwendung im Automobil ausgelegt werden soll, werden zunächst die dort anzutreffenden Umgebungsbedingungen betrachtet und auf Grundlage dessen relevante Größen des zu entwerfenden Sensors festgelegt. Der Sensorentwurf zielt auf das direkte Messverfahren ab, da dieses im Zusammenspiel mit Photonenkorrelation und einem optimalen Algorithmus zur Auswertung der Rohdaten die beste Leistung verspricht. In diesem Kapitel werden das Design des Sensors allgemein sowie die Funktionsweise von Schlüsselschaltungsteilen detailliert beschrieben. Weiterhin werden die Ergebnisse der Sensorcharakterisierung präsentiert.

7.1 Assistenzsysteme im Automobil

Fahrerassistenzsysteme dienen der Erhöhung der Sicherheit und des Komforts im Straßenverkehr [67], [68]. Dabei handelt es sich um elektronische Zusatzeinrichtungen, welche durch Eingriffe in die Fahrzeugsteuerung den Fahrer unterstützen. Die Stärke des Eingriffs lässt sich in unterschiedliche Typen unterteilen: In unkritischen Situationen kann es ausreichen, den Fahrer durch eine

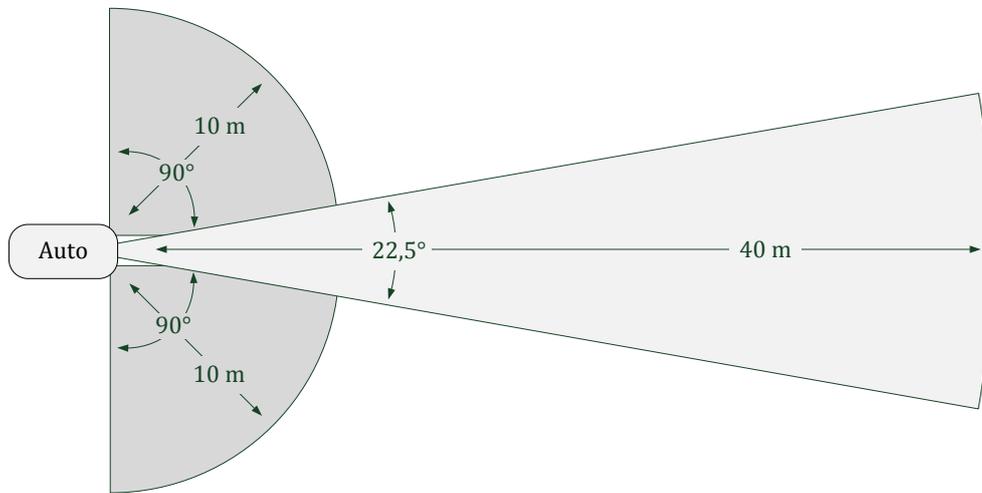


Abbildung 7.1: Sensoranordnung zur Überwachung des Nah- und Fernbereichs

geeignete Mensch-Maschinen-Schnittstelle auf die Gefahrensituation hinzuweisen. Ein Beispiel hierfür ist die elektronische Einparkhilfe, welche den Fahrer akustisch oder visuell auf Hindernisse aufmerksam macht. In kritischen Situationen kann das System in die Steuerung des Fahrzeuges aktiv eingreifen. Beispielsweise den Lenkwinkel des Fahrers verstärken um einem Hindernis auszuweichen oder eine Notbremsung einleiten um eine drohende Kollision zu vermeiden. Diese Systeme benötigen geeignete Sensoren sowie Hard- und Software zur Auswertung der Sensorsignale und Steuerung des Eingriffs. In den genannten Systemen werden Sensoren zur Überwachung der Fahrzeugumgebung benötigt, welche die Position eines Objektes und den Abstand zum Fahrzeug bestimmen können. Da Laufzeitsensoren genau diese Anforderungen erfüllen, stellen sie einen geeigneten Kandidaten für solche Systeme dar. Abbildung 7.1 zeigt eine beispielhafte Anordnung dreier Sensoren in einem Kraftfahrzeug zur Überwachung des Nah- und Fernbereichs. Im gezeigten Beispiel dienen die beiden äußeren Sensoren mit einer Reichweite von 10 m und einem horizontalen Blickfeld von 90° zur Beobachtung des Bereichs seitlich vor dem Fahrzeug. Die Überwachung dieses Bereichs ist im Stadtverkehr bei geringer Geschwindigkeit sinnvoll und dient zur Vermeidung von Kollisionen mit Fußgängern, sich seitlich nähernden Fahrzeugen oder festen Gegenständen wie Gebäuden oder parkenden Fahrzeugen. Der zentrale Sensor mit einer Reichweite von

40 m und einem Blickfeld von $22,5^\circ$ dient der Überwachung der vorausliegenden Fahrbahn und kann zur Auslösung einer Notbremsung oder Abstandsregelung bei mittleren Geschwindigkeiten im Stadtbereich eingesetzt werden.

Mit der immer weiter steigenden Anzahl von Assistenzsystemen und Automatisierung der Fahrzeuge wurde ein weiterer Trend geboren: das autonome Fahren. Hierbei wird die Steuerung des Fahrzeuges vollständig automatisiert, sodass die Notwendigkeit eines Fahrers entfällt. Für autonome Fahrzeuge spielt die Überwachung der Umgebung mit hoher Tiefen- und Winkelauflösung eine entscheidende Rolle. Nur wenn dies unter allen Umgebungsbedingungen fehlerfrei gelingt, kann autonomes Fahren Realität werden. Auch für diese Anwendungen eignet sich die lichtbasierte Distanzmessung aufgrund der hohen Winkelauflösung besonders. Viele Prototypen autonomer Fahrzeuge nutzen LiDAR um die Umgebung zu vermessen [69]. Dabei basieren diese Systeme zumeist auf scannenden Systemen, was sie sehr teuer und für den Massenmarkt untauglich macht. Die Entwicklung Flash-basierter Laufzeitsensoren ist daher ein wichtiger Schritt auf dem Weg zur Marktreife autonomer Fahrzeuge [20].

7.2 Sensorkonzeptionierung

Im ersten Schritt erfolgt die grundlegende Konzeptionierung des zu entwerfenden Sensors bezüglich des Messkonzepts, der Sensorarchitektur, der Auflösung und der Pixelarchitektur.

7.2.1 Messkonzept

Gemäß der Betrachtung in Abschnitt 6.4 soll der Sensor auf dem direkten Messverfahren beruhen, da dieses Verfahren die kürzesten Laserpulse und damit die höchste Emissionsleistung erlaubt. Gleichzeitig bietet es weitere Vorteile: Da die Ankunftszeit jedes Photons individuell erfasst wird, ist dieses Verfahren in der Lage Mehrfachreflexionen des Laserpulses zu erkennen, welche im indirekten Verfahren zu einer Verschiebung der gemessenen Distanz führen würden [70]. Weiterhin ist es möglich, Objekte hinter teiltransparenten Medien wie Glas oder Wasser zu erkennen, was Vorteile bei Nebel oder Regen erwarten lässt.

7.2.2 Dynamikbereich

Ein wichtiger Punkt bei der Auslegung des Sensorsystems ist der erforderliche Dynamikbereich der Lichtintensität, den die Pixel erfassen müssen. Da der Sensor im Automobil zum Einsatz kommen soll, ist besonderes Augenmerk auf das Hintergrundlicht zu legen, da dieses aufgrund der hohen Intensität des Sonnenlichtes die Intensität des Lasers überstrahlen kann. Da die Ereignisrate des reflektierten Laserlichtes quadratisch mit der Distanz abnimmt und nur hohes Hintergrundlicht der Messqualität abträglich ist, genügt es zur Auslegung des Sensorsystems den Fall maximaler Distanz und Hintergrundintensität zu betrachten. Dennoch ändern sich die Raten auch in diesem Fall durch eine Änderung der Reflexionscharakteristik des Zielobjektes. Der Sensor soll einen Reflexionsgrad von 5 % bis 100 % abdecken können, was einer Änderung der Raten um den Faktor 20 bzw. 26 dB entspricht.

Um eine möglichst hohe Reichweite des Systems zu erzielen, muss das SNR nach (4.78) einen möglichst hohen Wert über dem gesamten Dynamikbereich aufweisen. Entsprechend den Betrachtungen in Abschnitt 6.1.3 erreicht eine einzelne SPAD ein hohes SNR jedoch nur über einem kleinen Bereich, woraus sich eine geringere Reichweite an den Bereichsgrenzen ableitet. Um das SNR über dem gesamten Bereich konstant zu halten, wird Photonenkoinzidenz verwendet. Hierfür werden vier SPADs für jedes Pixel eingesetzt, womit der geforderte Bereich abgedeckt und der schaltungstechnische Aufwand zur Realisierung der Photonenkorrelationsdetektion überschaubar gehalten werden kann. Als weitere Stellgrößen zur Justierung der Koinzidenz werden zudem die Koinzidenzzeit in vier Stufen und die Anzahl der aktiven SPADs variabel gestaltet. Hierbei muss sichergestellt werden, dass die Anzahl aktiver SPADs stets mindestens der Koinzidenztiefe entspricht.

7.2.3 Sichtfeld und Auflösung

Typische Hindernisse, welche im Automobilbereich zuverlässig erkannt werden müssen, sind Fußgänger und andere Fahrzeuge. Da diese über eine große vertikale Ausdehnung verfügen, wird für die Erprobung des Konzepts der Hintergrundlichtunterdrückung ein Zeilensensor entworfen. Ein Vorteil hierbei ist die

Möglichkeit, die In-Pixel-Elektronik außerhalb der Zeile zu platzieren, was einen hohen Füllfaktor und eine nahezu lückenlose Überwachung des horizontalen Sichtfeldes ermöglicht. Zur Verbesserung der Sicherheit und vertikalen Auflösung ist ein Multizeilensensor denkbar, welcher es erlaubt, die Elektronik zwischen den einzelnen Zeilen zu platzieren, um die genannten Vorteile beizubehalten. Der Sensor muss in der Lage sein, die Zielobjekte in maximaler Distanz sicher zu erkennen. Um dies zu gewährleisten wird die Auflösung auf die minimale Objektgröße in horizontaler Richtung ausgelegt. Um eine horizontale Auflösung von d_{\min} in maximaler Distanz d_{\max} zu erreichen, ist eine Zeile mit

$$N_{\text{Pixel}} = \left\lceil \frac{d_{\max} \alpha_H}{d_{\min}} \right\rceil \quad (7.1)$$

Pixeln notwendig. Wird für das System in Abbildung 7.1 mit $\alpha_H = 0,39$ rad und $d_{\max} = 40$ m eine Auflösung von 10 cm gefordert, werden 158 Pixel benötigt. Je nach Formfaktor des Pixels definiert dies auch den vertikalen Öffnungswinkel. Da in dieser Arbeit die Breite des Sensors auf ca. 10 mm aufgrund des verwendeten Gehäuses begrenzt ist und SPADs mit rundem Aktivgebiet des Durchmessers $12 \mu\text{m}$ mit einem Pitch von $40,56 \mu\text{m}$ in horizontaler bzw. $52,4 \mu\text{m}$ in vertikaler Richtung verwendet werden, müssen die einzelnen Dioden übereinander angeordnet werden, um die geforderte horizontale Mindestauflösung zu erreichen. Die Anzahl der Pixel wird auf 192 bei zwei Zeilen festgelegt. Jede Zeile bietet bei einer Brennweite des Objektivs von 12 mm damit ein Sichtfeld (Field-of-View, FOV) von $35,95^\circ \times 1^\circ$ bei einem Füllfaktor von 5,32 %. Der Winkel zwischen den Zeilen ergibt sich aus deren Abstand von $810 \mu\text{m}$ zu $3,87^\circ$.

7.3 Schaltungsentwurf

Dieses Kapitel beschreibt den Entwurf des 2×192 Pixel Zeilensensors. Wie in Abbildung 7.2 gezeigt, lässt sich der Sensor in einen globalen und lokalen Schaltungsteil in den Pixeln unterteilen. Der Systemtakt von 200 MHz wird global mittels einer Phase-Locked-Loop (PLL) aus dem Eingangstakt von 25 MHz erzeugt. Die weiteren Schaltungsteile sowie die zugehörigen Unterblöcke werden in den folgenden Unterkapiteln näher betrachtet.

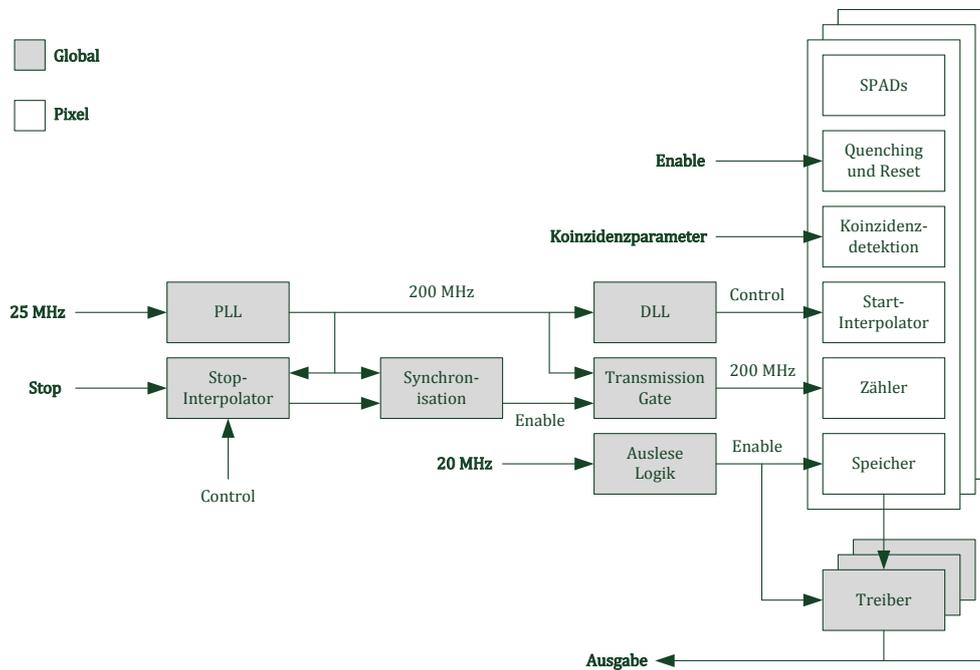


Abbildung 7.2: Blockschaltbild des 2 x 192 Pixel Doppelzeilensensors

7.3.1 Pixelschaltung

Die Schaltung des Pixels lässt sich in zwei Blöcke untergliedern: Der erste Block enthält die SPADs sowie alle Schaltungsteile, die für deren Betrieb und Auswertung der Detektionssignale zuständig sind. Das Ausgangssignal dieses Blocks dient dem zweiten Block, welcher den lokalen Teil des TDCs umfasst, als Eingangssignal zum Starten der Zeitmessung bzw. zur Ereigniszählung. Der zweite Teil wird im Kapitel 7.3.2 im Kontext des TDCs näher erläutert.

Abbildung 7.3 zeigt das Blockschaltbild des ersten Teils der Pixelschaltung aufgeteilt in die Schaltungskomponenten zur Ereignisdetektion sowie zu deren Steuerung. Jedes Pixel des Sensors enthält vier SPADs mit jeweils eigener Schaltung zur Unterbrechung des Stromflusses durch die Diode nach der Detektion eines Photons und anschließendem Wiederaufladen. Dessen folgt jeweils eine Gating-Schaltung, welche dazu dient, die Detektionssignale nur während eines definierten Zeitfensters an die nachfolgende Schaltung weiterzugeben. Ein Pulsformer mit variabler Pulsdauer sowie die anschließende logische Verknüpfung

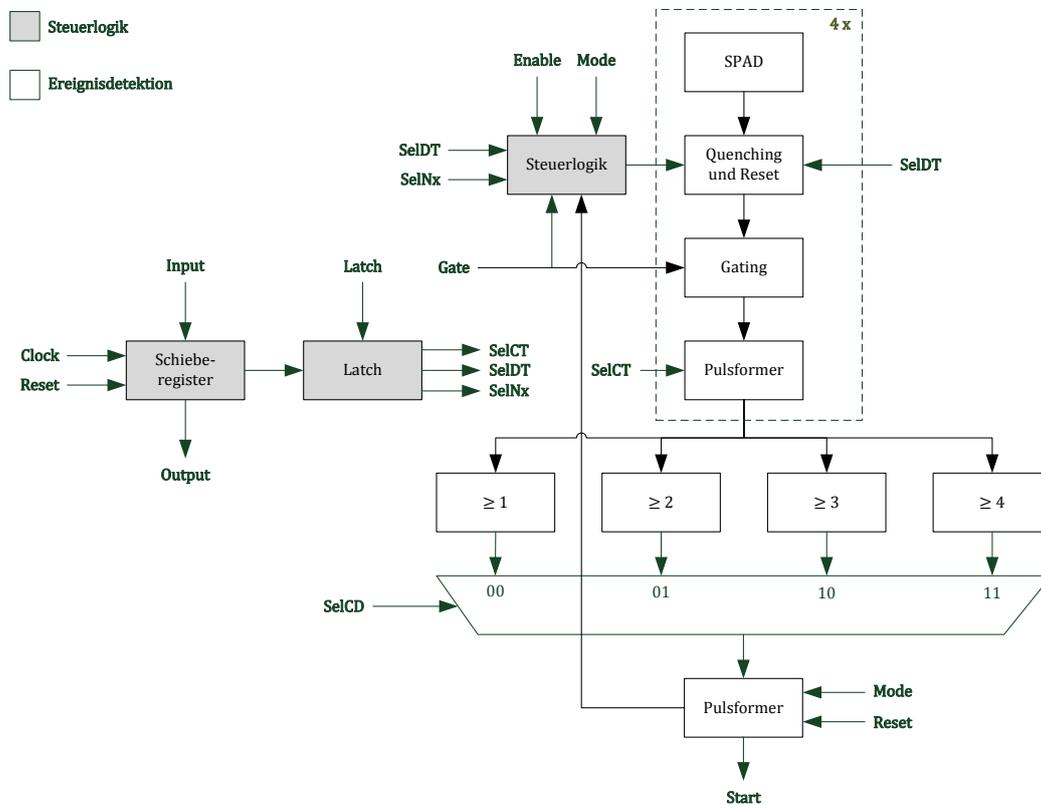


Abbildung 7.3: Blockschaltbild der Pixelschaltung zur Ansteuerung der SPADs und Photonenkorrelationsdetektion

der Signale dienen der Detektion von Photonenkorrelationen mit einstellbarer Koinzidenzzeit und Koinzidenztiefe. Ein Flipflop mit automatischem zeitverzögertem Zurücksetzen erzeugt Pulse mit definierter Breite und Pause für eine fehlerfreie Weiterverarbeitung der Signale.

7.3.1.1 Quenching und Reset

Nach der Detektion eines einfallenden Photons muss der Stromfluss durch die Diode unterbrochen werden. Da dies durch das Absenken der Spannung unter die Durchbruchspannung geschieht, muss in einem zweiten Schritt die Spannung wieder über den Schwellwert angehoben werden, um ein weiteres Photon detektieren zu können. Für den Sensor kommt die Schaltung nach [47] zur Anwendung. Diese verwendet das Prinzip des aktiven Quenching und Reset der Diode. Der Vorteil dieser Methode ist eine definierte Totzeit, welche nicht durch

ein während des Aufladevorgangs einfallendes Photon beeinflusst wird. Zudem ist es möglich, die Diode für eine gewisse Zeit inaktiv zu halten um die Wahrscheinlichkeit für Afterpulsing zu reduzieren. Um die SPADs deaktivieren zu können, wird der Rückkopplungspfad mittels eines logischen Gatters gesteuert. Ist dieser deaktiviert, wird der Stromfluss nach der Detektion eines Photons zwar unterbrochen, jedoch erfolgt kein automatisches Wiederaufladen. Auf diese Weise kann die Diode ohne Leistungsaufnahme inaktiv gehalten werden.

7.3.1.2 Korrelationsdetektion

Um zeitliche Korrelation zwischen den detektierten Photonen eines Pixels zu erfassen, wird ein entsprechender Schaltungsblock verwendet. In Anlehnung an [65] werden hierfür die Ausgangssignale der einzelnen SPADs über logische Gatter miteinander verknüpft. Um die Koinzidenztiefe variabel zu halten, werden mehrere logische Verknüpfungen für die verschiedenen Koinzidenztiefen realisiert. Da für den Sensor die Tiefen 1 bis 4 möglich sein sollen, kommen entsprechend vier Subblöcke zu Einsatz. Um zwischen den Tiefen zu wechseln wird ein Multiplexer verwendet, welcher über ein 2-Bit-Signal gesteuert wird. Um Fehler durch Laufzeitunterschiede zu vermeiden, sind die Logikschaltungen so ausgelegt, dass jedes Signal dieselben Gatter durchlaufen muss. Für die Koinzidenztiefe 1 wird ein einzelnes OR-Gatter verwendet, während für die Tiefe 4 ein AND-Gatter zu Einsatz kommt. Für die Tiefen 2 und 3 werden alle möglichen Kombinationen mit separaten AND-Gattern abgedeckt und anschließend mit einem OR-Gatter verknüpft. Für eine Tiefe von 2 ergeben sich sechs und für eine Tiefe von 3 vier Kombinationen. Die booleschen Gleichungen sind

$$\begin{aligned}
 C_0 &= S_3 + S_2 + S_1 + S_0 \\
 C_1 &= S_3S_2 + S_3S_1 + S_3S_0 + S_2S_1 + S_2S_0 + S_1S_0 \\
 C_2 &= S_3S_2S_1 + S_3S_1S_0 + S_2S_1S_0 + S_3S_2S_0 \\
 C_3 &= S_3S_2S_1S_0
 \end{aligned} \tag{7.2}$$

mit den von den SPADs kommenden Eingangssignalen S_i und den Ausgangssignalen der vier Logikblöcke C_i .

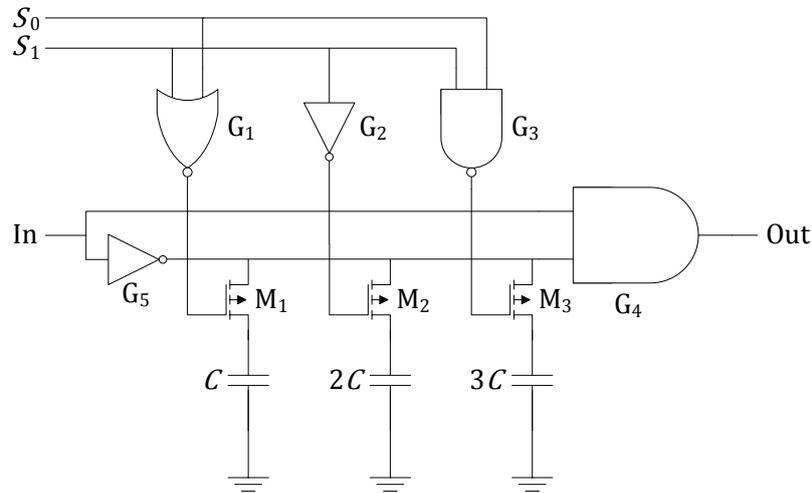


Abbildung 7.4: Pulsformer mit variabler Pulsdauer zur Koinzidenzzeitvariation

Zur Einstellung der Koinzidenzzeit werden Pulsformer mit variabler Pulsdauer benutzt. Die Quenching-Schaltung liefert Pulse mit der Dauer der Totzeit, welche mittels der Pulsformer auf die gewünschte Koinzidenzzeit reduziert wird. Die maximale Koinzidenzzeit ist daher auf die Totzeit begrenzt, welche jedoch über ein analoges Steuersignal einstellbar ist. Analog zur Koinzidenztiefe soll auch die Zeit in 4 Stufen von 1,5 ns, 4 ns, 8 ns und 16 ns variiert werden können. Die Einstellung erfolgt durch eine Änderung der kapazitiven Last am Eingangsknoten des pulsgenerierenden Gatters wie Abbildung 7.4 zeigt.

Da durch die logische Verknüpfung der Signale sehr kurze Pulse entstehen können, welche die Setup- und Hold-Zeit der nachfolgenden Signalverarbeitung verletzen, folgt ein weiterer Pulsformer. Dieser besteht aus einem RS-Flipflop, welches durch die steigende Flanke des Ausgangssignals der logischen Verknüpfung gesetzt und zeitverzögert zurückgesetzt wird. Die Verzögerung beträgt 2 ns, was folglich der Pulsdauer sowie der Pulspause entspricht. Das automatische Zurücksetzen kann zudem unterbrochen werden, da es in bestimmten Situationen nicht erwünscht ist. So kann bei der Bestimmung der Laufzeit nur die Ankunftszeit des ersten Ereignisses erfasst werden, womit eine Detektion weiterer Ereignisse überflüssig ist und das Zurücksetzen entfallen kann. Zusätzlich werden in diesem Fall die SPADs nach der Ereignisdetektion deaktiviert um die Leistungsaufnahme des Sensors zu reduzieren.

Tabelle 7.1: Wahrheitstabelle zur Aktivierung der SPADs

Anzahl		Tiefe		Aktive SPADs			
A_1	A_0	T_1	T_0	S_3	S_2	S_1	S_0
0	0	0	0	0	0	0	1
0	0	0	1	0	0	1	1
0	0	1	0	0	1	1	1
0	0	1	1	1	1	1	1
0	1	0	0	0	0	1	1
0	1	0	1	0	0	1	1
0	1	1	0	0	1	1	1
0	1	1	1	1	1	1	1
1	0	0	0	0	1	1	1
1	0	0	1	0	1	1	1
1	0	1	0	0	1	1	1
1	0	1	1	1	1	1	1
1	1	0	0	1	1	1	1
1	1	0	1	1	1	1	1
1	1	1	0	1	1	1	1
1	1	1	1	1	1	1	1

7.3.1.3 SPAD Steuerlogik

Eine weitere Stellgröße zur Anpassung der resultierenden Ereignisrate im Pixel ist die Anzahl der aktiven SPADs. Zur Vereinfachung der Steuerlogik ist nur die Anzahl aktiver SPADs wählbar statt einer individuellen Steuerung der einzelnen Dioden. Diese Einschränkung halbiert die Anzahl der notwendigen Steuersignale. Um Bedienfehler, welche eine Messung verhindern würden, zu vermeiden, stellt die Logik zudem sicher, dass mindestens so viele SPADs entsprechend der gewählten Koinzidenztiefe aktiv sind. Tabelle 7.1 zeigt die Wahrheitstabelle zur Aktivierung der SPADs. Die Anzahl sowie die Koinzidenztiefe werden von jeweils 2-Bit-Signalen vorgegeben. Das Ausgangssignal ist ein 4-Bit-Wert, wobei jedes Bit für die Aktivierung einer SPAD zuständig ist. Aus der Wahrheitstabelle lassen sich die folgenden Gleichungen in boolescher Algebra ermitteln:

$$\begin{aligned}
 S_0 &= 1 \\
 S_1 &= A_1 + A_0 + T_1 + T_0 \\
 S_2 &= A_1 + T_1 \\
 S_3 &= \overline{A_1 A_0} \overline{T_1 T_0}
 \end{aligned}
 \tag{7.3}$$

Um die Leistungsaufnahme zu reduzieren, sind die SPADs nur während des definierten Messfensters aktiv und werden deaktiviert, sobald ein Ereignis erkannt und der TDC gestartet wurde. Letzteres gilt nur im Zeitmodus, während im Zählmodus die SPADs während des gesamten Messfensters aktiv bleiben.

7.3.1.4 Schieberegister

Zur Einstellung der Koinzidenzzeit, Koinzidenztiefe und Anzahl aktiver SPADs kommen Schieberegister zum Einsatz. Für die beiden Pixelzeilen des Sensors werden separate Register verwendet, sodass insgesamt sechs Register mit einer Breite von jeweils 384 Bit verbaut werden. Die Register selbst bestehen aus getakteten D-Flipflops, welche über ein externes Taktsignal gesteuert werden. Um die seriell eingegebenen Daten synchron an alle Pixel zu übergeben, werden Latches verwendet. Dadurch ist es möglich, neue Daten bereits während einer laufenden Messung in die Register zu schreiben und diese nach dem Ende der Messung den Pixeln zu übergeben.

7.3.2 Time-to-Digital-Converter

Der TDC dient der Umwandlung der Laufzeit des Laserpulses in einen digitalen Wert und stellt damit eine entscheidende Baugruppe des Sensors dar. Für den TDC wird die Architektur nach [71] verwendet. Diese Architektur bietet im Vergleich mit anderen Architekturen [65], [72], [73], [74], [75], [76] eine für die Zielanwendung ausreichende zeitliche Auflösung bei geringem Verdrahtungsaufwand und Platzbedarf. Die Begrenzung in der Anzahl der möglichen Signale der Pixel ergibt sich durch den geringen Pitch von 40,56 μm . Ein geringer Platzbedarf kommt dem Füllfaktor sowie dem Abstand der Pixelzeilen zu Gute.

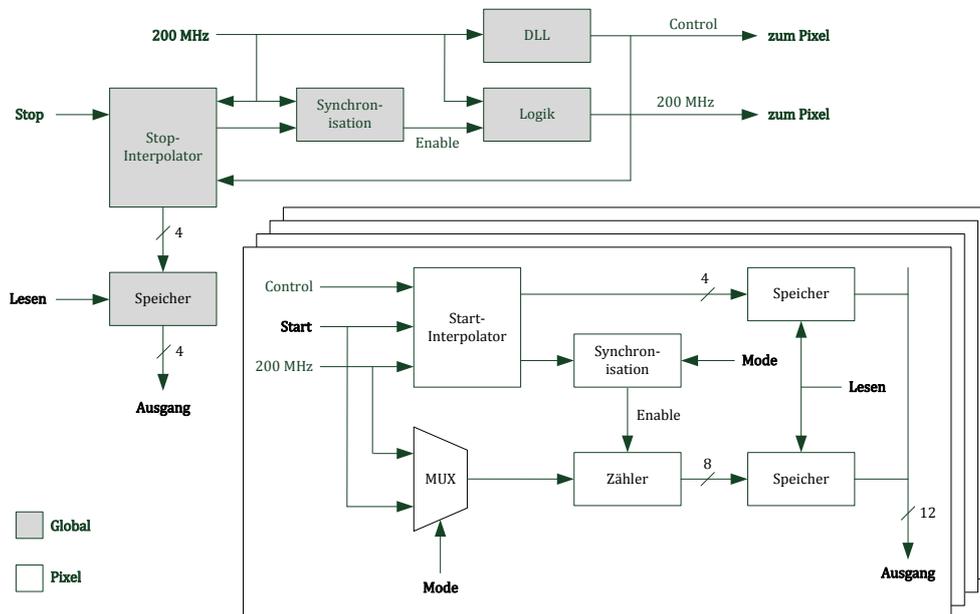


Abbildung 7.5: Blockschaltbild des TDCs mit globalem und lokalem Schaltungsteil

Der TDC besteht aus zwei Stufen für grobe und feine Zeiterfassung. In der ersten Stufe dient ein Zähler der Erfassung der verstrichenen Taktperioden des Haupttaktes durch Zählung der positiven Signalfanken. Die grobe Zeitauflösung ist daher durch die Taktfrequenz bestimmt, während die Breite des Zählers die maximal messbare Dauer festlegt. Aufgrund der verwendeten 0,35- μm -CMOS-Technologie wird der in [71] verwendete Takt von 560 MHz auf 200 MHz reduziert. Bei einem 8-Bit-Zähler ergibt sich damit eine grobe Auflösung von 5 ns bei einem Dynamikumfang von 1,28 μs , was Distanzen von 74,95 cm bzw. 191,87 m entspricht. Zur Verbesserung der Auflösung kommt eine zweite Stufe zur Interpolation der Zeit zwischen zwei Taktflanken zu Einsatz. In der verwendeten Architektur wird hierfür das Taktsignal mittels einer Delay-Locked-Loop (DLL) in 16 Schritten verzögert, womit sich eine Auflösung des Least Significant Bits (LSB) von 312,5 ps oder 4,68 cm ergibt.

Das Blockschaltbild des TDCs ist in Abbildung 7.5 dargestellt. Dabei befindet sich der globale Schaltungsteil nur einmal auf den Chip, während der lokale Teil in jedem Pixel repliziert wird. Die Steuerspannung wird in der globalen DLL generiert und an die lokalen Verzögerungselemente der Start-Interpolatoren sowie des globalen Stopp-Interpolators verteilt. Die Synchronisationsschaltung

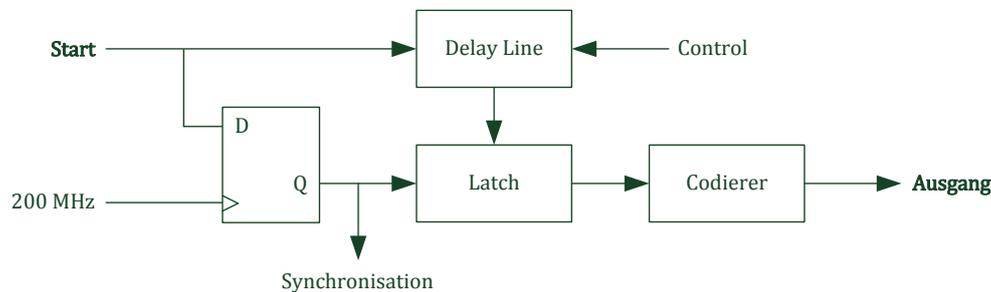


Abbildung 7.6: Blockschaltbild des Start- bzw. Stopp-Interpolators

dient der Vermeidung von fehlerhaften Signalzuständen durch Verletzungen der Setup- oder Hold-Zeit, welche auftreten können, wenn das Start- bzw. Stopp-Signal nahe an der steigenden Taktflanke auftritt. Nach dem Ende der Messung werden die Werte der Interpolatoren sowie des Zählers in Speicherelemente übertragen und anschließend ausgelesen.

7.3.2.1 Interpolation

Abbildung 7.6 zeigt das Blockschaltbild des Interpolators. Für Start- und Stopp-Interpolator wird die gleiche Architektur verwendet. Mit dem Auftreten des Start-Signals beginnt dieses sich durch die Kette aus Verzögerungselementen fortzupflanzen. Mit der nächsten positiven Taktflanke des 200-MHz-Taktes wird das D-Flipflop gesetzt, wodurch die Latches geschlossen werden und den aktuellen Zustand der Verzögerungselemente einfrieren. Die Elemente werden von der Steuerspannung der globalen DLL gesteuert und weisen damit ebenfalls eine Verzögerung von 312,5 ps auf. Der eingefrorene Zustand der Verzögerungselemente liefert einen 15-Bit-Thermometercode, welcher anschließend in einen 4-Bit-Graycode gewandelt wird. Der Interpolator liefert damit die Zeitspanne zwischen dem Start-Signal und der nächsten positiven Taktflanke.

7.3.2.2 Timing

Das Timing-Diagramm des TDCs ist in Abbildung 7.7 dargestellt. Der Start-Interpolator misst die Zeit zwischen dem Start-Signal und der nächsten positiven Taktflanke. Mit der darauffolgenden negativen Taktflanke wird der Zähler aktiviert und zählt die positiven Taktflanken. Die Aktivierung des Zählers mit

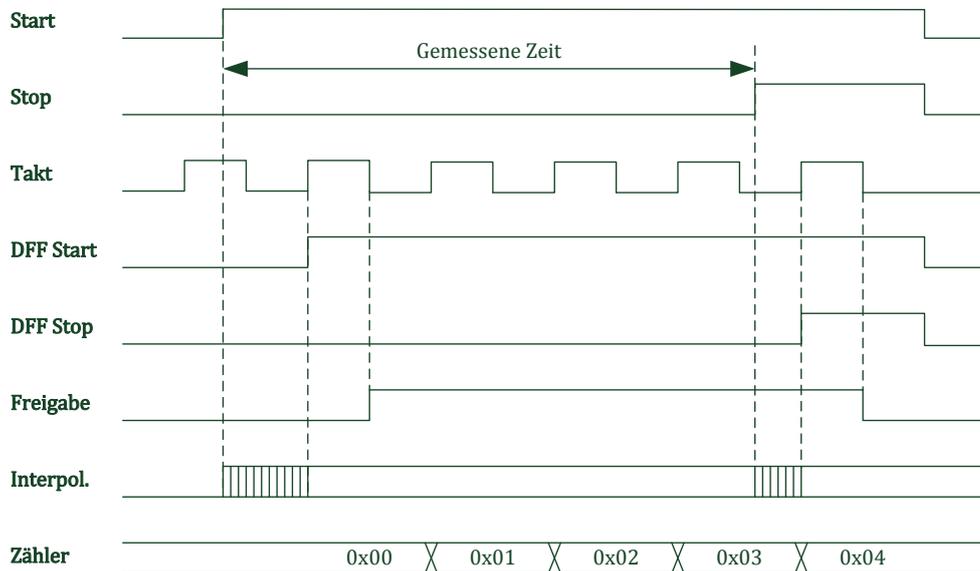


Abbildung 7.7: Timing-Diagramm des TDCs

der negativen Flanke erfolgt durch die Synchronisationsschaltung und dient der Vermeidung von Zeitverletzungen. Die analoge Vorgehensweise erfolgt am Ende der Messung: Der Stopp-Interpolator misst die Zeitspanne zwischen Auftreten des Stopp-Signals und der nächsten positiven Taktflanke. Das Anhalten des Zählers erfolgt durch die Unterbrechung des Taktsignals mit der nächsten negativen Flanke. Aus den Werten des Start- und Stopp-Interpolators N_{start} und N_{stop} sowie des Zählers N_{coarse} ergibt sich die gemessene Zeitspanne zu

$$T_{\text{meas}} = \left(\frac{N_{\text{start}} - N_{\text{stop}}}{16} + N_{\text{coarse}} \right) T_{\text{clk}} \quad (7.4)$$

mit der Periodendauer des Taktsignals T_{clk} . Um eine erneute Messung im nächsten Zyklus zu ermöglichen, wird das Taktsignal vor jeder Messung aktiviert.

7.3.3 Datenauslese

Jedes Pixel des Doppelzeilensensors hat einen 12-Bit-Datenausgang, welcher über einen vertikalen Datenbus ausgelesen wird. Dieser Bus wird spaltenweise zwischen den Zeilen geteilt, sodass jeder über Tristate-Treiber mit zwei Pixeln verbunden ist. Die vertikalen Leitungen sind wiederum über Tristate-Treiber an

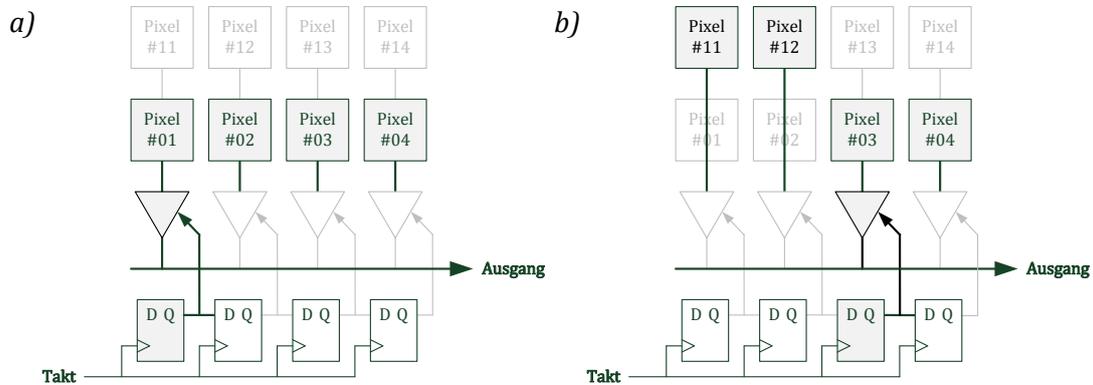


Abbildung 7.8: Ausleseschema des Doppelzeilensensors

den horizontalen Datenbus angeschlossen, welcher zur Datenauslese mit den Ausgangspads des Chips verbunden ist.

Zu Beginn der Auslese ist entsprechend Abbildung 7.8 (a) die untere Zeile mit den vertikalen 12-Bit-Datenleitungen verbunden. Diese werden mittels eines Schieberegisters nacheinander mit dem horizontalen Datenbus verbunden. Ist die untere Zeile zur Hälfte ausgelesen, wird, wie Abbildung 7.8 (b) zeigt, diese Hälfte von den vertikalen Datenleitungen getrennt und die erste Hälfte der zweiten Zeile verbunden. Analoges gilt für die zweite Hälfte der Zeilen. Durch dieses Schema haben die Treiber der vertikalen Datenleitungen ausreichend Zeit um diese ggf. umzuladen. Dies erlaubt Treiber kleinerer Dimension in Pixel, was den Platzbedarf reduziert, sowie eine schnellere Auslese der Daten [14]. Erfolgt die Auslese mit 20 MHz, beträgt die verfügbare Zeit zum Auf- bzw. Entladen der Datenleitungen $4,8 \mu\text{s}$. Die Aktivierung der Pixel erfolgt ebenfalls durch ein Schieberegister, wobei der Takt aus den Ausgängen der Flipflops 1 und 97 des horizontalen Schieberegisters erzeugt wird. Bei dieser Taktfrequenz benötigt die Auslese aller 384 Pixel $19,2 \mu\text{s}$.

7.3.4 Layout

Abbildung 7.9 zeigt das Layout und ein Foto des Sensors mit einer Größe von $8950 \times 5200 \mu\text{m}^2$. Die untere der beiden Zeilen ist so platziert, dass die SPADs exakt in der Mitte des Chips liegen. Die obere Zeile ist eine Kopie dieser und befindet sich darüber. Unterhalb der Zeilen befinden sich die Tristate-Treiber

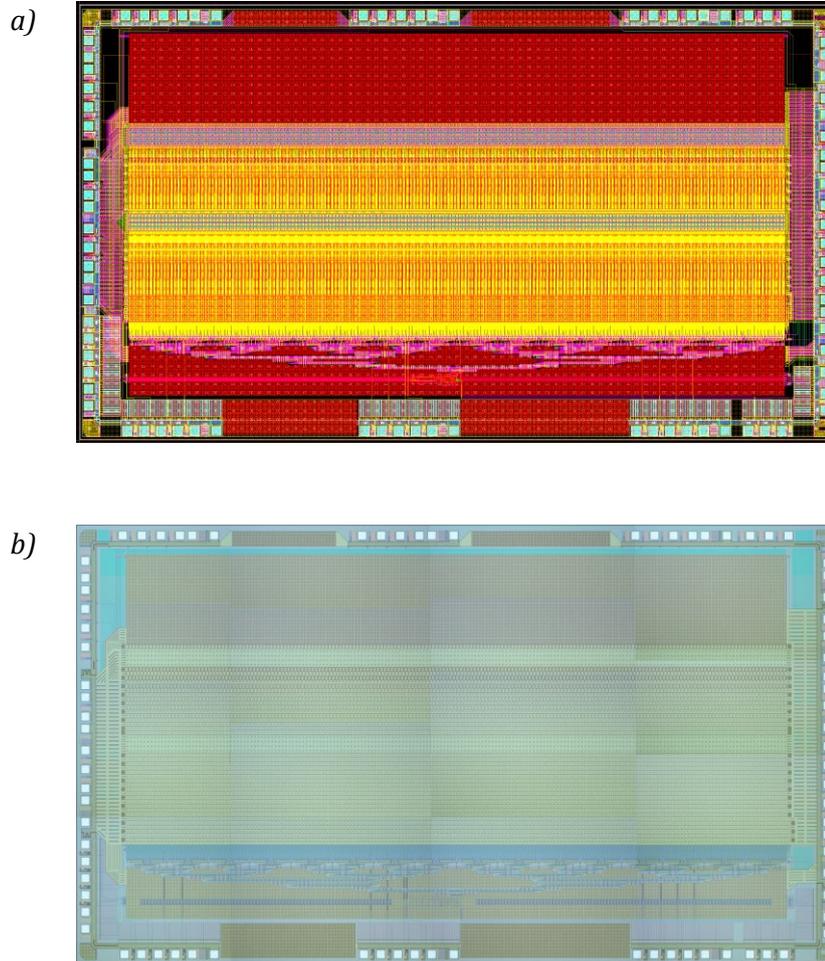


Abbildung 7.9: Layout (a) und Fotografie (b) des Doppelzeilensensors

für den horizontal verlaufenden Datenbus sowie das horizontale Schieberegister für die Datenauslese. Zentral darunter befindet sich der globale Schaltungsteil des Sensors mit PLL, DLL und Stop-Interpolator. Von diesem ausgehend werden alle zeitkritischen Steuersignale über eine Baumstruktur an die Pixel der unteren Zeile verteilt und über vertikale Leitungen an die obere Zeile weitergegeben. Links neben den Zeilen befindet sich das vertikale Schieberegister für die Auslese der Daten. Oberhalb der Zeilen befinden sich Füllstrukturen.

Tabelle 7.2: Kenngrößen des entworfenen Sensors

Parameter	Wert	Einheit
Verfahren	Direkt	–
Methode	Flash	–
Technologie	350	nm
Pixelarray	2 x 192	–
Pitch	40,56	μm
Zeilenabstand	1019,78	μm
SPAD Größe	12	μm
SPAD Form	Kreis	–
SPADs je Pixel	4	–
Füllfaktor	5,32	%
Zeitauflösung	312,5	ps
Messbereich	1,28	μs

7.3.5 Kenngrößen

Analog zum Stand der Technik in Tabelle 1.1 sind zur Übersicht die Kenngrößen des entworfenen Sensors in Tabelle 7.2 zusammenfassend dargestellt. Die zeitliche Auflösung des TDCs von 312,5 ps entspricht einer Distanz von 4,68 cm, während der Messbereich von 1,28 μs einer nominalen Distanz von 191,86 m entspricht. Mit 5,32 % liegt der Füllfaktor nur geringfügig über den Flächensensoren derselben Technologie aus Tabelle 1.1. Die Hauptursache hierfür liegt darin, dass zur Unterdrückung von Crosstalk jede SPAD in einer separaten n-Wanne liegt und aufgrund der hohen Kathodenspannung von 31 V entsprechend hohe Abstände eingehalten werden müssen. Gegenwärtig wird bereits an SPADs mit gemeinsamen Wannern gearbeitet, sodass für zukünftige Sensoren hohe Füllfaktoren erreicht werden können. Welchen Einfluss der geringe Füllfaktor auf die Schärfentiefe und das Aliasing hat, wird in Anhang B betrachtet.

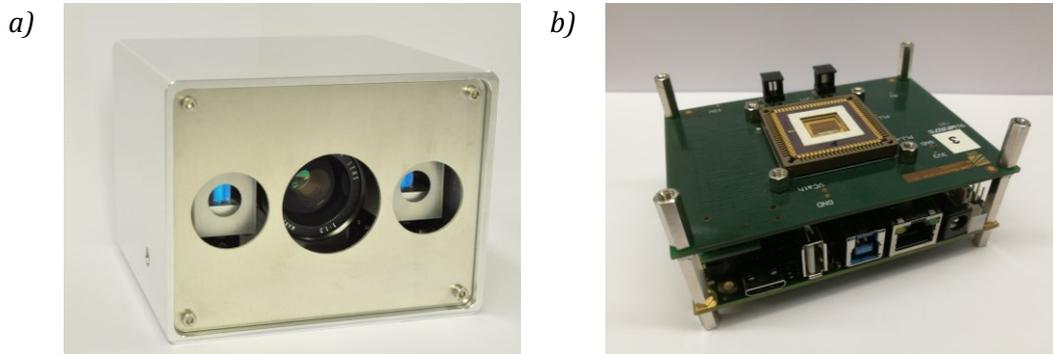


Abbildung 7.10: Kamerasystem „Owl“ (a) und Platinenstack der Kamera (b)

7.4 Kamerasystem

Für Charakterisierungs- und Demonstrationszwecke wird der Sensor in das am Fraunhofer IMS entwickelte Kamerasystem „Owl“ eingebaut. Zur Ansteuerung des Sensors und Auslese der Daten enthält dieses System ein Evaluation-Board des System-on-Chip (SoC) ZYNQ-7020 von Xilinx, welches sowohl einen ARM-basierten Prozessor als auch programmierbarer Logik enthält [77]. Auf einer zweiten Platine befindet sich der Sensorsockel, ein Digital-Analog-Umsetzer zur Bereitstellung analoger Spannungen sowie Komponenten zur Spannungsversorgung. Für die Belichtung der Szene kommen zwei Laserdioden des Typs SPL PL90-3 von Osram mit einer Spitzenleistung von 75 W zum Einsatz [78], welche, wie Abbildung 7.10 (a) zeigt, rechts und links des Sensors platziert sind. Die Ansteuerung der Dioden erfolgt mit der Lasertreiberschaltung LDP-AV 40-70 von PicoLAS GmbH. Diese erzeugt Pulse mit einer Dauer von 5 ns bei einem maximalen Ausgangsstrom von 40 A und einer Wiederholungsrate von bis zu 100 kHz [79]. Für die Sensoroptik kommt ein Standardobjektiv des Herstellers Computar mit einer Brennweite von 12 mm zum Einsatz, womit sich ein FOV jeder Zeile von $35,95^\circ \times 1^\circ$ ergibt. Zur Formung des Laserstrahls wird eine Kollimationslinse in Verbindung mit einem Diffusor verwendet. Für die Bedienung des Kamerasystems und Visualisierung der Messdaten wurde ein grafisches Benutzerinterface (GUI) mit Labview programmiert.

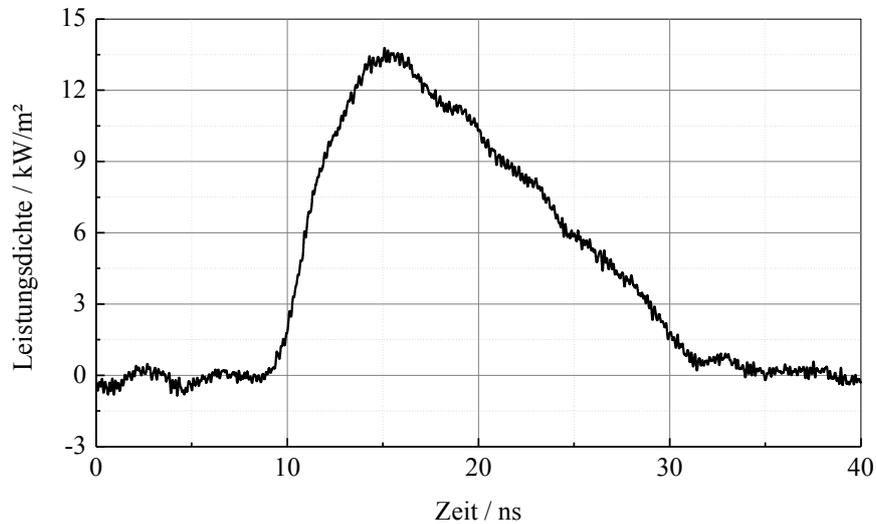


Abbildung 7.11: Gemessene Leistungsdichte des Laserpulses in 1 m Entfernung

7.4.1 Laserquelle

Als Laserquelle kommen zwei Laserdioden des Typs SPL PL90-3 von Osram zum Einsatz. Diese emittieren bei einer typischen Wellenlänge von 905 nm mit einer Bandbreite (FWHM) von 7 nm und einer Spitzenleistung von 75 W [78]. Zur Reduktion der Hintergrundlichtintensität auf dem Sensor wird ein optischer Bandpassfilter mit einer Mittenwellenlänge von 905 nm und einer 3-dB-Bandbreite von 90 nm verwendet. Zur Bestimmung der Pulsform wird eine Photodiode in 1 m Entfernung zur Laserquelle platziert. Abbildung 7.11 zeigt die gemessene Leistungsdichte einer Laserdiode mit einer Halbwertsbreite von 13 ns. Eine Umrechnung der Spitzenleistung von $13,5 \text{ kW/m}^2$ auf die emittierte optische Leistung unter Annahme einer Punktquelle mit Öffnungswinkeln von 40° und 1° mittels (5.1) ergibt eine Leistung von 164,6 W, was deutlich über dem Wert nach Datenblatt von 75 W liegt. Diese Leistung bleibt bis zu einem Winkel von 20° näherungsweise konstant. Bei einem Winkel von 25° lässt sich noch eine Spitzenleistung von 112,6 W und bei 30° von 42,1 W messen. Des Weiteren zeigt die Messung eine Halbwertsbreite von 13 ns und liegt damit deutlich über den 5 ns nach Datenblatt [79]. Die Form des Pulses, insbesondere die langsam abfallende Flanke, ist dem Aufbau des Lasertreibers geschuldet, bei

welchem der Stromfluss mittels einer Induktivität kontrolliert wird. Beim Einschalten ist der Widerstand der Diode groß, sodass die Zeitkonstante aus Widerstand und Induktivität klein ist und ein steiler Anstieg des Stromes möglich ist. Beim Ausschalten ist der Widerstand der Diode dagegen klein, was die langsam abfallende Flanke verursacht. Eine zusätzliche Zuleitungsinduktivität der Laserdiode hat bei diesem Treibertyp großen Einfluss und sorgt für die Verbreiterung des Pulses [80]. Eine Verringerung der Pulsbreite lässt sich folglich durch möglichst kurze Zuleitungen der Diode erreichen, was die Auswahl auf Laserdioden mit integrierten Treibern richtet. Bei dem in der Arbeit verwendeten direkten Messverfahren auf Basis des ersten detektierten Ereignisses ist eine kurze Anstiegszeit von entscheidender Bedeutung für eine präzise Laufzeitermittlung. Dennoch, kürzere Pulse sind auch aus Gründen der Augensicherheit zu bevorzugen, da sie eine höhere Spitzenleistung erlauben.

Die Laserpulsrate liegt im typischen Betrieb bei 10 kHz, womit die maximale Laserleistung entsprechend (5.18) durch die gewichtete Energie eines Einzelpulses begrenzt ist. Unter der Annahme einer Punktquelle ist die zulässige Pulsenergie nach (5.15) 28,91 nJ. Aufgrund des vertikalen Öffnungswinkels von 1° ergibt sich nur eine teilweise Durchflutung der Messapertur von 31,4 %, was einer Fläche von $12,09 \text{ mm}^2$ entspricht. Daraus ergibt sich eine zulässige Energiedichte des Pulses von $2,39 \text{ mJ/m}^2$. Die gemessene Pulsenergiedichte beträgt $170,4 \text{ } \mu\text{J/m}^2$ in einem Abstand von 1 m, was umgerechnet auf die Messdistanz von 10 cm einem Wert von $17,04 \text{ mJ/m}^2$ entspricht und damit um den Faktor 7,13 über dem zulässigen Wert für die Laserklasse 1M liegt. Für Messungen in abgesperrten Bereichen zu Charakterisierungszwecken wird das Kamerasystem bei voller Leistung betrieben.

7.4.2 Benutzerinterface

Für die Ansteuerung und Datenauslese des Sensors sowie Visualisierung der Messdaten wurde ein Programm mit LabView samt GUI geschrieben. Das SoC dient im Wesentlichen der Ansteuerung des Sensors, die ausgelesenen Rohdaten werden unverändert an das Labview-Programm ausgegeben. In diesem erfolgt die weitere Verarbeitung der Daten: Umrechnung der 16-Bit-Rohdaten in

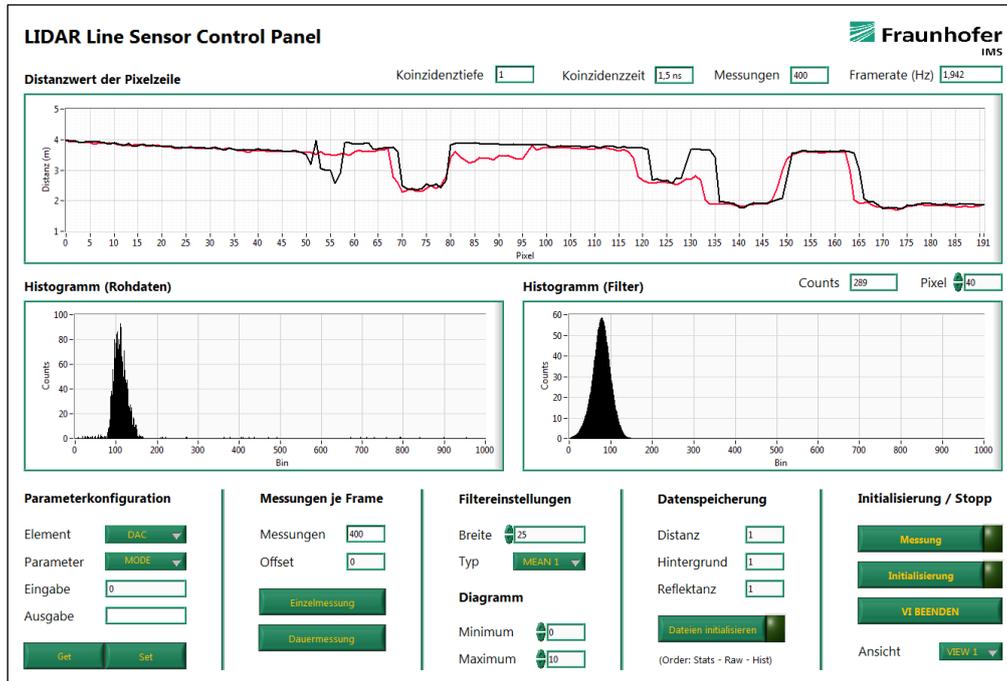


Abbildung 7.12: GUI zur Bedienung der Kamera und Messergebnisvisualisierung

binäre Laufzeitwerte gemäß (7.4), Bildung der pixelindividuellen Histogramme und Bestimmung der gemessenen Distanzen aus den Histogrammen. Das Interface zur Bedienung des Kamerasystems ist in Abbildung 7.12 dargestellt. Dabei zeigt das obere Diagramm die gemessenen Distanzwerte der beiden Sensorzeilen an, während die unteren Diagramme das Histogramm eines ausgewählten Pixels vor und nach der Filterung zeigen. Zusätzlich werden weitere allgemeine und pixelbezogene Informationen angezeigt: Aktuelle Bildrate, Anzahl der Laserpulse einer Distanzmessung, Einstellung der Koinzidenz und Anzahl der Detektionen im Histogramm. Die Schaltflächen im unteren Bereich dienen der Einstellung der verschiedenen Parameter, dem Starten und Stoppen der Distanzmessung sowie der Datenspeicherung zur späteren Analyse.

7.4.3 Datenauswertung

Der Algorithmus zur Auswertung der Rohdatenhistogramme hat einen großen Einfluss auf die Güte der Messung. Wie in Abbildung 7.13 gezeigt, liegt entsprechend (4.69) bei hoher Hintergrundlichtintensität der durch den Hintergrund

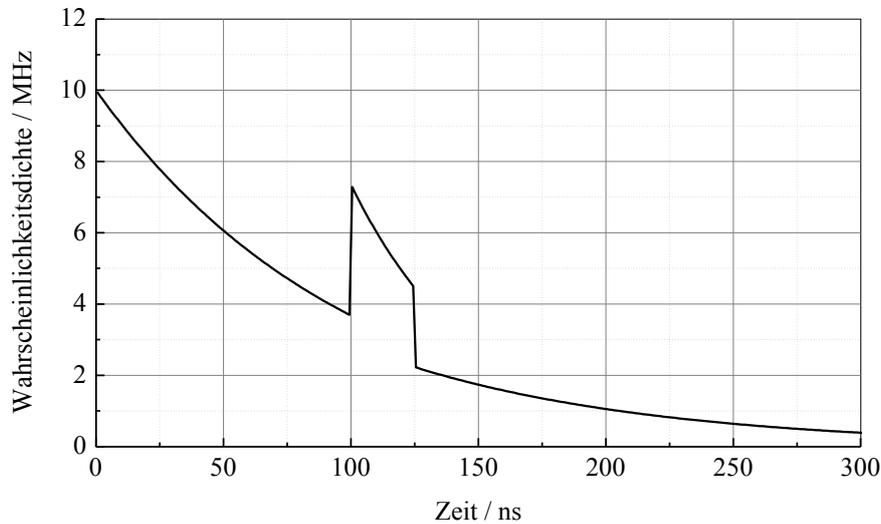


Abbildung 7.13: Berechnete PDF des ersten Ereignisses im direkten Verfahren

generierte Zählwert in den vorderen Klassen über dem Wert der lasergenerierten Ereignisse. Folglich resultiert die Verwendung des in Abschnitt 4.3.3 betrachteten Algorithmus, welcher die Distanz durch einfache Bestimmung des Maximums ermittelt, in fehlerhaften Messwerten.

Zur Lösung dieses Problems wird der exponentielle Verlauf des Hintergrundes aus dem Histogramm subtrahiert, sodass sich das Maximum, wie in Abbildung 7.14 gezeigt, bei der wahren Pulslaufzeit befindet. Die PDF der Ereignisdetektion des Hintergrundes ist gegeben zu

$$P_1(t)|_{r_L=0} = r_B \exp(-r_B t) \quad (7.5)$$

mit der Ereignisrate des Hintergrundes r_B . Aus dieser lässt sich mittels Integration der durch den Hintergrund generierte Erwartungswert jeder Klasse des Histogramms berechnen zu

$$\mu_i|_{r_L=0} = N_Z \int_{(i-1)T_{\text{Bin}}}^{iT_{\text{Bin}}} P_1(t)|_{r_L=0} dt \quad (7.6)$$

mit der Anzahl akkumulierter Einzelmessungen N_Z und der Breite der Klassen T_{Bin} . Da die Ereignisrate des Hintergrundes r_B im Allgemeinen unbekannt ist,

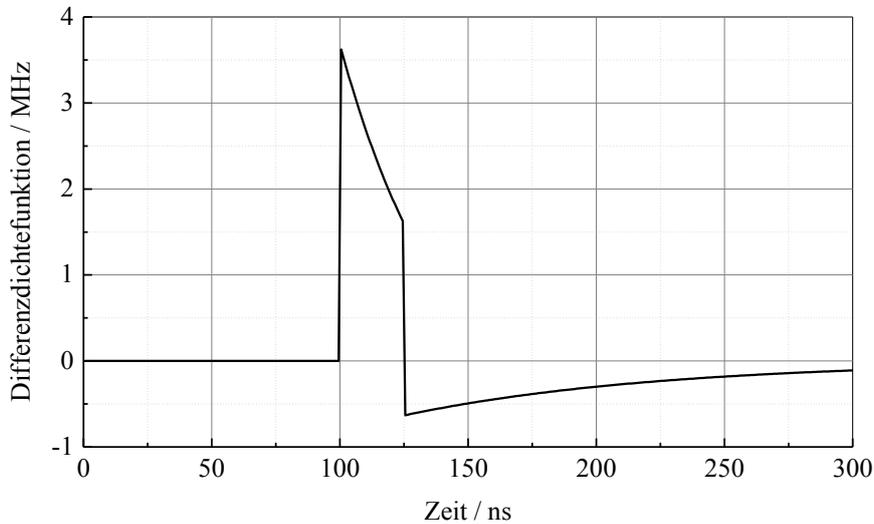


Abbildung 7.14: Berechnete Dichtefunktion nach der Hintergrundsubtraktion

muss diese zunächst bestimmt werden. Dafür gibt es zwei Möglichkeiten: Durch den Betrieb des Sensors im Zählmodus ohne aktiven Laser kann die Rate durch Zählung der Ereignisse in einem definierten Zeitfenster gemessen werden. Zum anderen kann die Rate aus dem Histogramm geschätzt werden.

Zur Schätzung der Ereignisrate des Hintergrundes wird die Gesamtzahl der Ereignisse im Histogramm betrachtet, welche unter Vernachlässigung des Laserpulses gegeben ist zu

$$N_B = \sum_{i=1}^{N_{\text{Bin}}} \mu_i |_{r_L=0} = N_Z (1 - \exp(-r_B N_{\text{Bin}} T_{\text{Bin}})) \quad (7.7)$$

mit der Anzahl der Klassen N_{Bin} . Da die Gesamtzahl N_B sich durch Summenbildung des Histogramms bestimmen lässt und damit bekannt ist, kann durch entsprechende Umformung die Rate r_B bestimmt werden. Es gilt für den Schätzwert der Ereignisrate

$$r_{B,\text{Est}} = -\ln\left(1 - \frac{N_B}{N_Z}\right) \frac{1}{N_{\text{Bin}} T_{\text{Bin}}}. \quad (7.8)$$

Vor der Subtraktion des Hintergrundes wird das Rohdatenhistogramm einer Mittelwertfilterung unterzogen. Dies senkt die Varianz der einzelnen Klassen,

wodurch das Histogramm sich der gewünschten Form entsprechend Abbildung 7.13 annähert und die Subtraktion das gewünschte Ergebnis liefert. Die Breite des Filters ist dabei von entscheidender Bedeutung: Ein zu schmaler Filter reduziert die Varianz nicht ausreichend, während ein zu breiter Filter den Puls verschmiert und eine präzise Ermittlung der Laufzeit erschwert. Ein guter Richtwert ist die Halbwertsbreite des Pulses. Die Bestimmung des Maximums im gefilterten und um den Hintergrund erniedrigten Histogramm liefert einen groben Wert für die Distanz.

Zur Steigerung der Präzision wird in einem zweiten Schritt ein Optimalfilter eingesetzt. Dieser zeichnet sich durch eine Impulsantwort entsprechend der gespiegelten gesuchten Signalform aus. Die Signalform lässt sich aus der zeitabhängigen Ereignisrate entsprechend (4.67) bestimmen. Hierfür ist die Ereignisrate des reflektierten Laserpulses erforderlich, welche, anders als die Ereignisrate des Hintergrundes, nur aus dem Rohdatenhistogramm geschätzt werden kann. Um eine möglichst präzise Schätzung zu erzielen, wird lediglich der Teilbereich des Histogramms betrachtet, in welchem sich der Puls befindet. Die Auswahl des Bereichs erfolgt auf Basis der zuvor beschriebenen groben Laufzeitbestimmung. Für die Ereignisse im Intervall $t = [t_0, t_1]$ gilt

$$N_B + N_L = N e^{-r_B t_0} \left(1 - \exp \left(-r_B (t_1 - t_0) - \int_0^{T_P} r_L(t) dt \right) \right) \quad (7.9)$$

mit der zeitabhängigen Ereignisrate des Laserpulses $r_L(t)$, welche der Form des Pulses entspricht, und der Pulsdauer T_P . Umgeformt ergibt sich

$$\int_0^{T_P} r_L(t) dt = -\ln \left(1 - \frac{N_B + N_L}{N \exp(-r_B t_0)} \right) - r_B (t_0 - t_1). \quad (7.10)$$

Die Schätzung der Ereignisrate des Laserpulses erfordert die Rate des Hintergrundes. Wurde diese mittels einer Schätzung ermittelt, kann zu deren Präzisierung eine erneute Schätzung aus dem Rohdatenhistogramm erfolgen, wobei nur der Teil des Histogramms vor der Ankunft des Pulses betrachtet wird. In diesem Bereich werden Ereignisse allein durch den Hintergrund verursacht, wodurch

eine Schätzung ohne Verfälschung durch den Laserpuls erreicht wird. Aus den Raten r_L und r_B wird im Weiteren die PDF des ersten Ereignisses nach (4.67) und auf dessen Basis die Koeffizienten des Optimalfilters bestimmt. Für den Wert der i -ten Klasse des gefilterten Histogramms gilt

$$f_i = \frac{1}{\sum_{i=1}^{N_F} \chi_i} \sum_{k=i}^{i+N_F-1} h_k \chi_{k-i} \quad (7.11)$$

mit der Breite des Optimalfilters N_F , den Werten des Rohdatenhistogramms h_i und den Filterkoeffizienten $\chi_i = P_1(T_{\text{TOF}} + iT_{\text{Bin}})$. Anschließend erfolgen die Subtraktion des Hintergrundes sowie die Bestimmung der Laufzeit anhand der Position des Maximums im Histogramm.

7.4.3.1 Ereignisratenschätzung

Da die Schätzung der Hintergrundereignisrate nach (7.8) die lasergenerierten Ereignisse im Rohdatenhistogramm vernachlässigt, wird der Hintergrund höher geschätzt als er tatsächlich ist. Der systematische Fehler lässt sich aus (7.8) und dem Integral der PDF nach (4.69) berechnen zu

$$\Delta r_{B,\text{Est}} = r_{B,\text{Est}} - r_B = \frac{r_L T_P}{N_{\text{Bin}} T_{\text{Bin}}}. \quad (7.12)$$

Es zeigt sich, dass der Fehler unabhängig von der Ereignisrate des Hintergrundes ist. Im in Abbildung 7.13 gezeigten Beispiel mit Ereignisraten von jeweils 10 MHz für den Hintergrund und den Laserpuls sowie einer Pulsdauer von 25 ns, beträgt der Schätzwert des Hintergrundes 10,83 MHz. Abbildung 7.15 zeigt die resultierende Dichtefunktion nach der Subtraktion des wahren Hintergrundes sowie dessen Schätzwertes. Dabei zeigt sich, ein höherer Schätzwert hat negative Werte im vorderen Bereich des Histogramms zur Folge, dennoch zeigt sich in der Höhe des empfangen Laserpulses kein signifikanter Unterschied, wodurch dieser in beiden betrachteten Fällen fehlerfrei lokalisiert werden kann. Folglich stellt ein höherer Schätzwert der Hintergrundrate kein Problem für die weitere Datenauswertung dar.

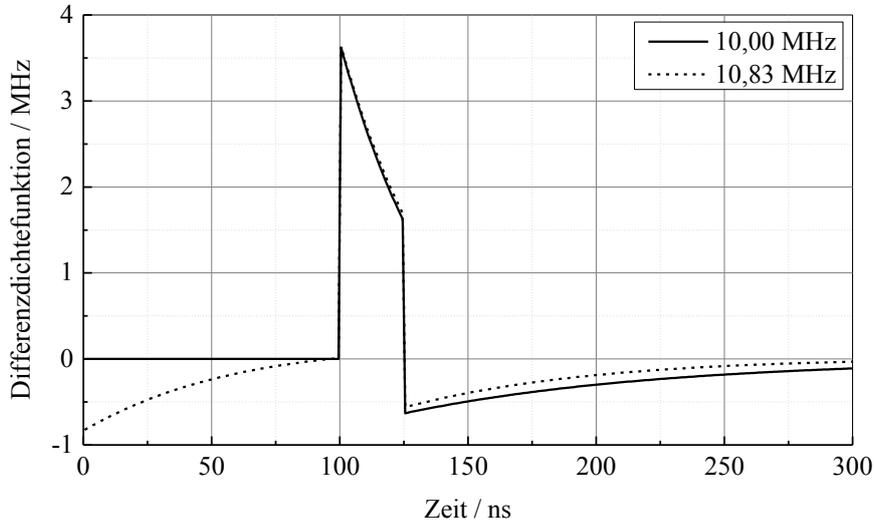


Abbildung 7.15: Berechnete Dichtefunktion nach Subtraktion des tatsächlichen sowie geschätzten Hintergrundes

Um allgemein zu beurteilen, wie sich eine Abweichung der Ratschätzung auf die Distanzmessung auswirkt, wird die Dichtefunktion des ersten Ereignisses nach der Hintergrundsubtraktion betrachtet. Wird die Rate zu gering geschätzt, kann der Wert der vorderen Klassen des Histogramms den Wert des Pulses überschreiten. Der untere Grenzwert wird festgelegt als der Wert, bei welchem die PDF bei $t = 0$ und $t = T_{\text{TOF}}$ den gleichen Wert aufweist. Mit (4.69) folgt

$$r_{\text{B,Est}}(1 - e^{-r_{\text{B,Est}} T_{\text{TOF}}}) = r_{\text{B}} - (r_{\text{L}} + r_{\text{B}})e^{-r_{\text{B}} T_{\text{TOF}}}. \quad (7.13)$$

Auch bei einem zu hohen Schätzwert können Klassen, welche im Histogramm vor der Ankunft des Pulses liegen, den höchsten Wert aufweisen. Das Maximum tritt hierbei jedoch nicht in der ersten Klasse auf, sondern an der Position

$$t_1 = \frac{2 \ln(r_{\text{B,Est}}) - 2 \ln(r_{\text{B}})}{r_{\text{B,Est}} - r_{\text{B}}}. \quad (7.14)$$

Analog zum unteren Grenzwert wird der Wert der Dichtefunktion zum Zeitpunkt t_1 und T_{TOF} gleichgesetzt. Damit ergibt sich

$$r_{\text{B,Est}}(e^{-r_{\text{B,Est}} T_{\text{TOF}}} - e^{-r_{\text{B,Est}} t_1}) = (r_{\text{L}} + r_{\text{B}})e^{-r_{\text{B}} T_{\text{TOF}}} - r_{\text{B}}e^{-r_{\text{B}} t_1}. \quad (7.15)$$

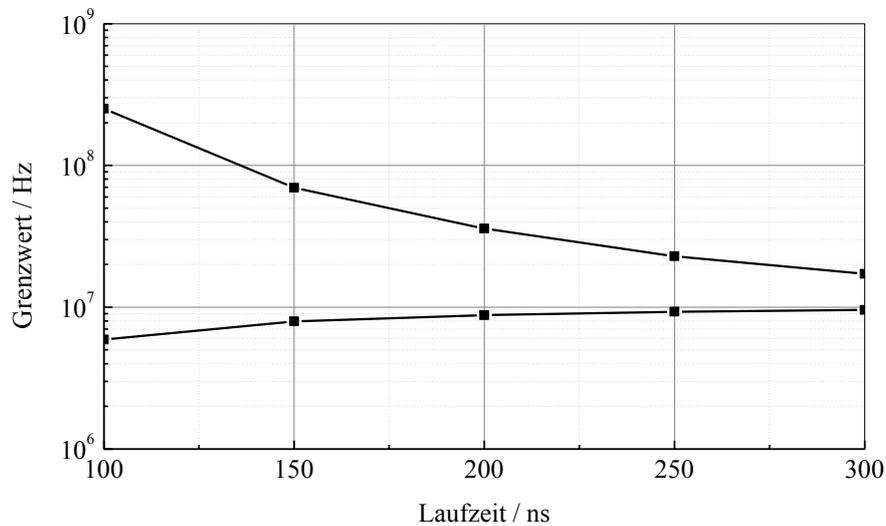


Abbildung 7.16: Grenzwerte der Ereignisratenschätzung

Da die beiden Gleichungen (7.13) und (7.15) keine analytische Lösung erlauben, erfolgt beispielhaft eine grafische Lösung. Abbildung 7.16 zeigt den oberen und unteren Grenzwert der geschätzten Ereignisrate für Ereignisraten des Hintergrundes und Laserpulses von jeweils 10 MHz für verschiedene Pulslaufzeiten. Dabei zeigt sich, dass die Grenzwerte stark von der Laufzeit abhängen und Abweichungen zu höheren Werten deutlich besser toleriert werden können. Abbildung 7.17 zeigt zur Veranschaulichung die Dichtefunktionen der beiden Grenzwerte bei einer Pulslaufzeit von 200 ns. Bei einer Rate entsprechend des oberen Grenzwertes liegt das relevante Maximum gemäß (7.14) bei 98,7 ns.

Im nächsten Schritt wird die Varianz der Ratenschätzung bestimmt. Hierfür wird das Histogramm ohne lasergenerierte Ereignisse betrachtet. Die Varianz setzt sich aus dem Quantisierungsfehler, welcher aus der diskreten Anzahl der detektierten Ereignisse N_B resultiert, und der Abweichung der Ereigniszahl selbst zusammen. Unter der Annahme gleichverteilter Werte ist der Quantisierungsfehler durch die Varianz der Gleichverteilung entsprechend

$$\sigma_Q^2 = \frac{\Delta r^2}{12} \quad (7.16)$$

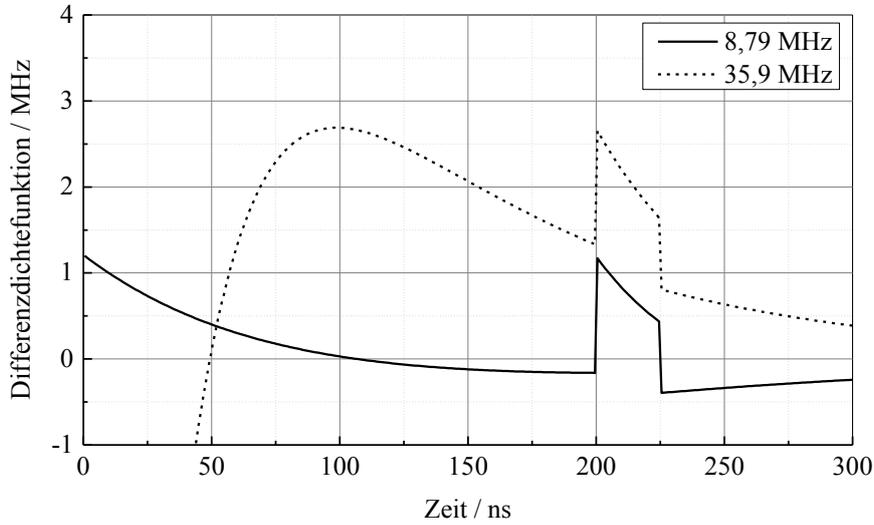


Abbildung 7.17: Berechnete Dichtefunktion nach Abzug der Grenzwertereignisraten

mit der Breite des Quantisierungsintervalls Δr gegeben. Die Breite des Intervalls ergibt sich aus der partiellen Ableitung von (7.8) nach der Anzahl der detektierten Ereignisse zu

$$\Delta r = \frac{\partial r_{B,Est}}{\partial N_B} = \frac{1}{N_{Bin} T_{Bin} (N - N_B)}, \quad (7.17)$$

womit die Breite des Quantisierungsintervalls mit steigendem N_B ansteigt. Mit der Ereigniszahl nach (7.7) folgt für den Quantisierungsfehler

$$\sigma_Q^2 = \frac{1}{12(N_{Bin} T_{Bin} N_Z \exp(-r_B N_{Bin} T_{Bin}))^2}. \quad (7.18)$$

Die Varianz der Ereigniszahl ist durch die Varianz der Binomialverteilung nach (4.4) gegeben, wobei die Anzahl der Versuch durch N_Z und die Wahrscheinlichkeit eines erfolgreichen Versuchs p durch die Poission-Verteilung mit dem Erwartungswert $r_B N_{Bin} T_{Bin}$ für einen Wert größer null entsprechend (4.8) gegeben ist. Damit folgt

$$\sigma_{N_B}^2 = N_Z (1 - \exp(-r_B N_{Bin} T_{Bin})) \exp(-r_B N_{Bin} T_{Bin}). \quad (7.19)$$

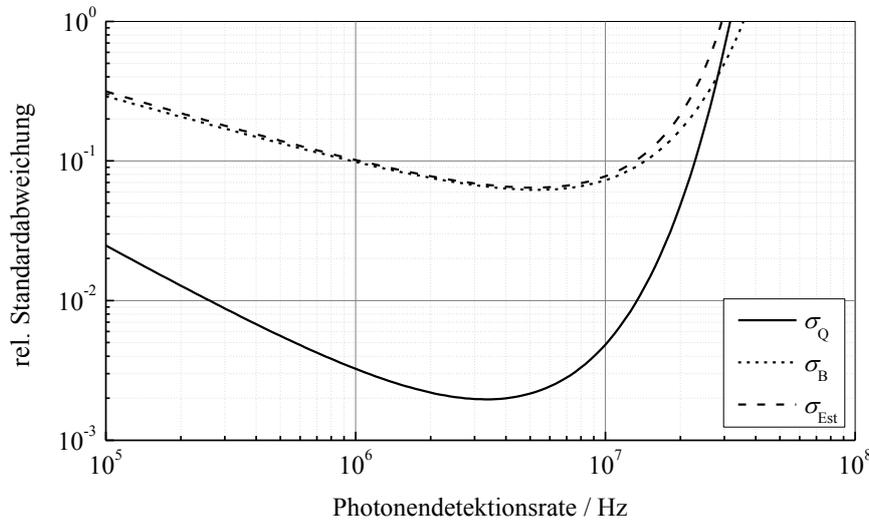


Abbildung 7.18: Theoretische Gesamt- und Teilabweichungen der Ereignisratenschätzung auf Basis des Histogramms

Die durch die Varianz der Ereigniszahl resultierende Varianz der Ereignisratenschätzung ergibt sich durch Anwendung der Fehlerfortpflanzung auf (7.8) zu

$$\sigma_B^2 = \frac{1 - \exp(-r_B N_{\text{Bin}} T_{\text{Bin}})}{N_{\text{Bin}}^2 T_{\text{Bin}}^2 N_Z \exp(-r_B N_{\text{Bin}} T_{\text{Bin}})} \quad (7.20)$$

Die Gesamtvarianz der Schätzung ist durch die Summe beider Teilvarianzen zu

$$\sigma_{\text{Est}}^2 = \frac{1 + 12N_Z(1 - \exp(-r_B N_{\text{Bin}} T_{\text{Bin}})) \exp(-r_B N_{\text{Bin}} T_{\text{Bin}})}{12(N_{\text{Bin}} T_{\text{Bin}} N_Z \exp(-r_B N_{\text{Bin}} T_{\text{Bin}}))^2} \quad (7.21)$$

gegeben. Abbildung 7.18 zeigt die Standardabweichung der Schätzung im Verhältnis zur Ereignisrate sowie die beiden Teilkomponenten der Abweichung für $N_Z = 400$ sowie eine Messdauer $N_{\text{Bin}} T_{\text{Bin}}$ von 300 ns. Die Abbildung zeigt, dass die Standardabweichung hauptsächlich durch die Varianz der Ereigniszahl bestimmt wird. Lediglich für hohe Ereignisraten dominiert der Quantisierungsfehler. Ursächlich hierfür ist die Logarithmusfunktion in (7.8), diese strebt gegen unendlich wenn N_B gegen N_Z strebt.

Um eine zuverlässige Schätzung höherer Raten zu ermöglichen, wird zur Schätzung nicht das vollständige Histogramm, sondern lediglich der vordere

Teilbereich betrachtet. Dies reduziert die Anzahl der detektierten Ereignisse und entspricht einer Verschiebung der Varianzkurven zu höheren Raten. Weiterhin erhöht sich dadurch die maximal schätzbare Ereignisrate, welche durch (7.8) mit $N_B = N_Z - 1$ zu

$$r_{B,Est,max} = -\ln\left(\frac{1}{N_Z}\right) \frac{1}{N_{Bin} T_{Bin}} \quad (7.22)$$

gegeben ist. Die Betrachtungen zeigen, dass die Wahl des Schätzbereichs von der zu erwartenden bzw. tolerierbaren Ereignisrate des Hintergrundes abhängt. Für geringe Ereignisraten des Hintergrundes entfällt das Problem der fehlerhaften Detektion des Maximums, wodurch dieser Fall für die Auslegung vernachlässigt werden kann.

Die Standardabweichung für eine Ereignisrate von 10 MHz liegt im Beispiel bei 7,76 % und überschreitet damit den unteren Grenzwert der tolerierbaren Abweichung nach Abbildung 7.16 von 4,5 % deutlich. Basierend auf diesen Betrachtungen kann zur Lösung dieses Problems die Standardabweichung nach (7.21) zum Schätzwert der Ereignisrate addiert werden. Dies reduziert die Abweichung nach unten, wodurch sich der Wert wahrscheinlicher im tolerierbaren Bereich bewegt. Eine erhöhte Abweichung zu höheren Ereignisraten hat entsprechend den Betrachtungen keinen negativen Einfluss auf die Qualität der Messung. Analog kann die Standardabweichung auch im Falle einer Messung der Hintergrundereignisrate zum Messwert addiert werden. Hierbei ist die Varianz gemäß der Poisson-Verteilung durch den Erwartungswert gegeben.

Zur Verifikation der theoretischen Untersuchung wird eine Simulation des direkten Laufzeitverfahrens unter Anwendung des beschriebenen Algorithmus zur Rohdatenauswertung durchgeführt. Abbildung 7.19 zeigt die ermittelte Erfolgswahrscheinlichkeit der Distanzmessung als Funktion der zur Eliminierung des Hintergrundes subtrahierten Ereignisrate für Eingangsraten von je 10 MHz für Hintergrund und Laserpuls sowie Laufzeiten von 150 ns und 200 ns. Die resultierenden Kurven decken sich mit den in Abbildung 7.16 gezeigten Grenzwerten. Für eine Laufzeit von 150 ns bzw. 200 ns liegen die Grenzwerte bei einer Erfolgswahrscheinlichkeit von ca. 15 % bzw. 20 %. Dies bestätigt die Erwartung aus der Theorie, dass eine Abweichung zu höheren Raten deutlich besser

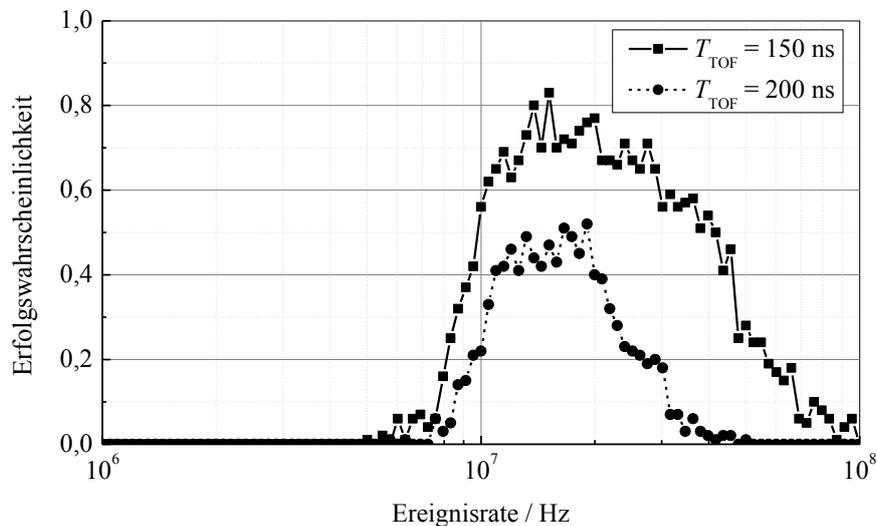


Abbildung 7.19: Simulierte Erfolgswahrscheinlichkeit als Funktion der zur Subtraktion des Hintergrundes verwendeten Ereignisrate

toleriert werden kann. Die Simulation zeigt darüber hinaus, dass eine höhere subtrahierte Ereignisrate die Güte der Messung sogar verbessert.

In Abbildung 7.20 ist die Simulation der Standardabweichung der Raten-schätzung dargestellt. Dabei zeigt sich eine gute Übereinstimmung mit der theoretischen Berechnung. Für höhere Raten lässt sich kein Wert ermitteln, da die Anzahl detektierter Ereignisse gleich der Anzahl Messzyklen ist, eine unendliche Rate geschätzt wird und damit eine Berechnung des Mittelwertes sowie der Standardabweichung verhindert wird.

7.4.3.2 Filterung

Neben der Subtraktion des Hintergrundes ist die Filterung des Histogramms entscheidend für die Zuverlässigkeit und Güte der Distanzmessung. Zur Wahl eines optimalen Filters werden hierfür die resultierenden PDFs des Pulses nach der Filterung mit verschiedenen Filtertypen betrachtet. Zunächst erfolgt die Betrachtung für einen idealen Rechteckpuls. Unter Vernachlässigung der Hintergrundereignisse sowie der Dunkelzählrate ist die PDF des ersten Ereignisses bei einem Rechteckpuls gegeben durch

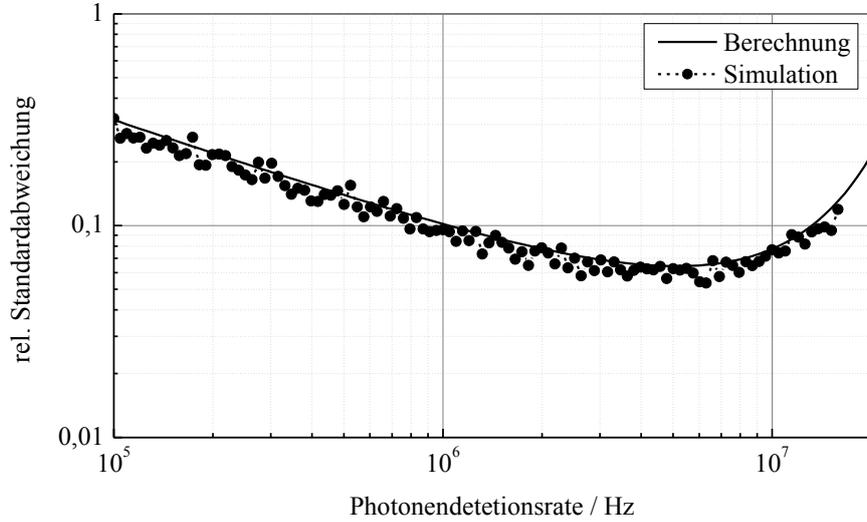


Abbildung 7.20: Simulierte und berechnete Standardabweichung der Ereignisraten schätzung

$$P_1 = r_L \exp(-r_L(t - T_{\text{TOF}})) (\Theta(t - T_{\text{TOF}}) - \Theta(t - T_{\text{TOF}} - T_P)). \quad (7.23)$$

Nach der Filterung mit einem Mittelwertfilter der Breite $T_F \leq T_P$ ergibt sich

$$P_1 * P_F = \begin{cases} 1 - e^{-r_L(t+T_F)} & \text{für } -T_F \leq t \leq 0 \\ e^{-r_L t} - e^{-r_L(t+T_F)} & \text{für } 0 < t < T_P - T_F \\ e^{-r_L t} - e^{-r_L T_P} & \text{für } T_P - T_F \leq t \leq T_P \end{cases} \quad (7.24)$$

für $T_{\text{TOF}} = 0$ und mit der Impulsantwort des Mittelwertfilters

$$P_F = \Theta(t) - \Theta(t - T_F). \quad (7.25)$$

Das Maximum der resultierenden PDF liegt bei $t = 0$, was der angenommenen Pulslaufzeit entspricht. Für hohe Ereignisraten des Laserpulses r_L reduziert sich die Halbwertsbreite der Funktion nach (7.23) entsprechend

$$FWHM = \min\left(-\frac{\ln(0,5)}{r_L}, T_P\right), \quad (7.26)$$

was zur Bildung eines Plateaus mit zunehmender Rate führt. Dies hat zur Folge, dass die Erwartungswerte der zur wahren Pulslaufzeit benachbarten Klassen

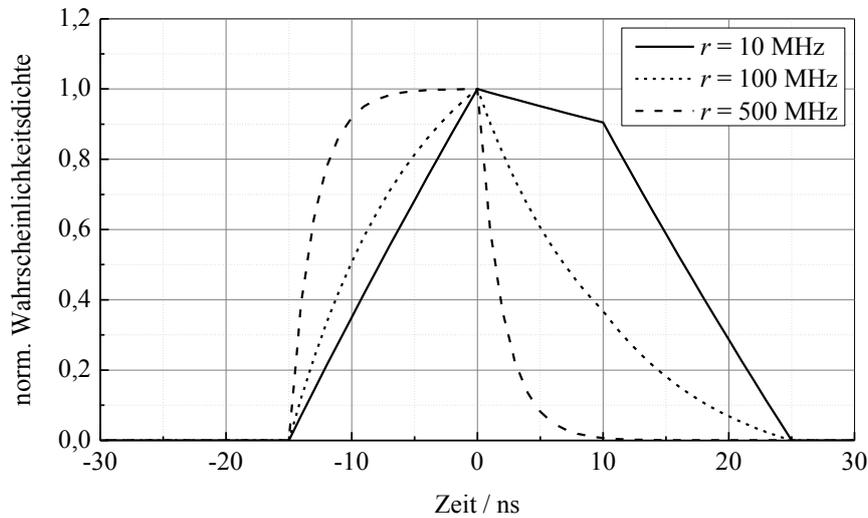


Abbildung 7.21: PDF nach Mittelwertfilterung für verschiedene Ereignisraten

sich dem Maximum annähern. Bedingt durch die Varianz der Zählwerte steigt damit die Wahrscheinlichkeit eines Maximums ungleich der Pulslaufzeit. Die maximal zu erwartenden Abweichung ist näherungsweise durch die Breite des Plateaus und somit des Mittelwertfilters T_F gegeben.

Abbildung 7.21 zeigt die PDF nach der Filterung gemäß (7.24) für verschiedene Ereignisraten des reflektierten Laserpulses r_L für eine Pulsweite von 25 ns und eine Filterbreite von 15 ns. Eine Filterbreite unterhalb der Pulsdauer resultiert in einem Plateau im Intervall $t = [T_{\text{TOF}}, T_{\text{TOF}} + T_P - T_F]$ für geringe Raten, da sich hierbei die PDF des Pulses einem Rechteck annähert. Für hohe Raten zeigt sich dagegen ein Plateau im Intervall $t = [T_{\text{TOF}} - T_F, T_{\text{TOF}}]$, da die PDF gegen die Diracsche Deltafunktion strebt. Für mittlere Ereignisraten erhält man eine Funktion mit einem definierten Maximum, was folglich die geringste Varianz und das beste Messergebnis erwarten lässt.

Eine mögliche Optimierung der Filterung ist eine adaptive Breite des Mittelwertfilters. Für die Wahl der idealen Breite wird die Steigung der resultierenden PDF nach (7.24) vor und nach der Pulslaufzeit betrachtet. Für eine möglichst fehlerfreie Detektion des Maximums müssen die Steigungen invers zueinander sein. Daraus folgt

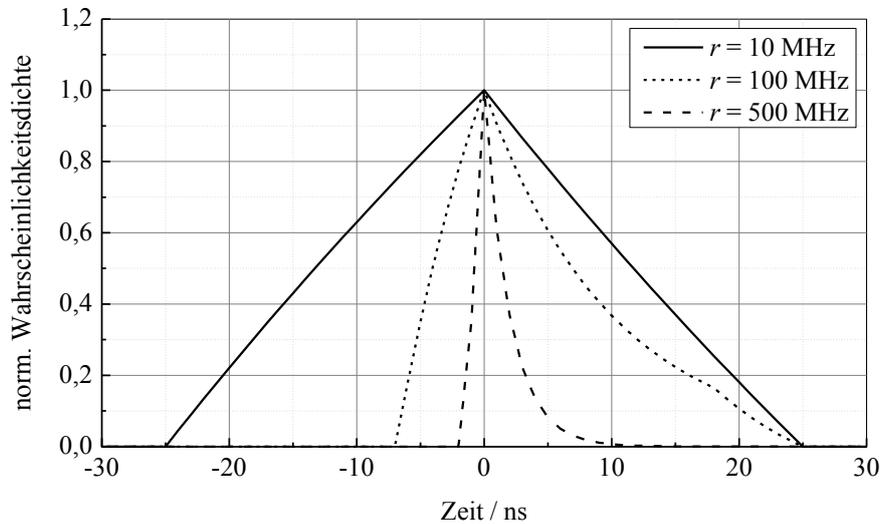


Abbildung 7.22: PDF nach adaptiver Mittelwertfilterung für verschiedene Ereignisraten

$$T_{F,Opt} = -\frac{\ln(0,5)}{r_L}, \quad (7.27)$$

was der Halbwertsbreite nach (7.26) entspricht. In Analogie wird auch die Filterbreite auf die Pulsweite begrenzt. Abbildung 7.22 zeigt die PDF nach der Filterung mit einem adaptiven Mittelwertfilter entsprechend (7.26) bzw. (7.27). Für die drei Ereignisraten 10 MHz, 100 MHz und 500 MHz ergeben sich dabei Filterbreiten von 25 ns, 6,93 ns und 1,39 ns. Durch die Anpassung der Filterbreite werden die Plateaus in den resultierenden Kurven verhindert und definierte Maxima erzeugt, was eine verbesserte Detektion derer ermöglicht. Eine Bestimmung der Steigung am Maximum ergibt

$$\left. \frac{\partial(P_1 * P_F)}{\partial t} \right|_{t=0} = r_L \frac{e^{-r_L T_{F,Opt}}}{1 - e^{-r_L T_{F,Opt}}} = r_L. \quad (7.28)$$

Eine andere Möglichkeit ist ein Optimal- bzw. Korrelationsfilter, bei diesem entspricht die Impulsantwort der zeitlich gespiegelten gesuchten Signalform [81]. Wird ein solches Filter verwendet, ist die resultierende Signalform durch die Autokorrelation des Pulses nach (7.23) gegeben zu

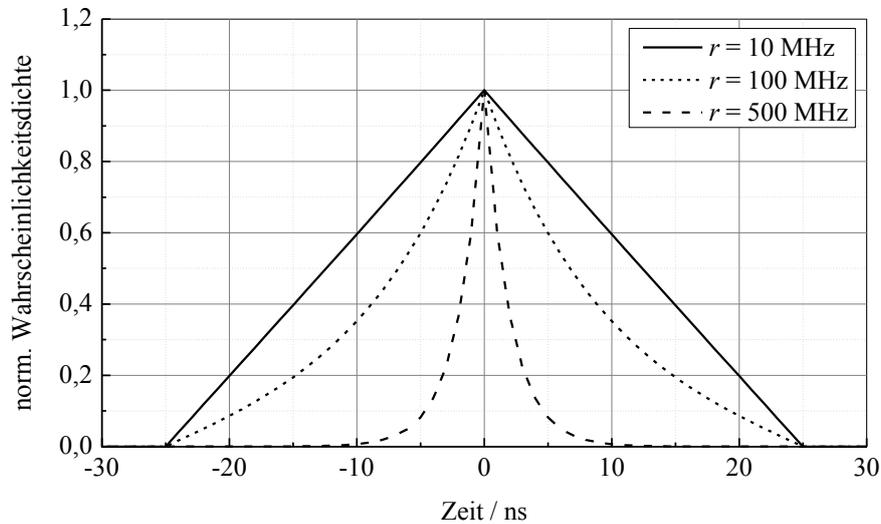


Abbildung 7.23: PDF nach Korrelationsfilterung für verschiedene Ereignisraten

$$P_1 * P_1 = \begin{cases} \frac{r_L}{2} (e^{-r_L t} - e^{r_L(t-2T_P)}) & \text{für } 0 \leq t \leq T_P \\ \frac{r_L}{2} (e^{r_L t} - e^{r_L(-t-2T_P)}) & \text{für } -T_P \leq t < 0 \end{cases} \quad (7.29)$$

womit die Steigung vor und nach dem Maximum identisch ist. Abbildung 7.23 zeigt die Signalform nach (7.29) nach der Filterung mittels Korrelationsfilter für verschiedene Ereignisraten. Hierbei ergibt sich ratenunabhängig ein definiertes Maximum, welches für höhere Raten spitzer ausfällt und damit eine bessere Detektion erlaubt. Generell liefert die Autokorrelationsfunktion stets eine achsensymmetrische Kurve, d. h. es gilt $f(x) = f(-x)$. Eine Berechnung der Steigung am Maximum liefert

$$\lim_{r_L \rightarrow \infty} \left(\frac{\partial(P_1 * P_1)}{\partial t} \Big|_{t=0} \right) = \lim_{r_L \rightarrow \infty} \left(r_L \frac{1 + e^{-2r_L T_P}}{1 - e^{-2r_L T_P}} \right) = r_L. \quad (7.30)$$

Hierbei zeigt sich, dass die Steigung nur für Ereignisraten gegen unendlich der Steigung des adaptiven Mittelwertfilters nach (7.28) entspricht und sonst geringer ausfällt. Neben der Steigung ist auch die Varianz der Klassen nach der Filterung für eine fehlerfreie Detektion des Maximums relevant. Hierbei gilt zu beachten, dass benachbarte Klassen eine hohe Korrelation aufweisen, da für

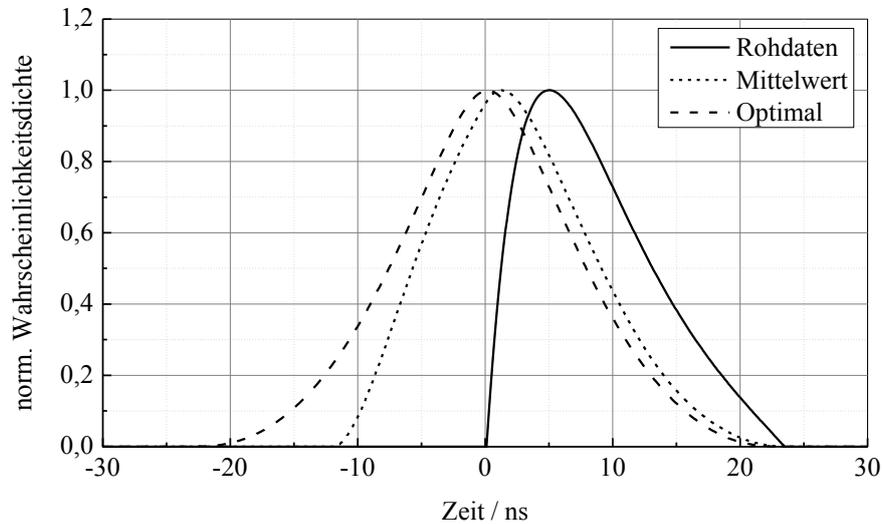


Abbildung 7.24: PDF der realen Pulsform vor und nach Filterung

deren Berechnung im Wesentlichen dieselben Rohdatenwerte verwendet werden. Die Betrachtung zeigt, dass für einen idealen Rechteckpuls ein adaptiver Mittelwertfilter eine geeignete Wahl darstellt, da dieser für geringe sowie sehr hohe Raten einem Optimalfilter nahe kommt.

Im Weiteren wird der Effekt eines Mittelwert- sowie Optimalfilters auf eine reale Pulsform, welche durch Näherung aus Abbildung 7.11 gewonnen wurde, betrachtet. Diese ist gegeben zu

$$r_L(t) = r_{L,\max} A (BC^2 t \exp(-Ct) - \exp(Dt)) \text{ für } 0 \leq t \leq 23,45 \text{ ns} \quad (7.31)$$

mit den Skalierungsfaktoren $A = 3,26 \text{ MHz}$ und $B = 4,375 \text{ } \mu\text{s}$ sowie den Formfaktoren $C = 192 \text{ MHz}$ und $D = 160 \text{ MHz}$. Der Skalierungsfaktor A ist dabei so gewählt, dass das Maximum der Funktion bei $r_{L,\max}$ liegt. Abbildung 7.24 zeigt die PDF des ersten Ereignisses berechnet mittels (4.67) für die zeitabhängige Ereignisrate nach (7.31) für $r_{L,\max} = 100 \text{ MHz}$ sowie die PDFs nach der Filterung mit einem Mittelwertfilter der Breite 13 ns und einem Optimalfilter. Dabei zeigt sich neben einer näherungsweise identischen Form der Maxima eine Verschiebung dessen im Falle der Mittelwertfilterung. Da diese Verschiebung von der Form der PDF, d. h. von der Ereignisrate des reflektierten Pulses, abhängt, stellt für Pulse beliebiger Form der Optimalfilter die bessere Wahl dar.

7.5 Charakterisierung

Dieses Kapitel beschreibt die Charakterisierung des Sensors. Alle durchgeführten Messungen sind dabei den vier Kategorien SPAD, TDC, Photonenkoinzidenz und Kamerasystem zugeordnet und nachfolgend beschrieben.

7.5.1 Single-Photon Avalanche Diode

Zur Charakterisierung der SPADs wurden die pixelindividuelle DCR, die Abhängigkeit der Totzeit von der Steuerspannung und die PDE als Funktion der Wellenlänge bestimmt.

7.5.1.1 Dunkelzählrate

Zur Bestimmung der DCR wird der Sensor bei Raumtemperatur in einer lichtisolierten Kammer im Zählmodus betrieben. Durch Zählung der Ereignisse ohne aktive Photonenkorrelation und mit allen SPADs über eine bekannte Zeitspanne wird die DCR jedes Pixels bestimmt. Da die Zählwerte der Poisson-Verteilung unterliegen, werden zur Reduktion der Messvarianz die Ereignisse über eine Gesamtdauer von 25 s gezählt. Da die SPADs nicht individuell deaktiviert werden können, wird die DCR des gesamten Pixels ermittelt. Abbildung 7.25 zeigt die DCR der Pixel in aufsteigender Reihenfolge über dem Anteil der Pixel. Daraus wird ersichtlich, dass 94,5 % der Pixel eine DCR unter 100 Hz aufweisen und 3,4 % einen Wert über 1 kHz zeigen. Weiterhin ergibt sich ein Mittelwert von 1,78 kHz und ein Median von 40,56 Hz. Die Werte entsprechen den in [9] ermittelten DCRs für die identischen SPADs mit einem geringeren Durchmesser von 10 μm bei einer Temperatur von 30 °C.

7.5.1.2 Totzeit

Die Totzeit hängt von der Beschaltung der Diode ab und ist daher keine direkte Eigenschaft der SPAD. Da die Totzeit jedoch für die weiteren Messungen relevant ist, wird sie an dieser Stelle gemessen. Die Totzeit wird in der Quenching-Schaltung durch Steuerung des Ladestromes einer internen Kapazität variiert. Als Stromquelle dient ein NMOS-Transistor an dessen Gate die Steuerspannung

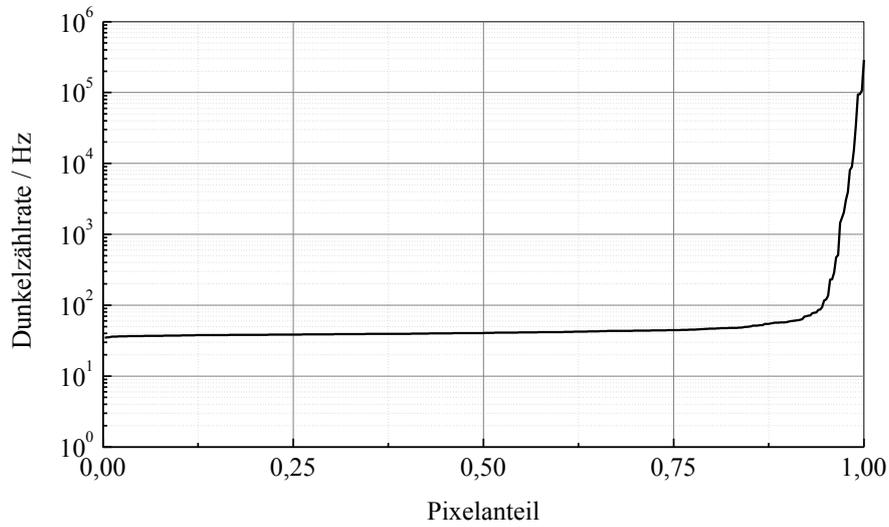


Abbildung 7.25: Dunkelzählrate der Pixel des Sensors

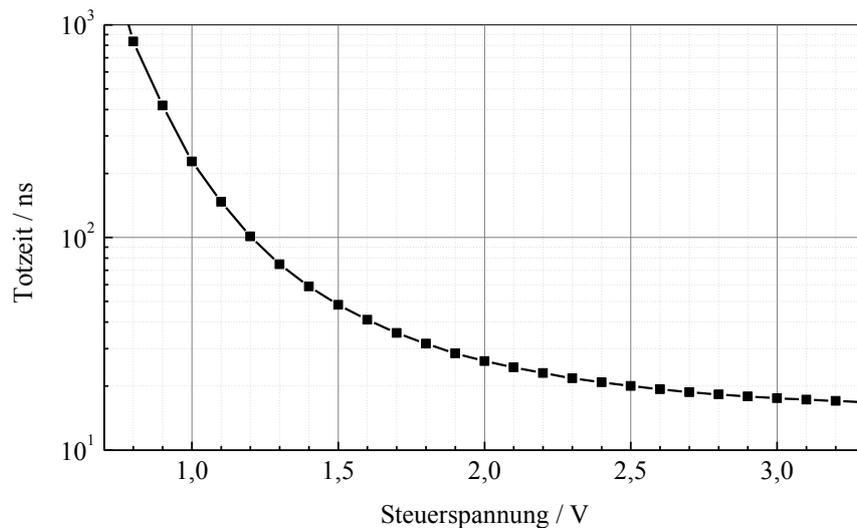


Abbildung 7.26: Totzeit als Funktion der Steuerspannung

anliegt, daher erhöht sich mit steigender Spannung der Ladestrom der Kapazität, wodurch die Totzeit sinkt. Die Bestimmung der Totzeit erfolgt durch Messung der Ereignisrate im Zählmodus ohne Photonenkorrelation und mit nur einer aktiven SPAD. In diesem Fall wird die Ereignisrate allein von der Totzeit gemäß (4.23) beeinflusst. Für sehr hohe Bestrahlungsstärken ist die Ereignisrate durch den Kehrwert der Totzeit gegeben. Abbildung 7.26 zeigt das Ergebnis

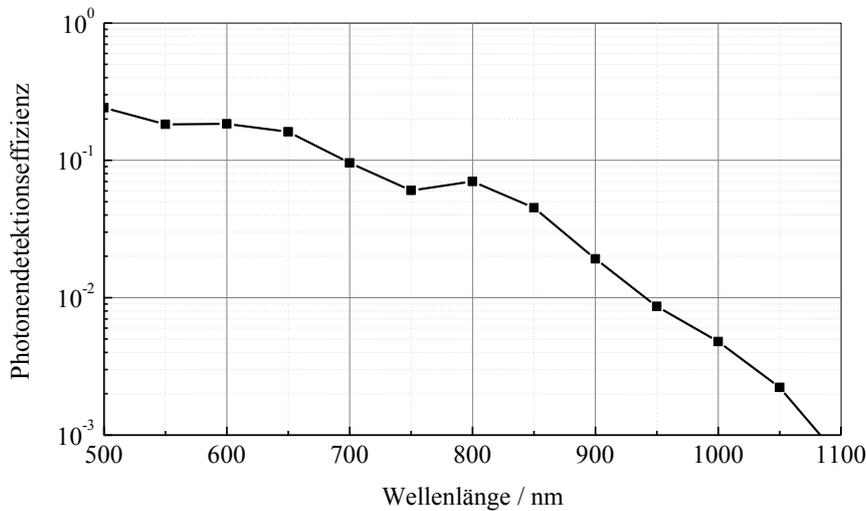


Abbildung 7.27: PDE der SPADs über der Wellenlänge bei 4,7 V Überspannung

der Messung. Für eine maximale Steuerspannung von 3,3 V liegt die Totzeit bei 16,79 ns. Mit abnehmender Spannung steigt die Totzeit exponentiell an und strebt für Spannungen unterhalb der Transistorschwellenspannung von etwa 0,7 V gegen unendlich.

7.5.1.3 Photonendetectionseffizienz

Für die Bestimmung der PDE wird der Sensor im Zählmodus ohne aktive Photonenkorrelation betrieben. Um Verfälschungen der gemessenen Rate durch die Verschaltung der vier SPADs zu vermeiden, ist während der Messung nur eine SPAD aktiv. Aus der gemessenen Rate ergibt sich die PDE zu

$$\eta_{\text{PDE}} = \frac{rhc}{S_L A_{\text{SPAD}} \lambda (1 - rT_T)} \quad (7.32)$$

mit der gemessenen Ereignisrate r , der Bestrahlungsstärke der Lichtquelle S_L , der Wellenlänge λ , der aktiven Fläche der SPAD A_{SPAD} und der Totzeit T_T . Gemäß den Betrachtungen in Abschnitt 2.4.1 zeigt das Messergebnis in Abbildung 7.27 einen Abfall der PDE mit steigender Wellenlänge. Die SPADs werden im Sensor mit einer Kathodenspannung von 31 V betrieben, was bei einer Durchbruchspannung von 26,3 V [9] einer Überspannung von 4,7 V entspricht. Bei

der Betriebswellenlänge des Lasers im Kamerasystem von 905 nm liefert die Messung eine PDE von 1,89 %. Das Resultat entspricht dem in [9] ermittelten Wert für die vom Aufbau her identischen SPADs mit leicht abweichenden Abmessungen und Überspannungen.

7.5.2 Time-to-Digital-Converter

Zur Charakterisierung des TDCs werden die Linearität sowie die Unsicherheit bestimmt. Die Bestimmung der Linearität erfolgt mit Hilfe des Testmodus sowie durch Messung gleichverteilter Ereignisse.

7.5.2.1 Linearität

Für eine erste Bestimmung der Linearität des TDCs wird der Sensor im Testmodus betrieben. In diesem wird das Start-Signal des TDCs nicht durch ein oder mehrere Photonendetektionen der SPADs erzeugt, sondern von einer externen Quelle zugeführt. Für die vorliegende Messung wird das Start-Signal vom FPGA generiert. Da dieser mit einer internen Frequenz von 160 MHz arbeitet, kann das Start-Signal nur in Schritten von 6,25 ns variiert werden. Für die Bestimmung der Linearität wird der vollständige Messbereich des TDCs von 1,28 μ s durchlaufen, was 205 Schritten entspricht. Bei jedem Schritt werden 1000 Messungen der Zeit durchgeführt, aus welchen der Stichprobenmittelwert gebildet wird, welcher die gemessene Laufzeit repräsentiert. Zur Bestimmung der integralen Nichtlinearität (INL) wird von den Stichprobenmittelwerten N_i der mittels einer linearen Approximation ermittelte Sollwert N_{linear} entsprechend

$$INL_i = N_i - N_{\text{linear}} \quad (7.33)$$

subtrahiert. Durch die Approximation wird ebenfalls ein konstanter Offset, welcher durch Laufzeiten auf der Platine und dem Chip entsteht, eliminiert.

Die INL des TDCs als Funktion der zu messenden Zeitspanne zeigt Abbildung 7.28. Hierbei zeigt sich ein wellenförmiger Verlauf der INL. Die Ursache hierfür findet sich in der internen Frequenz der PLL. Diese weist zu Beginn des Messfensters, welcher durch eine positive Flanke am Gate-Signal gekennzeichnet ist, eine geringere Frequenz auf als zum Ende hin. Da die Messung stets mit

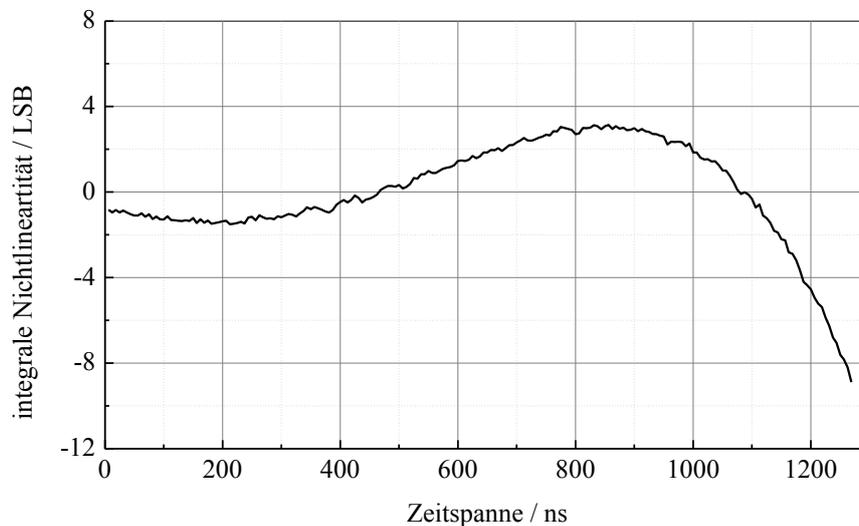


Abbildung 7.28: Integrale Nichtlinearität des TDCs ermittelt im Testmodus

dem Ende des Messfensters endet, zeigt sich im mittleren Zeitbereich eine positive Steigung der INL, da aufgrund der erhöhten Frequenz die Anzahl der Taktperioden, welche innerhalb der zu messenden Zeitspanne liegen, mit steigender Messdauer überproportional zunimmt und folglich die Differenz zwischen Messwert und Sollwert steigt. Ein analoges Verhalten zeigt sich am Ende des Diagramms: Die niedrigere Frequenz resultiert in einer negativen Steigung der INL. Infolge der Frequenzänderung innerhalb des Messfensters liefert die lineare Approximation der Ausgangskennlinie eine Steigung von $3,2108 \text{ LSB/ns}$, was einem LSB von $311,45 \text{ ps}$ entspricht. Der quadratische Mittelwert (RMS) des INL über alle 205 Zeitschritte ergibt sich zu $2,42 \text{ LSB}$.

Da, wie wiederholte Messungen der INL zeigen, dieses Verhalten systematischer Natur ist, kann es durch entsprechende Anpassung der Datenauswertung herausgerechnet werden. Mit einem Polynom vierter Ordnung lässt sich der Kurvenverlauf exakt nachbilden. Abbildung 7.29 zeigt die verbleibende Nichtlinearität bei Subtraktion des Polynoms vierter Ordnung, welche sich auf die Jitter des Taktsignals, der Laufzeitinterpolation und weiterer Schaltungskomponenten zurückführen lässt. Der RMS reduziert sich dabei auf $0,113 \text{ LSB}$. Ursächlich für die Frequenzänderung über dem Messfenster scheint eine durch die positive Flanke des Gate-Signals induzierte Änderung der Regelspannung

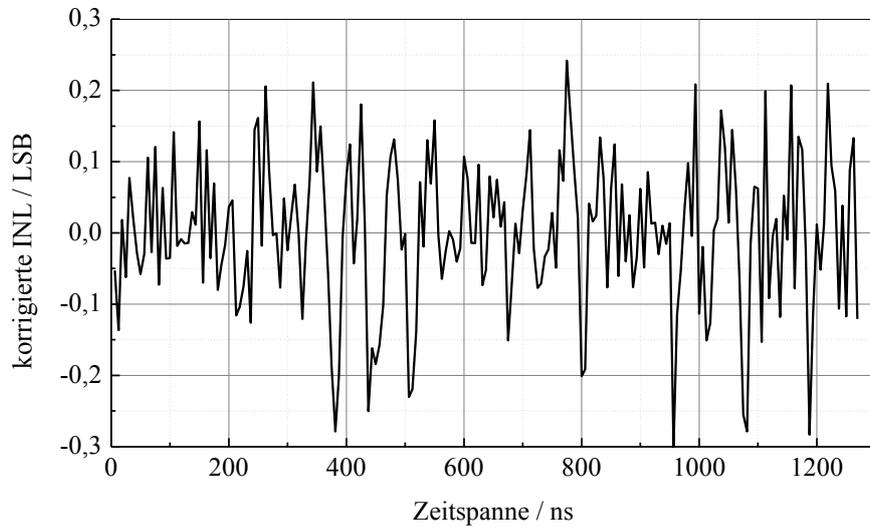


Abbildung 7.29: Korrigierte integrale Nichtlinearität des TDC

der PLL zu sein, was auch die einem Einschwingvorgang ähnelnde Form der INL erklärt. Da die Regelspannung der PLL nicht zugänglich ist, kann diese Vermutung nicht abschließend verifiziert werden.

Im Weiteren wird aus den Messdaten die differentielle Nichtlinearität (DNL) bestimmt. Diese berechnet sich für jeden Zeitschritt gemäß

$$DNL_i = \frac{N_{i+1} - N_i}{\Delta N} - 1 \quad (7.34)$$

mit dem gemessenen digitalen Zeitwert N_i beim Zeitschritt i und der Breite der Zeitschritte ΔN in LSB. Die Dauer der Zeitschritte von 6,25 ns entspricht 20 LSB. Abbildung 7.30 zeigt die mittels (7.34) berechnete DNL des TDCs mit einem RMS über alle Zeitschritte von 0,0089 LSB. Da die Zeitschritte deutlich über einem LSB liegen, wird die Abweichungen über mehrere Quantisierungsstufen des TDCs gemittelt, weshalb die auf diese Weise ermittelte DNL nur bedingt eine Aussage über die Linearität des TDCs erlaubt. Auch hier zeigt sich das durch die Frequenzvariation verursachte Verhalten.

Für eine zweite Bestimmung der Nichtlinearität des TDCs werden über dem Messbereich gleichverteilte Ereignisse generiert und aus der Abweichung zur tatsächlichen Verteilung die Nichtlinearität berechnet [82]. Zur Erzeugung

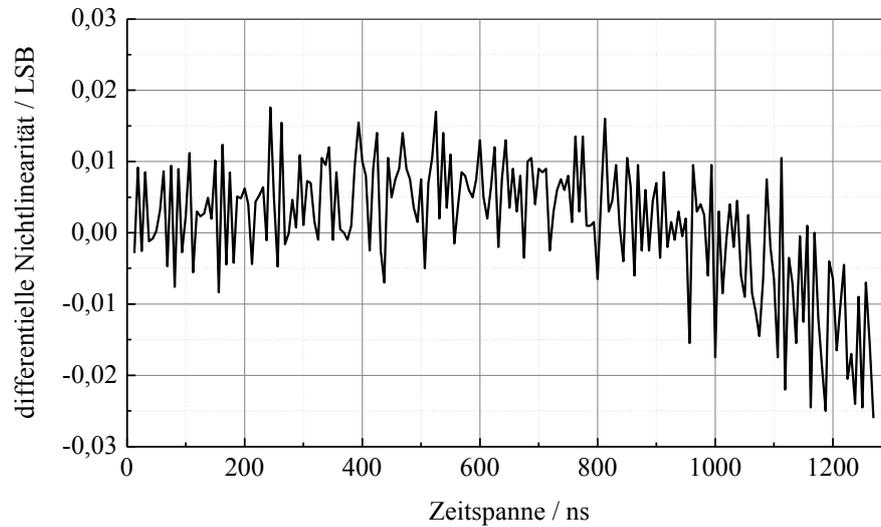


Abbildung 7.30: Differentielle Nichtlinearität des TDCs ermittelt im Testmodus

der Ereignisse werden die gemessenen Zeitwerte über viele Messungen im Histogramm akkumuliert. Erfolgt dies bei einer konstanten Bestrahlungsstärke, folgen die Ereignisse gemäß (4.67) einer Exponentialverteilung. Wird die Bestrahlungsstärke reduziert, geht diese in eine Gleichverteilung über. Da jedoch nur bei einer Beleuchtungsstärke von null eine Gleichverteilung erreicht wird, ist bei einer von null verschiedenen Bestrahlungsstärke stets eine exponentielle Verteilung der Ereignisse vorhanden. Um Verfälschungen zu minimieren, wird der exponentielle Verlauf vor der Nichtlinearitätsberechnung herausgerechnet. Die Berechnung der Nichtlinearität erfolgt für jedes der 384 Pixel individuell. Die Anzahl der Ereignisse in jeder Klasse liegt im Mittel bei 4500, was bei Poisson-Verteilung der Zählwerte einer Standardabweichung von 1,5 % entspricht, welche sich zur DNL addiert. Abbildung 7.31 (a) zeigt die nach

$$DNL_i = \frac{N_i}{\bar{N}} - 1 \quad (7.35)$$

mit dem Mittelwert aller Klassen des Histogramms \bar{N} berechnete DNL eines Pixels. Der RMS der DNL liegt bei 0,0909 LSB. Bei näherer Betrachtung der DNL in Abbildung 7.31 (b) zeigt sich ein in 16er-Schritten wiederholendes Muster, was die Vermutung zulässt, die Ursache in der Laufzeitinterpolation zu

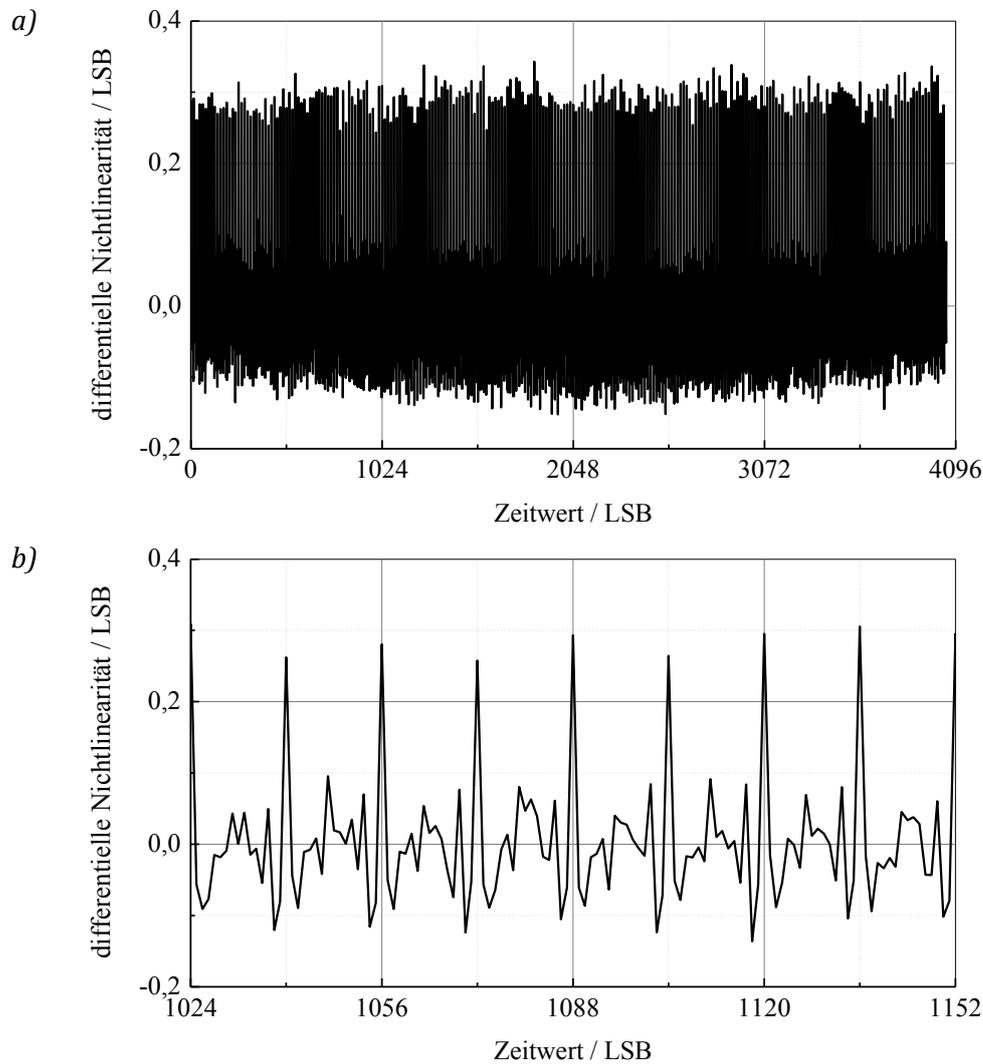


Abbildung 7.31: Differenzielle Nichtlinearität des TDCs (a) und Ausschnitt dieser zur Darstellung des sich wiederholenden Musters (b)

finden. Zur Ursachenfindung wird die Verteilung des Start-Interpolators betrachtet. Die Verteilung in Abbildung 7.32 zeigt eine deutlich erhöhte relative Häufigkeit bei einem Wert von 0, während die Werte 14 und 15 selten bis gar nicht auftreten. Dieses Verhalten resultiert aus einem Laufzeitunterschied zwischen dem Start-Signal, welches sich durch die Verzögerungskette fortpflanzt, und dem Signal, welches bei einer positiven Taktflanke generiert wird und den aktuellen Zustand der Verzögerungselemente einfriert. Durch eine geringere Verzögerung im zweitgenannten Signalpfad ist der Zeitbereich, in welchem der

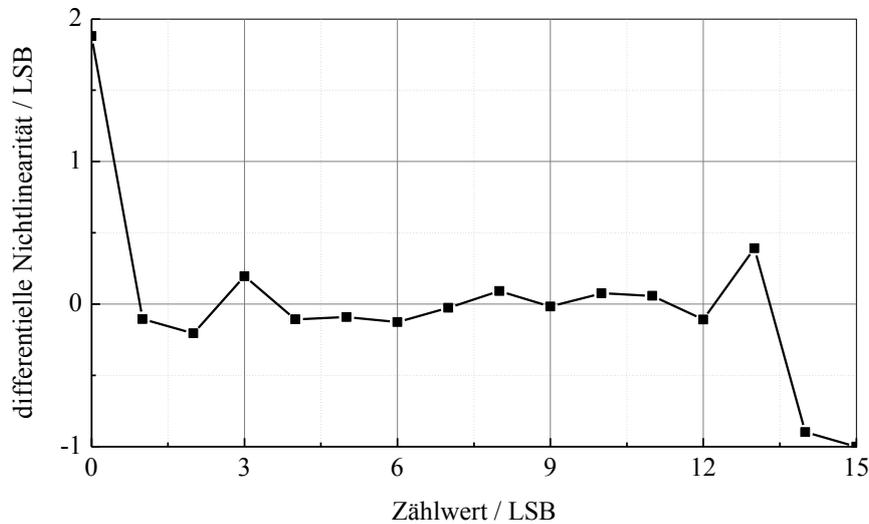


Abbildung 7.32: Differentielle Nichtlinearität des Start-Interpolators

Ausgang aller Elemente null ist, länger. Dies führt dazu, dass die Wahrscheinlichkeit, ein Ereignis in der ersten Klasse zu erhalten, höher ist. Da die Verzögerung der Kettenelemente dem Sollwert entspricht, ist im Gegenzug die Wahrscheinlichkeit, ein Ereignis in den hinteren Klassen zu empfangen, geringer. Diese zeitliche Verschiebung entspricht einem Offset, welcher zur Kompensation zum resultierenden Zeitwert addiert werden kann. Für eine Korrektur dieses Offsets in zukünftigen Sensoren muss eine zusätzliche Verzögerung in den Signalpfad für das Einfrieren des Verzögerungskettenzustandes bei einer positiven Taktflanke implementiert werden. Aus der DNL des Start-Interpolators lässt sich die notwendige Verzögerung zu 580 ps berechnen. Um Temperaturabhängigkeiten kompensieren zu können, wäre auch eine über eine analoge Spannung einstellbare Verzögerung denkbar, was allerdings eine zusätzliche Signalleitung erfordern würde, was aufgrund des geringen Pitches unerwünscht ist.

Aus der DNL wird die INL durch Summenbildung gemäß

$$INL_i = \sum_{k=1}^i DNL_k \quad (7.36)$$

berechnet. Abbildung 7.33 zeigt die aus der DNL in Abbildung 7.31 (a) berechnete INL mit einem RMS von 3,419 LSB. Dabei zeigt sich analog zu Abbildung

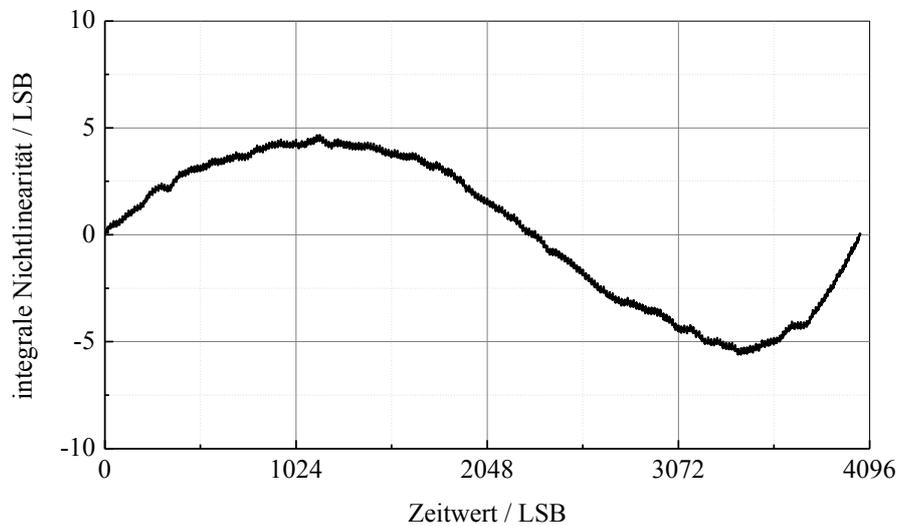


Abbildung 7.33: Integrale Nichtlinearität des TDCs ermittelt durch Messung gleichverteilter Ereignisse

7.28 das durch die Variation der Frequenz der PLL verursachte Verhalten. Die höhere Frequenz im mittleren Bereich sorgt für geringere Breiten der Klassen, weshalb diese geringere Werte enthalten, was bei der Bildung der INL eine negative Steigung zur Folge hat. Die Abweichung in der Form ergibt sich aus der exponentiellen Approximation des Histogramms. Weiterhin ist der sinusförmige Verlauf aufgrund der unterschiedlichen Berechnungsweise negiert. Wird die INL des Start-Interpolators berechnet, erhält man einen näherungsweise konstanten Wert, was einem konstanten Offset entspricht.

Der RMS der INL und DNL aller Pixel des Doppelzeilensensors ist in Abbildung 7.34 dargestellt. Im Mittel aller Pixel ergeben sich eine INL von 3,39 LSB und eine DNL von 0,084 LSB.

7.5.2.2 Präzision

Die Messung der Zeitwertestreuung des TDCs wird im Testmodus ermittelt. Hierfür wird die Standardabweichung von 10.000 Messwerten bei konstantem Sollwert bestimmt. Abbildung 7.35 zeigt das Resultat der Messung. Dabei zeigt sich, dass die Unsicherheit mit der gemessenen Zeitspanne ansteigt, da sich der Jitter des Taktsignals, welches der groben Zeitmessung dient, aufsummiert. Zur

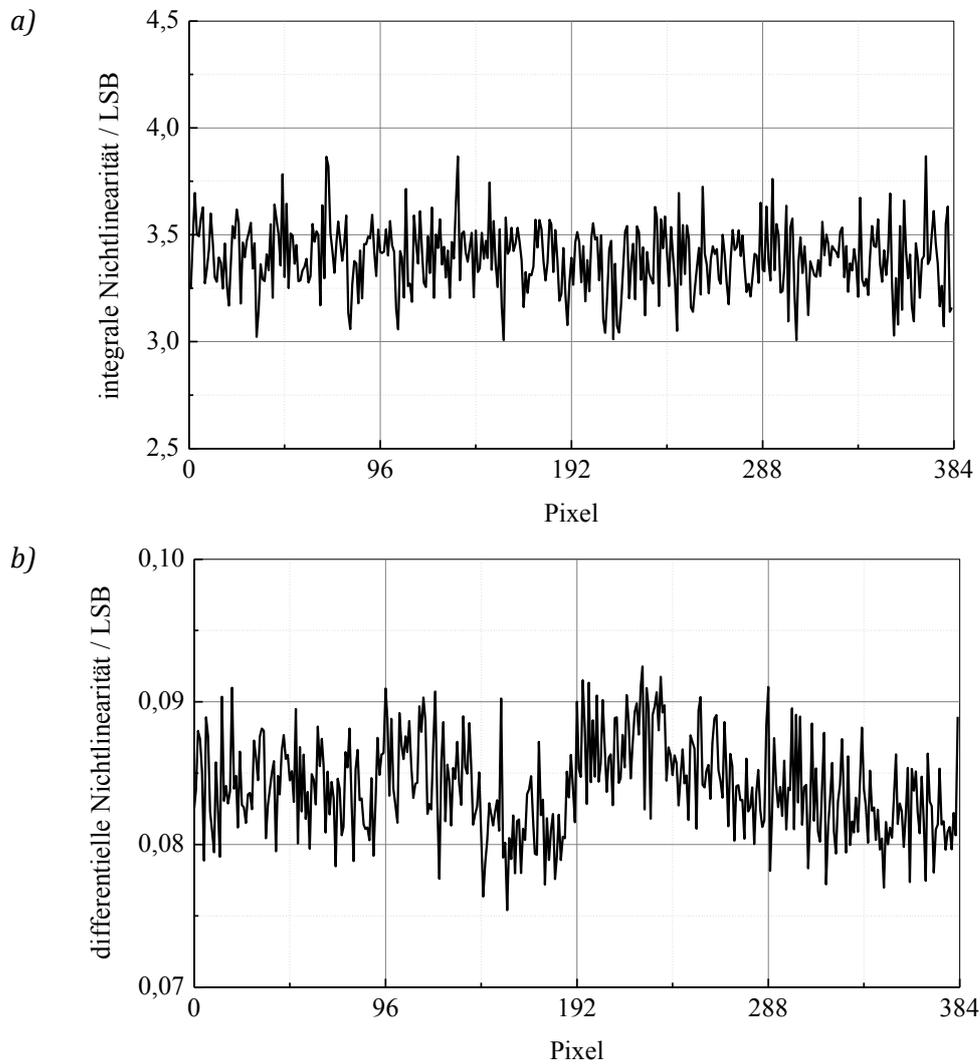


Abbildung 7.34: RMS der integralen (a) und differentiellen (b) Nichtlinearität des Sensors ermittelt durch Messung gleichverteilter Ereignisse

Verifikation der Ursache wird der Jitter des Taktsignals, das über einen Taktteiler mit dem Faktor 8 an ein Ausgangspad des Chips geführt ist, gemessen. Für eine Taktperiode von 40 ns ergibt sich eine Standardabweichung des Jitters von 137,7 ps, was 0,44 LSB entspricht. Werden 32 Taktperioden betrachtet, was der maximalen Messdauer von 1,28 μ s entspricht, wird ein Jitter von 1,553 ns gemessen. Dies entspricht 4,97 LSB und damit näherungsweise der bei dieser Messdauer gemessenen Unsicherheit des TDCs. Die Messung zeigt, dass sich die Abweichung aus dem Jitter des Taktsignals sowie der Interpolation, welche bei

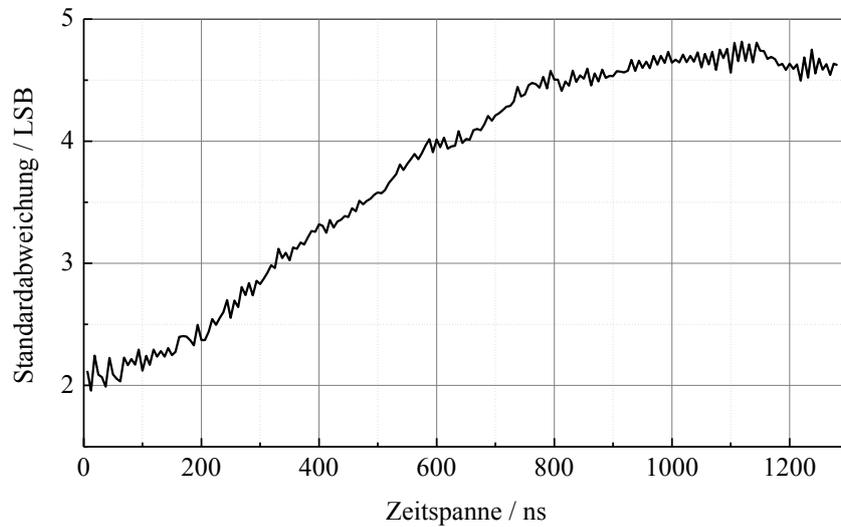


Abbildung 7.35: Standardabweichung der Laufzeitmessung

geringen Zeiten dominiert, zusammensetzt. Der Jitter des TDCs sorgt für eine Verbreiterung des detektierten Signalpulses im Histogramm der Zeitwerte, was eine präzise Bestimmung der Laufzeit erschwert sowie das SNR reduziert. Da im Sensor das Starten der Laufzeitmessung bei einer Ereignisdetektion und das Stoppen am Ende des Messfensters erfolgt, tritt besonders bei geringen Distanzen eine hohe Unsicherheit auf. Zur Lösung kann das Starten der Laufzeitmessung global zu Beginn der Messung erfolgen, während bei einer Ereignisdetektion die Messung pixelindividuell gestoppt wird. In diesem Fall entspricht die gemessene Zeit direkt der Laufzeit, wodurch geringere Unsicherheiten bei kurzen Distanzen erreicht werden. Allerdings erhöht sich die Unsicherheit bei höheren Distanzen, was zu einer Reduktion der Reichweite führen kann.

7.5.3 Photonenkorrelation

Zur Überprüfung der Wirksamkeit der Detektion korrelierter Einzelphotonen zur Unterdrückung des Hintergrundlichtes werden Messungen zur Verifikation des in Abschnitt 6.1.1 gezeigten theoretischen Modells durchgeführt. Für diesen Zweck wird der Sensor in einer lichtisolierten Kammer platziert und mit einem Illuminator mit konstanter Beleuchtungsstärke bestrahlt. Zur Ermittlung der Ereignisraten wird der Sensor im Zählmodus mit verschiedenen Einstellungen

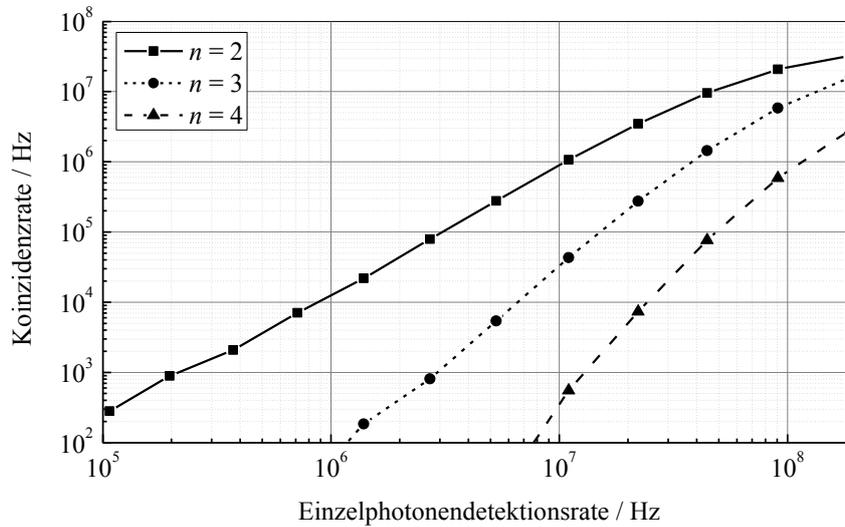


Abbildung 7.36: Ereignisrate der Photonenkoinzidenz für eine Koinzidenztiefe von 2 bis 4 bei konstanter Koinzidenzzeit von 16 ns und 4 aktiven SPADs

der Koinzidenz sowie ohne aktive Koinzidenz betrieben. Aus der Anzahl der in einer bekannten Zeitspanne gezählten Ereignissen wird die Ereignisrate bestimmt. Wird dieser Vorgang für eine steigende Beleuchtungsstärke durchgeführt, ergibt sich daraus eine Kurve der Ereignisrate unter Verwendung von Koinzidenz als Funktion der Einzelphotonendetektionsrate ohne Koinzidenz. Als Referenzzählrate dient die Zählrate ohne aktive Koinzidenz sowie bei minimaler Koinzidenzzeit und mit einer aktiven SPAD, da bei diesen Einstellungen die Messung allein durch die Totzeit der SPAD gemäß (4.23) beeinflusst wird und eine Korrektur sich entsprechend einfach gestaltet. Für die Messung wird der Sensor mit einer Bestrahlungsstärke von $0,1 \mu\text{W}/\text{cm}^2$ bis $204,8 \mu\text{W}/\text{cm}^2$ bei einer Wellenlänge von 850 nm belichtet. Bei jedem Schritt wird die Intensität verdoppelt, womit sich insgesamt 12 Schritte ergeben. Die Belichtungszeit beträgt 5 ms, womit die Auflösung bei 200 Hz liegt.

In Abbildung 7.36 sind die Ergebnisse der Messung dargestellt. Dem Modell aus Abschnitt 6.1.1 entsprechend zeigt sich eine abnehmende Koinzidenzrate mit steigender Koinzidenztiefe. Dabei ist eine Abhängigkeit gemäß

$$r_{\text{Coin}} \propto r^n \quad (7.37)$$

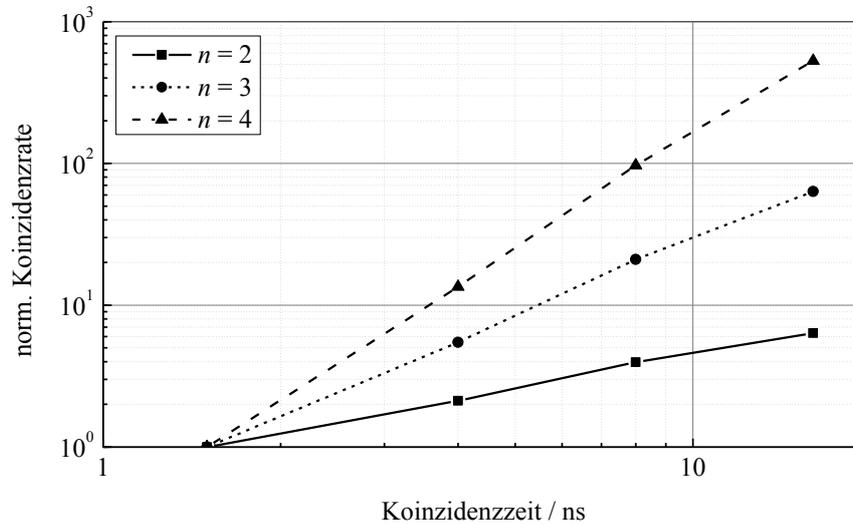


Abbildung 7.37: Normierte Ereignisrate der Photonenkoinzidenz für eine Koinzidenzzeit von 1,5 bis 16 ns bei Koinzidenztiefen von 2 bis 4 und 4 aktiven SPADs

zu erkennen. Für höhere Einzelphotonenraten zeigt sich aufgrund der Totzeit sowie der Tatsache, dass die Koinzidenzrate die Einzelphotonenrate nicht übersteigen kann, eine Sättigung der Koinzidenzrate. Abbildung 7.37 zeigt die normierte Koinzidenzrate als Funktion der Koinzidenzzeit für verschiedene Koinzidenztiefen. Hierbei zeigt sich dem Modell entsprechend eine von der Koinzidenztiefe abhängige Proportionalität der Koinzidenzrate gemäß

$$r_{\text{Coin}} \propto \vartheta^{n-1}. \quad (7.38)$$

Eine Änderung der Anzahl aktiver SPADs hat, abgesehen vom Einfluss der Totzeit, dieselben Auswirkungen wie eine erhöhte Einzelphotonenrate. Die Messungen der Koinzidenz bestätigen das in Abschnitt 6.1.1 entworfene Modell.

7.5.4 Kamerasystem

Für die weiteren Messungen zur Validierung der betrachteten Messkonzepte wird der Sensor in das in Abschnitt 7.4 beschriebene Kamerasystem „Owl“ eingebaut. Die beiden Laserquellen werden dabei so ausgerichtet, dass jede eine der beiden Zeilen des Sensors belichtet.

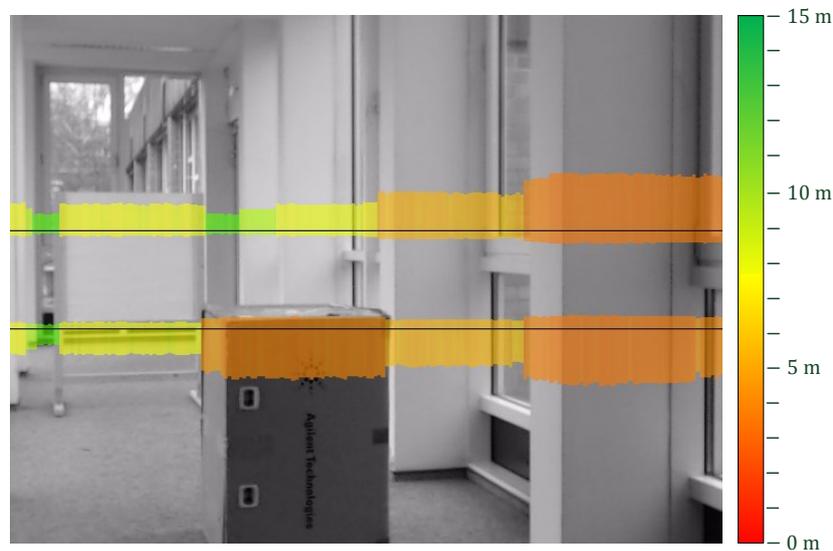


Abbildung 7.38: Zielszene zur Funktionsprüfung des LiDAR-Systems

7.5.4.1 Funktionstest

Im ersten Schritt wird die grundsätzliche Funktion des Kamerasystems geprüft. Hierfür wird das System im Innenraum, d. h. bei geringer Hintergrundlichtintensität, zur Distanzmessung einer beliebigen Szene verwendet. Abbildung 7.38 zeigt die verwendete Szene mit Hindernissen in verschiedenen Distanzen zur Kamera. Die Distanzinformation ist dem Grauwertbild der Szene mittels farbiger vertikaler Balken überlagert dargestellt. Die Farbe der Balken sowie deren Höhe gibt dabei Auskunft über die gemessene Distanz des jeweiligen Pixels. Die horizontalen Linien illustrieren die Position des FOV der Zeilen in der Szene. In der Abbildung lassen sich deutlich die unterschiedlichen Distanzen der Säulen erkennen. Die Kamera liefert die Distanzinformation jedes Pixels in Polarkoordinaten, d. h. die gemessene Distanz entspricht dem Betrag, während die Position des Pixels in der Zeile zusammen mit dem verwendeten Objektiv den Winkel ergibt. Für eine übersichtlichere Darstellung der Distanzwerte werden diese in kartesische Koordinaten umgerechnet. Abbildung 7.39 zeigt die Messergebnisse für die obere (Zeile 1) und untere Zeile (Zeile 2) des Sensors. Verglichen mit der oberen Zeile zeigt sich als abweichendes Objekt in der unteren Zeile der Karton bei einer Distanz von 2,13 m. Dieser wird, wie in Abbildung 7.38 gezeigt, nur

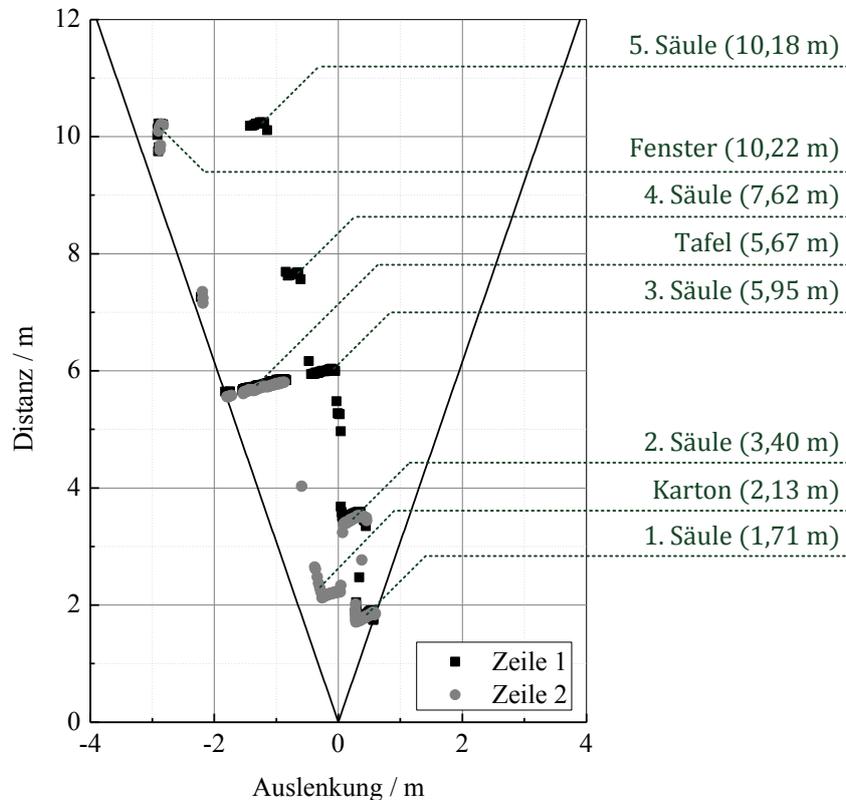


Abbildung 7.39: Distanzwerte der oberen und unteren Zeile

von der unteren Zeile erfasst. Alle weiteren Objekte der Szene werden von beiden Zeilen detektiert und sind entsprechend unter Angabe der Distanz gekennzeichnet. Neben den Säulen auf der rechten Seite finden sich auch Pixel, welche die dazwischenliegenden Fenster erkennen.

An Rändern einzelner Objekte erfassen die Pixel das reflektierte Lasersignal verschiedener Objekte in unterschiedlicher Distanz. Dies führt dazu, dass je nach Auswertung der Rohdaten, die ermittelte Distanz zwischen beiden Werten springt oder dazwischen liegt. Ein Pixel der unteren Zeile ist dafür ein Beispiel: Die gemessene Distanz von 4 m ist in der Szene nicht existent. Allgemein kann die Datenauswertung auf die Detektion mehrerer Maxima in Histogramm ausgelegt werden, sodass in solch einem Fall beide oder nur der geringe Distanzwert ausgegeben wird. Sind Objekte sehr nah beieinander, kann dies zu einer Verbreiterung des reflektierten Pulses durch die unterschiedlichen Laufzeiten

führen, was sich nachträglich nicht mehr korrigieren lässt. Auch unter diesem Gesichtspunkt sich kurze Laserpulse zu bevorzugen.

Generell zeigt die Messung, dass der entworfene Sensor funktionsfähig und eine Distanzmessung auf Basis der Lichtlaufzeit möglich ist. Weitere Messungen zu Beurteilung der Güte folgen in den weiteren Unterkapiteln.

7.5.4.2 Reichweite ohne Hintergrundlicht

Diese Messung dient der Bestimmung der Reichweite des Systems. Dafür wurde die Distanz eines Zielobjektes im Innenraum, d. h. bei geringem Hintergrundlicht, gemessen. Als Zielobjekte dienen Holzplatten im Format 80 x 80 cm², welche mit mattem weißem und schwarzem Papier beklebt sind, da dieses einem idealen Lambertschen Reflektor nahe kommt. Dabei wurde, begrenzt durch die zur Verfügung stehende Räumlichkeit, eine Distanz von 5 m bis 50 m in 5-m-Schritten gemessen. Als Referenzwert wurde die Entfernung mit dem Distanzmessgerät CX80 des Herstellers Precaster Enterprises Co. gemessen. Dieses hat eine Genauigkeit von 1,5 mm bei einer Reichweite von bis zu 80 m [83]. Als Kriterium zur Beurteilung der Messleistung werden bei jedem Distanzwert 100 Messungen durchgeführt, wobei für jede Messung 400 Laserpulse verwendet werden. Aus den 100 Distanzwerten werden der Mittelwert, die Standardabweichung sowie die Erfolgswahrscheinlichkeit bestimmt. Diese besagt, dass die wahre Distanz mit einer festgelegten Toleranz, welche für die nachfolgende Berechnung auf 2 % festgelegt wird, gemessen werden kann.

Die Abweichung der Mittelwerte von den Referenzdistanzen sowie die zugehörigen Standardabweichungen sind in Abbildung 7.40 dargestellt. Die Abweichung des Mittelwertes zeigt für das weiße Zielobjekt bis zu einer Distanz von 30 m eine Zunahme auf einen Wert von ca. 25 cm. Dieses Verhalten lässt sich durch die in Abschnitt 7.5.2.1 beschriebene schwankende Frequenz der PLL über dem Messfenster erklären: Durch die geringe Frequenz zu Beginn des Messfensters fallen bei einer Erhöhung der Messdistanz weniger Taktflanken aus dem Messbereich als bei der Sollfrequenz. Infolgedessen erhöht sich die gemessene Distanz weniger, was zu einer steigenden Differenz zwischen Sollwert und Messwert führt. Für das schwarze Objekt zeigt sich bereits ab einer Distanz von 25 m eine starke Abweichung vom Sollwert. Diese entsteht, da ab

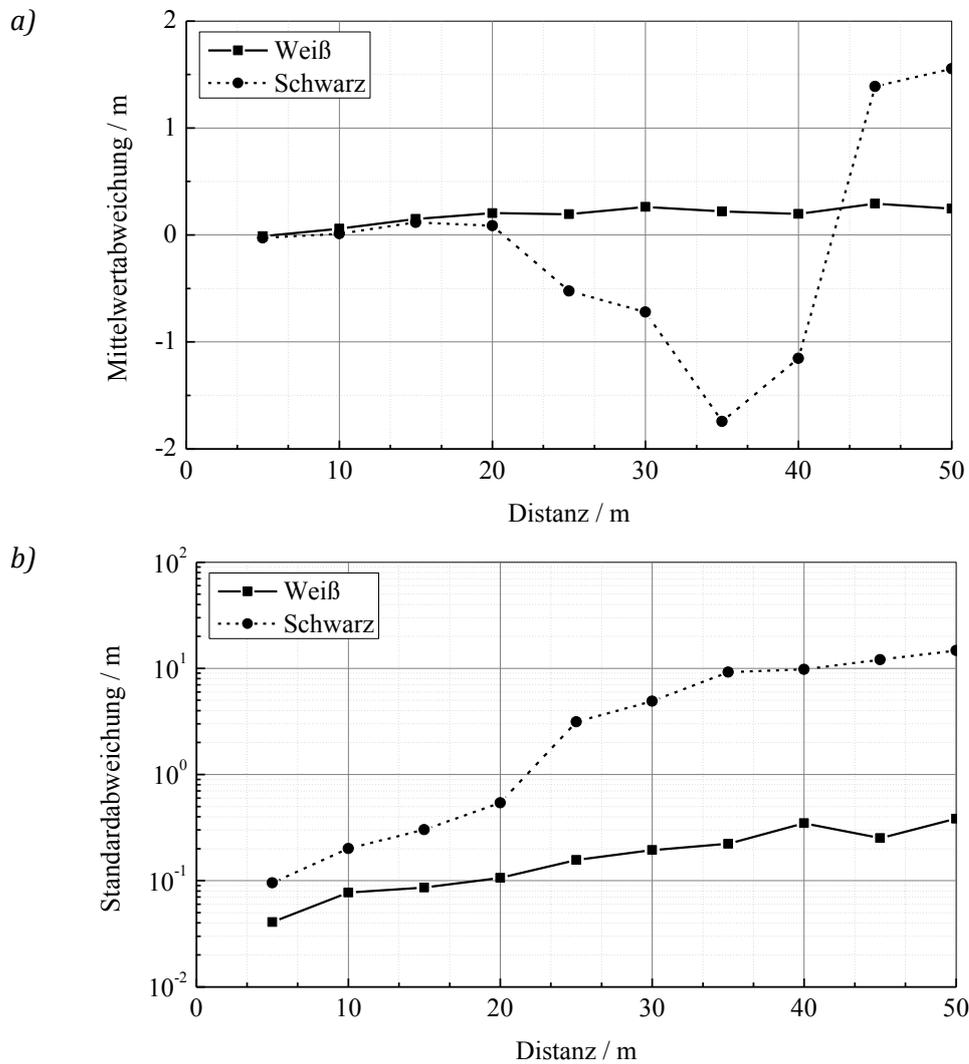


Abbildung 7.40: Mittelwertabweichung (a) und Standardabweichung (b) der Distanzmessung ohne Hintergrundlicht

dieser Distanz die Position des Laserpulses im Histogramm nicht immer gefunden wird. Dementsprechend entstehen einzelne Messwerte, welche sehr stark von Sollwert abweichen und den Mittelwert verfälschen. Dies spiegelt sich in der in Abbildung 7.40 (b) gezeigten Standardabweichung wieder: Ab 25 m steigt diese für das schwarze Objekt stark an. Für das weiße Objekt ergibt sich eine stetige Zunahme der mit der Entfernung, da der reflektierte Laserpuls eine geringere Intensität und damit eine stärkere Streuung im Histogramm aufweist.

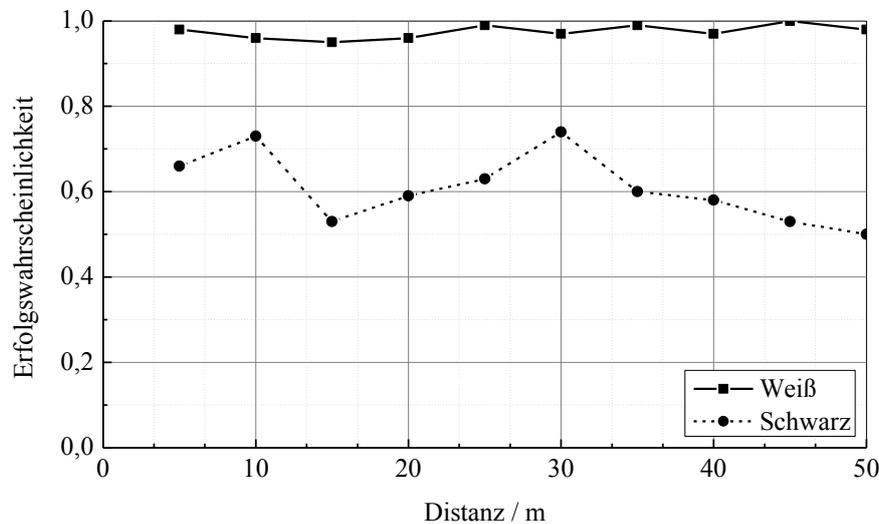


Abbildung 7.41: Erfolgswahrscheinlichkeit der Messung ohne Hintergrundlicht

Dies führt wiederum zu einer höheren Varianz der gemessenen die Distanz, da die Position des Laserpulses weniger präzise bestimmt werden kann.

Da einzelne Distanzwerte hoher Abweichung den Mittelwert und die Standardabweichung stark beeinflussen, wird zusätzlich die Erfolgswahrscheinlichkeit bestimmt. Für das weiße Zielobjekt liegt das in Abbildung 7.41 dargestellte Resultat über dem gesamten Distanzbereich zwischen 96 % und 100 %, während für das schwarze Objekt die Wahrscheinlichkeit zwischen 74 % und 50 % liegt und mit zunehmender Distanz abfällt. Werden nur starke Ausreißer mit einer Abweichung über 10 % vom Sollwert betrachtet, erhält man für das schwarze Objekt bis zu einer Distanz von 20 m eine Erfolgswahrscheinlichkeit von 100 %. Ab 25 m fällt dieser Wert von 97 % auf 61 % ab. Dies zeigt, dass die hohe Abweichung des Mittelwertes sowie der Standardabweichung bei 25 m von lediglich drei fehlerhaften Messwerten verursacht wird.

Die Messung zeigt, dass für das Objekt mit hoher Reflektanz Reichweiten von über 50 m erreicht werden können, wenn kein Hintergrundlicht vorhanden ist. Objekte mit geringer Reflektanz erreichen dagegen deutlich geringe Reichweiten. Durch die Verwendung einer größeren Apertur könnte die Reichweite für alle Reflexionsfaktoren gesteigert werden.



Abbildung 7.42: Messaufbau zur Verifikation der Hintergrundlichtunterdrückung

7.5.4.3 Hintergrundlichtunterdrückung

Zur Überprüfung der Wirksamkeit der Photonenkorrelation zur Unterdrückung von hohem Hintergrundlicht wird die Erfolgswahrscheinlichkeit der Messung eines Zielobjektes in zunehmender Distanz bestimmt, wobei jeder Distanzwert mit allen verfügbaren Koinzidenzeinstellungen gemessen wird. Unter Berücksichtigung der variablen Koinzidenztiefe, Koinzidenzzeit und der Anzahl aktiver SPADs, ergeben sich 28 Kombinationen. Die maximale Abweichung für die Bestimmung der Erfolgswahrscheinlichkeit wird zu 10 % festgelegt. Als Zielobjekte dienen erneut die mit Papier beklebten Holzplatten, wobei nun zusätzlich die Farbe Grau betrachtet wird. Eine Bestimmung der relativen Reflektanz der Zielobjekte über die detektierte Lichtintensität bei Bestrahlung mit einer konstanten Lichtquelle bei 880 nm liefert Werte von 100 %, 80 % und 12,5 %. Die drei Zielobjekte decken damit einen Dynamikbereich von 18 dB ab. Abbildung 7.42 zeigt den Messaufbau mit der Kamera sowie dem schwarzen Zielobjekt in 10 m Entfernung. Die Messung wurde an einem sonnigen Tag im Februar durchgeführt, an welchem eine senkrecht auf das Zielobjekt einfallende Sonnenlichtintensität von 65 klx gemessen wurde.

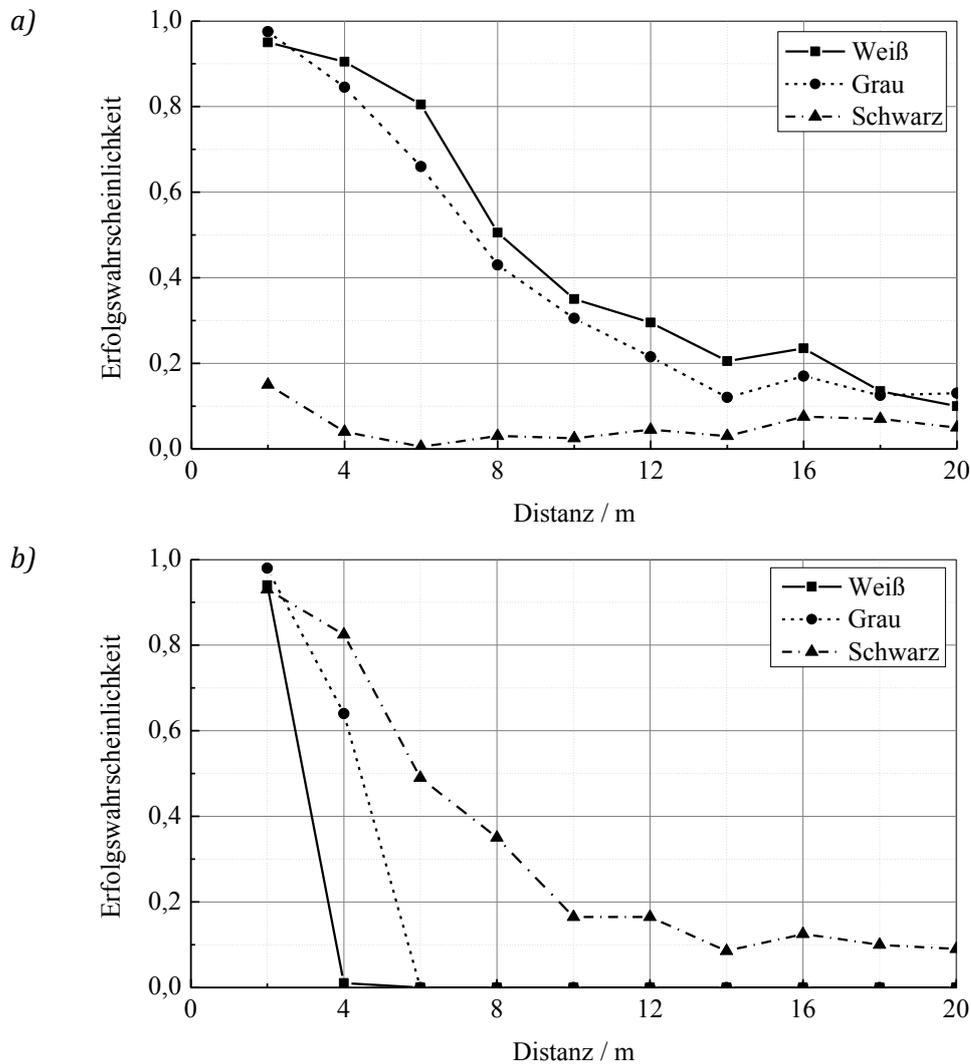


Abbildung 7.43: Erfolgswahrscheinlichkeit der Distanzmessung bei hohem Hintergrundlicht mit hoher (a) und geringer (b) Lichtunterdrückung

Die Ergebnisse der Messung bei hoher Hintergrundlichtintensität sind in Abbildung 7.43 dargestellt. Dabei wurden für jeden Distanzwert 400 Laserpulse akkumuliert, was zusammen mit der Pulswiederholungsrate von 10 kHz einer Bildrate von 25 Hz entspricht. Abbildung 7.43 (a) zeigt die Resultate bei hoher Unterdrückung des Hintergrundlichtes mit 4 aktiven SPADs, einer Koinzidenzzeit von 16 ns und einer Koinzidenztiefe von 4. Hierbei zeigt sich, dass bei Zielobjekten hoher Reflektanz eine deutlich bessere Leistung als bei Objekten geringer Reflektanz erzielt wird. Bei einer geringen Reflektanz ist die Intensität

des reflektierten Lasersignals gering, sodass durch die hohe Koinzidenztiefe das Signal zu stark gedämpft wird und im Histogramm nicht mehr detektiert werden kann. Bei hoher Reflektanz hingegen ist das empfangene Lasersignal ausreichend stark, um nach der Unterdrückung des Hintergrundlichtes noch detektiert werden zu können. Abbildung 7.43 (b) zeigt den Fall geringer Hintergrundlichtunterdrückung mit 4 aktiven SPADs, einer Koinzidenzzeit von 8 ns und einer Koinzidenztiefe von 2. Hierbei zeigt das Objekt mit geringerer Reflektanz ein besseres Ergebnis, da bei Objekten hoher Reflektanz das Hintergrundlicht nicht ausreichend reduziert wird und ein Empfang des reflektierten Laserpulses erschwert wird, da mit hoher Wahrscheinlichkeit das erste Ereignis aus dem Hintergrundlicht resultiert. Die weißen und grauen Objekte zeigen in beiden Graphen nur geringe Abweichungen, da deren Reflektanz sich nur um 2 dB unterscheidet. Zum schwarzen Objekt beträgt der Unterschied 16 dB, weshalb die Änderung hier deutlicher ausfällt.

Zur Darstellung des Effekts der Photonenkorrelation auf das Rohdatenhistogramm im Messbetrieb zeigt Abbildung 7.44 die Histogramme aus der Messung des weißen Objektes in 6 m Entfernung für 80.000 Zeitmessungen. Wie zuvor sind eine geringe Unterdrückung mit 4 SPADs, einer Koinzidenzzeit von 8 ns und einer Koinzidenztiefe von 2 (a) sowie eine höhere Unterdrückung mit 4 SPADs, einer Koinzidenzzeit von 16 ns und einer Koinzidenztiefe von 4 (b) eingestellt. Zusätzlich zeigt (c) das Histogramm bei einer noch höheren Unterdrückung mit 3 SPADs, einer Koinzidenzzeit von 1,5 ns und einer Koinzidenztiefe von 3. Der Vergleich zeigt deutlich, dass bei einer zu geringen Unterdrückung die Ereignisrate des Hintergrundlichtes zu hoch ist und daher der Laserpuls bei Klasse 165 kaum zu erkennen ist. Würden, wie in der tatsächlichen Messung, lediglich 400 Zeitmessungen akkumuliert, würde das Lasersignal im Rauschen der Hintergrundereignisse verschwinden. Bei einer stärkeren Unterdrückung wird der Hintergrund soweit abgesenkt, dass die SPADs bei der Ankunft des Pulses noch ausreichend häufig empfangsbereit sind, um das Signal zu detektieren. Folglich ist der Laserpuls im Histogramm deutlich sichtbar. Darüber hinaus zeigt sich zwischen Histogramm (a) und (b) eine Zunahme des SBR: Interpoliert man den exponentiellen Verlauf in Abbildung 7.44 (a) ergibt sich ein Zählwert des Hintergrundes von ca. 150 sowie des Laserpulses von ca. 50 und damit

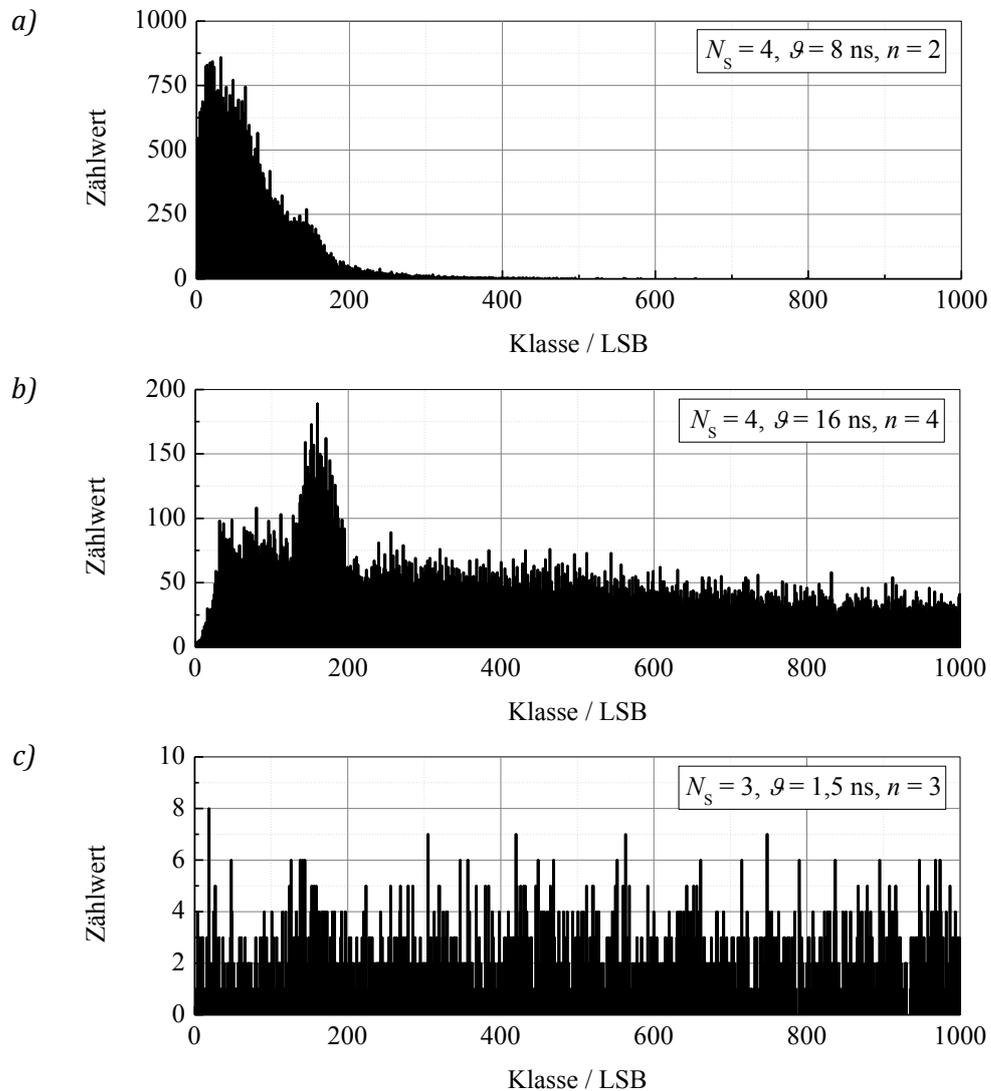


Abbildung 7.44: Rohdatenhistogramme der Distanzmessung bei hohem Hintergrundlicht mit geringer (a), optimaler (b) und hoher (c) Lichtunterdrückung

ein SBR von 0,33. Bei stärkerer Unterdrückung liegt der Hintergrund bei ca. 70 und der Pulse bei ca. 80, sodass sich ein SBR von 1,14 ergibt. Bei einer noch stärkeren Unterdrückung wird auch der reflektierte Laserpuls so stark gedämpft, dass dieser im Histogramm nicht mehr lokalisiert werden kann. Die Histogramme zeigen, dass es ja nach vorhandener Hintergrundlichtintensität eine optimale Einstellung der Photonenkoinzidenz gibt.

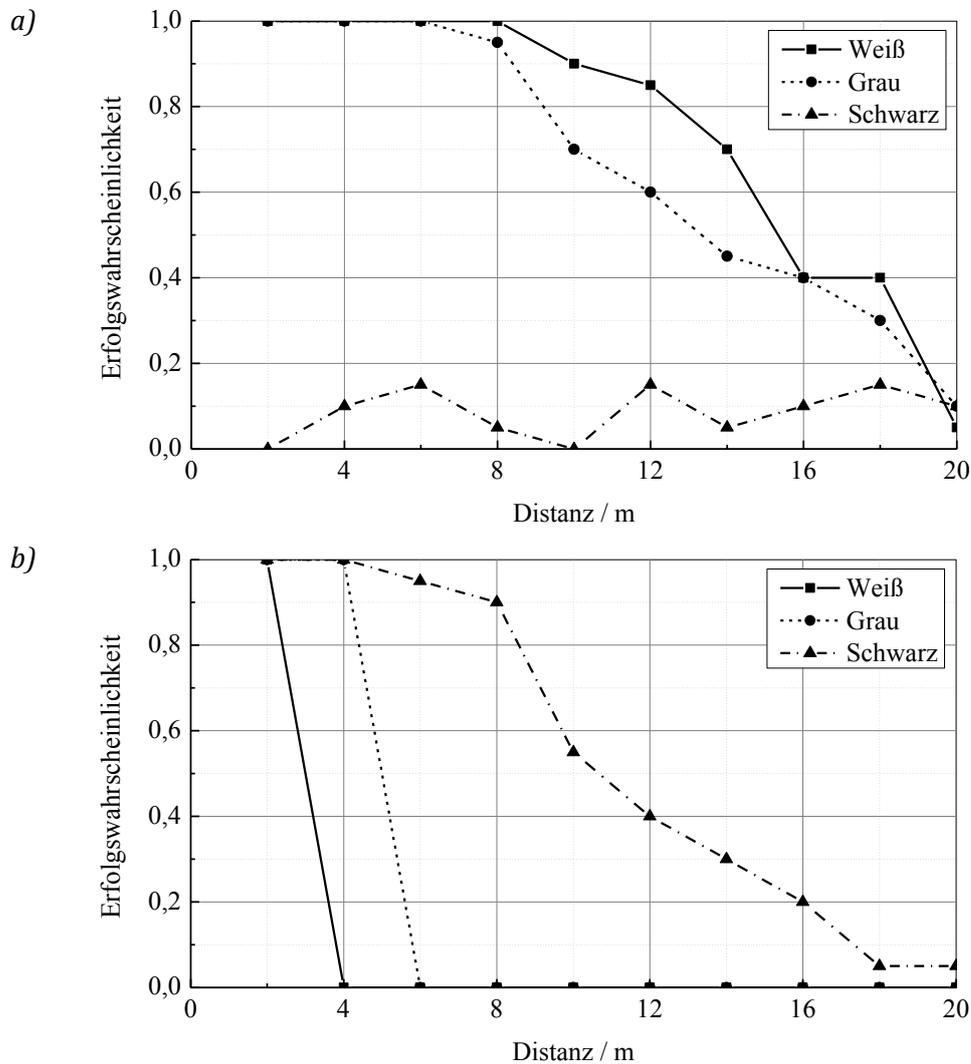


Abbildung 7.45: Erfolgswahrscheinlichkeit mit hoher (a) und geringer (b) Lichtunterdrückung bei der Verwendung von 4000 Laserpulsen

Generell bestätigen die Messungen das Konzept der Detektion zeitlich korrelierter Photonen zur Reduktion des Hintergrundeinflusses auf die Distanzmessung. Zudem zeigt sich, dass die in Abschnitt 6.1.3 beschriebene adaptive Koinzidenz, d. h. die Anpassung der Koinzidenzparameter an die Intensität des Hintergrundlichtes, es ermöglicht, eine näherungsweise konstante Messleistung bei unterschiedlichen Objekten zu erzielen. Die in den Messungen erzielte Systemreichweite liegt jedoch unter den Erwartungen. Eine Optimierungsmöglichkeit stellt dabei der TDC des Sensors da: Durch seinen hohen Jitter von 3,758 LSB

wird der detektierte Laserpuls im Histogramm verbreitert und in seiner Höhe reduziert. Dies sorgt dafür, dass bei höheren Distanzen oder bei hohem Hintergrundlicht der Puls im Histogramm nicht mehr lokalisiert werden kann, da die geringe Höhe der Klassen im Rauschen des Hintergrundes untergeht.

Da für eine Optimierung des TDCs eine Überarbeitung des Sensors erforderlich ist, wird zur Steigerung der Reichweite die Anzahl der akkumulierten Laserpulse erhöht. Entsprechend den Betrachtungen in Abschnitt 4.3.3 verbessert eine Erhöhung der Pulszahl das SNR der direkten Laufzeitmessung nach (4.78). Abbildung 7.45 zeigt die Resultate bei einer Erhöhung von 400 auf 4000 Laserpulse, was bei der verwendeten Laserrate einer Bildrate von 2,5 Hz entspricht. Die Koinzidenzeinstellungen der beiden Graphen sind dabei unverändert. Da gegenwärtig viele Hersteller von Laserkomponenten an Dioden und Treibern mit höheren Raten arbeiten, kann davon ausgegangen werden, dass in naher Zukunft gepulste Laser mit Wiederholungsraten von 100 kHz und mehr zur Verfügung stehen und somit auch bei der Akkumulation höherer Pulszahlen Bildraten von 25 Hz und mehr erreicht werden können. Durch die Erhöhung der Pulse zeigt sich eine deutliche Verbesserung der Reichweite, legt man eine Erfolgswahrscheinlichkeit von 90 % für die Reichweite zu Grunde, erhöht sich die Reichweite beim weißen Zielobjekt von 4 m auf 10 m, beim grauen Objekt von 3 m auf 8,5 m und beim schwarzen Objekt von 2,5 m auf 8 m. Die Steigerung beträgt bei allen Objekten somit mindestens 150 %.

7.5.5 Zusammenfassung

Tabelle 7.3 zeigt eine Zusammenfassung der Ergebnisse der Sensorcharakterisierung. Die gemessene DCR und PDE der SPADs entsprechen mit 40,56 Hz und 1,89 % den in früheren Messungen ermittelten Werten [9]. Mit 3,39 LSB zeigt die INL des TDCs aufgrund der Frequenzänderung der PLL über dem Messfenster einen hohen Wert. Durch entsprechende Anpassungen im Layout kann dieses Problem aber in zukünftigen Sensoren behoben werden. Die DNL zeigt mit 0,084 LSB dagegen ein gutes Ergebnis [82]. Die Unsicherheit des TDCs ist mit 3,758 LSB hoch. Zur Reduktion muss die Frequenz der PLL stabilisiert werden, was durch eine Anpassung des Schleifenfilters oder des Layouts zur Reduktion

Tabelle 7.3: Ergebnisse der Sensor- und Systemcharakterisierung

Parameter	Wert	Einheit	Bemerkung
DCR	40,56	Hz	Median
Totzeit	16,79	ns	$U_{DT} = 3,3 V$
PDE	1,89	%	$\lambda = 905 \text{ nm}, U_{ex} = 4,7 V$
INL	3,39	LSB	Aus Histogramm
DNL	0,084	LSB	Aus Histogramm
Präzision	3,758	LSB	Mittelwert
Reichweite	50	m	300 lx, 400 Pulse, weiß
	10	m	65 klx, 4000 Pulse, weiß

von Störungen erreicht werden kann. Eine weitere Option ist es den Haupttakt direkt von einer externen Taktquelle einzuspeisen, was die Notwendigkeit einer internen Taktgeneration eliminiert. Eine weitere Optimierung betrifft die Laufzeitinterpolation, durch eine Anpassung der Verzögerung im Signalpfad, welcher für das Einfrieren der Signalzustände der Verzögerungskette zuständig ist, lässt sich die Unsicherheit des TDCs weiter reduzieren. Die Bestimmung des Einflusses der Koinzidenzparameter auf die resultierende Ereignisrate bestätigt das auf Basis theoretischer Berechnungen entworfene Modell. Die Bestimmung der Reichweite des Kamerasystems zeigt, dass ohne oder mit nur sehr geringem Hintergrundlicht Reichweiten von über 50 m erzielt werden, während bei hoher Hintergrundlichtintensität lediglich 10 m erreicht werden. Eine Ursache für die geringe Distanz ist die hohe Unsicherheit des TDCs, welche in zukünftigen Sensoren gelöst werden muss. Durch die Erhöhung der Anzahl akkumulierter Pulse im Histogramm konnte die erzielte Reichweite deutlich gesteigert werden. Für zukünftige Systeme sind folglich höhere Laserpulsraten wünschenswert. Die Messung der Erfolgswahrscheinlichkeit bei verschiedenen Reflexionskoeffizienten des Zielobjektes zeigt, dass eine Anpassung der Koinzidenzparameter eine näherungsweise gleichbleibende Reichweite ermöglicht. Die adaptive Anpassung der Koinzidenzparameter auf Basis einer Messung der Ereignisrate des Hintergrundlichtes soll daher im Kamerasystem implementiert werden.

8 Zusammenfassung und Ausblick

Das Ziel der Arbeit war die Entwicklung eines SPAD-basierten Sensors für die laufzeitbasierte Distanzmessung bei hoher Hintergrundlichtintensität. Die Entwicklung zielte dabei auf Anwendungen im Automobilbereich ab, bei welchen besondere Anforderungen an die Hintergrundlichtresistenz gestellt werden. Zur Auswahl des geeignetsten Messkonzeptes wurden zunächst die in der Literatur verfügbaren Verfahren hinsichtlich der Messvarianz unter Berücksichtigung der SPAD-typischen Totzeit theoretisch untersucht. Auf Basis dieser wurden verschiedene Modifikationen zur Steigerung der Hintergrundlichtresistenz entworfen sowie theoretisch und simulativ auf die zu erwartende Verbesserung untersucht. Die Gesamtheit der Untersuchungen ließ den Schluss zu, dass das direkte Messverfahren in Verbindung mit der Detektion zeitlich korrelierter Einzelphotonen mit variablen Korrelationsparametern eine konstant hohe Reichweite bei hohem Hintergrund sowie wechselnder Reflexionscharakteristik ermöglicht. Dementsprechend wurde ein Doppelzeilensensor mit 2×192 Pixeln in $0,35\text{-}\mu\text{m}$ -CMOS-Technologie zur Verifikation der theoretischen Berechnungen entworfen und gefertigt. Die Charakterisierung des Sensors zeigte, dass die Detektion korrelierter Photonen zur Reduktion der Ereignisraten und damit zur Unterdrückung des Hintergrundlichtes verwendet werden kann. Bei Messungen unter realitätsnahen Bedingungen im Freien wurde eine Reichweite zwischen 8 m und 10 m bei einer Hintergrundlichtintensität von 65 klx für unterschiedliche Reflexionskoeffizienten des Zielobjektes erreicht.

Im Rahmen der Arbeit wurde ein Sensor mit zwei Zeilen entworfen. Eine logische Weiterentwicklung stellt der Übergang zu Multizeilensensoren oder zweidimensionalen Pixelarrays dar. Diese bieten einen höheren Informationsgehalt und eröffnen damit eine Vielzahl weiterer Anwendungsfelder. Besonders in der digitalen Bildverarbeitung erlaubt ein Pixelarray eine vielfältige Weiterverarbeitung der Daten. So können Objekte klassifiziert, im Raum verfolgt sowie deren Bewegungsrichtung und Geschwindigkeit ermittelt werden. Dies bietet im Automobilbereich eine Verbesserung der Hinderniserkennung, was die Sicherheit und Zuverlässigkeit der Systeme erhöht. Die Anordnung der Pixel in einem zweidimensionalen Array hat, wie in der Einleitung erwähnt, einen negativen Einfluss auf den Füllfaktor des Sensors. Eine Lösung dieses Problems stellt die 3D-Integration dar. Dabei befinden sich die SPADs und die zugehörige Elektronik auf separaten Wafern, welche nach der Fertigung zusammengefügt werden. Am Fraunhofer IMS wurde bereits ein derartiges Verfahren entwickelt, was für zukünftige Generationen von SPAD-basierten Laufzeitsensoren zur Verfügung steht. Neben dem Sensor selbst bietet auch die Datenauswertung und Datenverarbeitung weiteres Potential. Bei der in der Arbeit verwendeten Datenverarbeitung werden die Daten einer Messung stets isoliert betrachtet. Werden die Daten vorhergehender Bilder mit in die Auswertung einbezogen, könnte dies die Güte steigern. So könnte eine Reduktion des Hintergrundlichtes durch zeitliche Anpassung des Messfensters auf Basis der Position eines Objektes im vorhergehenden Bild erfolgen. Eine weitere Möglichkeit liegt in der Ausnutzung der Quantenmechanik. So könnte es durch das simultane Aussenden und Beobachten eines verschränkten Photonenpaares möglich sein, den Einfluss fremder Photonen aus dem Hintergrundlicht gänzlich zu eliminieren.

Literaturverzeichnis

- [1] S. Bellisai, D. Bronzi, F. A. Villa, S. Tisa, A. Tosi, und F. Zappa, „Single-photon pulsed-light indirect time-of-flight 3D ranging“, *Opt. Express*, Bd. 21, Nr. 4, S. 5086, Feb. 2013.
- [2] W. S. Boyle und G. E. Smith, „Charge Coupled Semiconductor Devices“, *Bell Syst. Tech. J.*, Bd. 49, Nr. 4, S. 587–593, Apr. 1970.
- [3] S. Mendis, S. E. Kemeny, und E. R. Fossum, „CMOS active pixel image sensor“, *IEEE Trans. Electron Devices*, Bd. 41, Nr. 3, S. 452–453, März 1994.
- [4] F. Remondino und D. Stoppa, *TOF Range-Imaging Cameras*. New York: Springer, 2013.
- [5] S. Bellisai, F. Guerrieri, S. Tisa, und F. Zappa, „3D ranging with a single-photon imaging array“, 2011, Bd. 7875, S. 78750M-78750M-6.
- [6] F. Zappa und A. Tosi, „MiSPIA: microelectronic single-photon 3D imaging arrays for low-light high-speed safety and security applications“, 2013, Bd. 8727, S. 87270L-87270L-11.
- [7] C. Scarcella, A. Tosi, F. Villa, S. Tisa, und F. Zappa, „Low-noise low-jitter 32-pixels CMOS single-photon avalanche diodes array for single-photon counting from 300 nm to 900 nm“, *Rev. Sci. Instrum.*, Bd. 84, Nr. 12, S. 123112, Dez. 2013.
- [8] A. Vil, A. Arbat, E. Vilella, und A. Dieguez, „Geiger-Mode Avalanche Photodiodes in Standard CMOS Technologies“, in *Photodetectors*, S. Gateva, Hrsg. InTech, 2012.
- [9] D. Bronzi *u. a.*, „Low-noise and large-area CMOS SPADs with timing response free from slow tails“, in *Solid-State Device Research Conference (ESSDERC), 2012 Proceedings of the European*, 2012, S. 230–233.
- [10] F. Villa *u. a.*, „High linearity SPAD and TDC array for TCSPC and 3D ranging applications“, 2015, Bd. 9370, S. 93701U-93701U-7.
- [11] C. Niclass, M. Soga, H. Matsubara, M. Ogawa, und M. Kagami, „A 0.18- μm CMOS SoC for a 100-m-Range 10-Frame/s 200 x 96-Pixel Time-of-Flight Depth Sensor“, *IEEE J. Solid-State Circuits*, Bd. 49, Nr. 1, S. 315–330, Jan. 2014.
- [12] H. Akita, I. Takai, K. Azuma, T. Hata, und N. Ozaki, „An imager using 2-D single-photon avalanche diode array in 0.18- μm CMOS for automotive LIDAR application“, in *2017 Symposium on VLSI Circuits*, 2017, S. C290–C291.
- [13] F. Villa *u. a.*, „CMOS Imager With 1024 SPADs and TDCs for Single-Photon Timing and 3-D Time-of-Flight“, *IEEE J. Sel. Top. Quantum Electron.*, Bd. 20, Nr. 6, S. 364–373, Nov. 2014.
- [14] D. Bronzi *u. a.*, „100 000 Frames/s 64 x 32 Single-Photon Detector Array for 2-D Imaging and 3-D Ranging“, *IEEE J. Sel. Top. Quantum Electron.*, Bd. 20, Nr. 6, S. 354–363, Nov. 2014.
- [15] F. Guerrieri, S. Tisa, A. Tosi, und F. Zappa, „Two-Dimensional SPAD Imaging Camera for Photon Counting“, *IEEE Photonics J.*, Bd. 2, Nr. 5, S. 759–774, Okt. 2010.

- [16] D. Portaluppi, E. Conca, und F. Villa, „32 x 32 CMOS SPAD Imager for Gated Imaging, Photon Timing, and Photon Coincidence“, *IEEE J. Sel. Top. Quantum Electron.*, Bd. 24, Nr. 2, S. 1–6, März 2018.
- [17] M. Perenzoni, D. Perenzoni, und D. Stoppa, „A 64 × 64-Pixels Digital Silicon Photomultiplier Direct TOF Sensor With 100-MPhotons/s/pixel Background Rejection and Imaging/Altimeter Mode With 0.14% Precision Up To 6 km for Spacecraft Navigation and Landing“, *IEEE J. Solid-State Circuits*, Bd. 52, Nr. 1, S. 151–160, Jan. 2017.
- [18] I. Vornicu, R. Carmona-Galan, und A. Rodriguez-Vazquez, „A CMOS 0.18 μm 64 x 64 single photon image sensor with in-pixel 11b time-to-digital converter“, in *Semiconductor Conference (CAS), 2014 International*, 2014, S. 131–134.
- [19] I. Vornicu, R. Carmona-Galán, und A. Rodríguez-Vázquez, „Photon counting and direct ToF camera prototype based on CMOS SPADs“, in *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2017, S. 1–4.
- [20] E. Ackerman, „Lidar that will make self-driving cars affordable [News]“, *IEEE Spectr.*, Bd. 53, Nr. 10, S. 14–14, Okt. 2016.
- [21] A. Süß, V. Rochus, M. Rosmeulen, und X. Rottenberg, „Benchmarking time-of-flight based depth measurement techniques“, 2016, Bd. 9751, S. 975118-975118–19.
- [22] S. Lindner, C. Zhang, I. M. Antolovic, J. Mata Pavia, M. Wolf, und E. Charbon, „Column-Parallel Dynamic TDC Reallocation in SPAD Sensor Module Fabricated in 180nm CMOS for Near Infrared Optical Tomography“, *Proc. 2017 Int. Image Sens. Workshop*, 2017.
- [23] Y. Zou, D. Bronzi, F. Villa, und S. Weyers, „Backside illuminated wafer-to-wafer bonding single photon avalanche diode array“, in *Microelectronics and Electronics (PRIME), 2014 10th Conference on Ph.D. Research in*, 2014, S. 1–4.
- [24] G. Intermite *u. a.*, „Enhancing the fill-factor of CMOS SPAD arrays using microlens integration“, 2015, Bd. 9504, S. 95040J-95040J–12.
- [25] J. Illade-Quinteiro, V. M. Brea, P. López, D. Cabello, und G. Doménech-Asensi, „Distance Measurement Error in Time-of-Flight Sensors Due to Shot Noise“, *Sensors*, Bd. 15, Nr. 3, S. 4624–4642, Feb. 2015.
- [26] G. Lutz, *Semiconductor radiation detectors: device physics*, 1st ed. New York: Springer, 2007.
- [27] B. El-Kareh, *Silicon devices and process integration: deep submicron and nano-scale technologies*. New York: Springer, 2009.
- [28] D. Stoppa, L. Pancheri, M. Scandiuozzo, L. Gonzo, G.-F. Dalla Betta, und A. Simoni, „A CMOS 3-D Imager Based on Single Photon Avalanche Diode“, *IEEE Trans. Circuits Syst. Regul. Pap.*, Bd. 54, Nr. 1, S. 4–12, Jan. 2007.
- [29] D. Bronzi *u. a.*, „Large-area CMOS SPADs with very low dark counting rate“, 2013, Bd. 8631, S. 86311B-86311B–8.

-
- [30] H. Zimmermann, *Integrated silicon optoelectronics*, 2nd ed. Berlin ; London: Springer, 2010.
- [31] A. Spinelli und A. L. Lacaita, „Physics and numerical simulation of single photon avalanche diodes“, *IEEE Trans. Electron Devices*, Bd. 44, Nr. 11, S. 1931–1943, Nov. 1997.
- [32] P. Seitz und A. J. P. Theuwissen, Hrsg., *Single-photon imaging*. Heidelberg ; New York: Springer, 2011.
- [33] C. F. Klingshirn, *Semiconductor optics*, 3rd ed. Berlin ; New York: Springer, 2007.
- [34] F. L. Pedrotti, Hrsg., *Optik für Ingenieure: Grundlagen ; mit 28 Tabellen*, 3., bearb. und aktualisierte Aufl. Berlin: Springer, 2005.
- [35] W. G. Oldham, R. R. Samuelson, und P. Antognetti, „Triggering phenomena in avalanche diodes“, *IEEE Trans. Electron Devices*, Bd. 19, Nr. 9, S. 1056–1060, Sep. 1972.
- [36] M. W. Fishburn, *Fundamentals of CMOS Single-Photon Avalanche Diodes*. fishburn, 2012.
- [37] D. P. Palubiak und M. J. Deen, „CMOS SPADs: Design Issues and Research Challenges for Detectors, Circuits, and Arrays“, *IEEE J. Sel. Top. Quantum Electron.*, Bd. 20, Nr. 6, S. 409–426, Nov. 2014.
- [38] D. Bronzi, „Front-side and Back-side Illuminated SPAD Arrays for 2D Imaging and 3D Ranging“, Politecnico di Milano, 2014.
- [39] J.-P. Colinge und C. A. Colinge, *Physics of semiconductor devices*. Boston: Kluwer Academic Publishers, 2002.
- [40] M. Razeghi, *Fundamentals of solid state engineering*, 3rd edition. New York, NY: Springer, 2009.
- [41] S. M. Sze und K. K. Ng, *Physics of semiconductor devices*, 3rd ed. Hoboken, N.J: Wiley-Interscience, 2007.
- [42] G. A. M. Hurkx, H. C. de Graaff, W. J. Kloosterman, und M. P. G. Knuvers, „A new analytical diode model including tunneling and avalanche breakdown“, *IEEE Trans. Electron Devices*, Bd. 39, Nr. 9, S. 2090–2098, Sep. 1992.
- [43] S. Cova, M. Ghioni, A. Lacaita, C. Samori, und F. Zappa, „Avalanche photodiodes and quenching circuits for single-photon detection“, *Appl. Opt.*, Bd. 35, Nr. 12, S. 1956, Apr. 1996.
- [44] A. Gallivanoni, I. Rech, und M. Ghioni, „Progress in Quenching Circuits for Single Photon Avalanche Diodes“, *IEEE Trans. Nucl. Sci.*, Bd. 57, Nr. 6, S. 3815–3826, Dez. 2010.
- [45] A. L. Lacaita, F. Zappa, S. Bigliardi, und M. Manfredi, „On the bremsstrahlung origin of hot-carrier-induced photons in silicon devices“, *IEEE Trans. Electron Devices*, Bd. 40, Nr. 3, S. 577–582, März 1993.
- [46] M. Assanelli, A. Ingargiola, I. Rech, A. Gulinatti, und M. Ghioni, „Photon-Timing Jitter Dependence on Injection Position in Single-Photon Avalanche Diodes“, *IEEE J. Quantum Electron.*, Bd. 47, Nr. 2, S. 151–159, Feb. 2011.

- [47] D. Bronzi, S. Tisa, F. Villa, S. Bellisai, A. Tosi, und F. Zappa, „Fast Sensing and Quenching of CMOS SPADs for Minimal Afterpulsing Effects“, *IEEE Photonics Technol. Lett.*, Bd. 25, Nr. 8, S. 776–779, Apr. 2013.
- [48] W. Freude, „185 MHz Count Rate, 139 dB Dynamic Range Single-Photon Avalanche Diode with Active Quenching Circuit in 130nm CMOS Technology“, gehalten auf der International Image Sensor Workshop, 2011.
- [49] A. M. Fox, *Quantum Optics: An Introduction*. Oxford New York: Oxford University Press, 2007.
- [50] F. Blais, „Review of 20 years of range sensor development“, *J. Electron. Imaging*, Bd. 13, Nr. 1, S. 231–243, 2004.
- [51] D. Bronzi u. a., „3D sensor for indirect ranging with pulsed laser source“, 2012, Bd. 8541, S. 85410T-85410T-5.
- [52] M. Löffler-Mang, *Optische Sensorik: Lasertechnik, Experimente, Light Barriers*, 1. Aufl. Wiesbaden: Vieweg + Teubner, 2012.
- [53] P. J. Besl, „Active, optical range imaging sensors“, *Mach. Vis. Appl.*, Bd. 1, Nr. 2, S. 127–152, Juni 1988.
- [54] S. Bellisai u. a., „Low-power 20-meter 3D ranging SPAD camera based on continuous-wave indirect time-of-flight“, 2012, Bd. 8375, S. 83750E-83750E-7.
- [55] P. R. Bevington und D. K. Robinson, *Data reduction and error analysis for the physical sciences*, 3rd ed. Boston: McGraw-Hill, 2003.
- [56] W. Feller, *Vol 1 - An introduction to probability theory and its applications*, 3. ed., rev. Print., [Nachdr.]. S.l.: Wiley, 1968.
- [57] D. F. Yu und J. A. Fessler, „Mean and variance of single photon counting with deadtime“, *Phys. Med. Biol.*, Bd. 45, Nr. 7, S. 2043–2056, Juli 2000.
- [58] M. Beer, O. M. Schrey, B. J. Hosticka, und R. Kokozinski, „Expected Value and Variance of the Indirect Time-of-Flight Measurement With Dead Time Afflicted Single-Photon Avalanche Diodes“, *IEEE Trans. Circuits Syst. Regul. Pap.*, Bd. 65, Nr. 3, S. 970–981, März 2018.
- [59] I. Vornicu, R. Carmona-Galan, und A. Rodriguez-Vazquez, „On the calibration of a SPAD-based 3D imager with in-pixel TDC using a time-gated technique“, in *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2015, S. 1102–1105.
- [60] T. Murphy, „Statistics of Photon Arrival Time“. 02-Feb-2001.
- [61] L. Papula, *Mathematik für Ingenieure und Naturwissenschaftler 3: Vektoranalysis, Wahrscheinlichkeitsrechnung, mathematische Statistik, Fehler- und Ausgleichsrechnung*, 6., überarb. und erw. Aufl. Wiesbaden: Vieweg + Teubner, 2011.
- [62] European Standard, „EN 60825-1:2007 Safety of Laser Products - Part 1: Equipment classification and requirements“. British Standards, 09-Jan-2007.
- [63] C. J. Chen, *Physics of solar energy*. Hoboken, NJ: Wiley, 2011.

- [64] M. M. Hayat, S. N. Torres, und L. M. Pedrotti, „Theory of photon coincidence statistics in photon-correlated beams“, *Opt. Commun.*, Bd. 169, Nr. 1–6, S. 275–287, Okt. 1999.
- [65] C. Niclass, M. Soga, H. Matsubara, S. Kato, und M. Kagami, „A 100-m Range 10-Frame/s 340 x 96-Pixel Time-of-Flight Depth Sensor in 0.18- μm CMOS“, *IEEE J. Solid-State Circuits*, Bd. 48, Nr. 2, S. 559–572, Feb. 2013.
- [66] I. Vornicu, R. Carmona-Galán, und Á. Rodríguez-Vázquez, „A SPAD-based 3D imager with in-pixel TDC for 145ps-accuracy ToF measurement“, 2015, Bd. 9403, S. 94030I-1-94030I-6.
- [67] T. Hummel, Hrsg., *Fahrerassistenzsysteme: Ermittlung des Sicherheitspotenzials auf Basis des Schadengeschehens der deutschen Versicherer*. Berlin: GDV, 2011.
- [68] A. Shaout, D. Colella, und S. Awad, „Advanced Driver Assistance Systems - Past, present and future“, in *Computer Engineering Conference (ICENCO), 2011 Seventh International*, 2011, S. 72–82.
- [69] J. Levinson u. a., „Towards fully autonomous driving: Systems and algorithms“, in *2011 IEEE Intelligent Vehicles Symposium (IV)*, 2011, S. 163–168.
- [70] S. Fuchs, „Multipath Interference Compensation in Time-of-Flight Camera Images“, in *2010 20th International Conference on Pattern Recognition*, 2010, S. 3583–3586.
- [71] M. Gersbach u. a., „A Time-Resolved, Low-Noise Single-Photon Image Sensor Fabricated in Deep-Submicron CMOS Technology“, *IEEE J. Solid-State Circuits*, Bd. 47, Nr. 6, S. 1394–1407, Juni 2012.
- [72] P. Chen, C. C. Chen, J. C. Zheng, und Y. S. Shen, „A PVT Insensitive Vernier-Based Time-to-Digital Converter With Extended Input Range and High Accuracy“, *IEEE Trans. Nucl. Sci.*, Bd. 54, Nr. 2, S. 294–302, Apr. 2007.
- [73] J. Richardson u. a., „A 32 x 32 50ps resolution 10 bit time to digital converter array in 130nm CMOS for time correlated imaging“, in *IEEE Custom Integrated Circuits Conference, 2009. CICC '09*, 2009, S. 77–80.
- [74] M. Kim u. a., „High-resolution and wide-dynamic range time-to-digital converter with a multi-phase cyclic Vernier delay line“, in *ESSCIRC (ESSCIRC), 2013 Proceedings of the*, 2013, S. 311–314.
- [75] C.-C. Chen, S.-H. Lin, und C.-S. Hwang, „An Area-Efficient CMOS Time-to-Digital Converter Based on a Pulse-Shrinking Scheme“, *IEEE Trans. Circuits Syst. II Express Briefs*, Bd. 61, Nr. 3, S. 163–167, März 2014.
- [76] B. Markovic, S. Tisa, F. A. Villa, A. Tosi, und F. Zappa, „A High-Linearity, 17 ps Precision Time-to-Digital Converter Based on a Single-Stage Vernier Delay Loop Fine Interpolation“, *IEEE Trans. Circuits Syst. Regul. Pap.*, Bd. 60, Nr. 3, S. 557–569, März 2013.
- [77] Xilinx, „Zynq-7000 All Programmable SoC Data Sheet: Overview“. 23-Okt-2017.
- [78] Osram, „Nanostack Pulsed Laser Diode in Plastic Package 75 W Peak Power“. 28-Nov-2014.

- [79] PicoLAS, „PicoLAS LDP-AV 40-70 User Manual“. 23-Okt-2017.
- [80] PicoLAS, „Application Note # 03: LD-Connections“. 05-Juli-2012.
- [81] G. Turin, „An introduction to matched filters“, *IRE Trans. Inf. Theory*, Bd. 6, Nr. 3, S. 311–329, Juni 1960.
- [82] F. Villa u. a., „SPAD Smart Pixel for Time-of-Flight and Time-Correlated Single-Photon Counting Measurements“, *IEEE Photonics J.*, Bd. 4, Nr. 3, S. 795–804, Juni 2012.
- [83] Precaster Enterprises, „Laser Distance Meter CX80 - Operation Manual“. 07-Apr-2016.
- [84] F. Träger, Hrsg., *Springer handbook of lasers and optics*, 2nd ed. Dordrecht ; New York: Springer, 2012.
- [85] T. Gockel, *Kompendium digitale Fotografie: von der Theorie zur erfolgreichen Fotopraxis*. Berlin: Springer, 2012.
- [86] J. M. Lloyd, *Thermal imaging systems*, 3. printing. New York: Plenum Press, 1982.

Veröffentlichungen

Im Rahmen der Arbeit sind nachfolgende Veröffentlichungen entstanden:

- M. Beer, B. J. Hosticka, and R. Kokozinski, "SPAD-based 3D sensors for high ambient illumination," in *2016 12th Conference on Ph.D. Research in Microelectronics and Electronics (PRIME)*, 2016, pp. 1–4.
- M. Beer, O. Schrey, B. J. Hosticka, and R. Kokozinski, "Dead time effects in the indirect time-of-flight measurement with SPADs," in *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2017, pp. 1–4.
- M. Beer, O. M. Schrey, B. J. Hosticka, and R. Kokozinski, "Coincidence in SPAD-based time-of-flight sensors," in *2017 13th Conference on Ph.D. Research in Microelectronics and Electronics (PRIME)*, 2017, pp. 381–384.
- M. Beer, O. M. Schrey, B. J. Hosticka, and R. Kokozinski, "Modelling of SPAD-based time-of-flight measurement techniques," in *2017 European Conference on Circuit Theory and Design (ECCTD)*, 2017, pp. 1–4.
- M. Beer, B. J. Hosticka, O. M. Schrey, W. Brockherde, and R. Kokozinski, "Range accuracy of SPAD-based time-of-flight sensors," in *2017 European Conference on Circuit Theory and Design (ECCTD)*, 2017, pp. 1–4.
- O. M. Schrey, M. Beer, W. Brockherde, and B. J. Hosticka, "SPAD-Based Flash Lidar with High Background Light Suppression," in *Advanced Microsystems for Automotive Applications 2017*, Springer, Cham, 2018, pp. 37–43.
- M. Beer, O. M. Schrey, C. Nitta, W. Brockherde, B. J. Hosticka, and R. Kokozinski, "1×80 pixel SPAD-based flash LIDAR sensor with background rejection based on photon coincidence," in *2017 IEEE SENSORS*, 2017, pp. 1–3.
- M. Beer *et al.*, "SPAD-based flash LiDAR sensor with high ambient light rejection for automotive applications," presented at the Quantum Sensing and Nano Electronics and Photonics XV, 2018, vol. 10540, p. 105402G.
- M. Beer, O. M. Schrey, B. J. Hosticka, and R. Kokozinski, "Expected Value and Variance of the Indirect Time-of-Flight Measurement With Dead Time Afflicted Single-Photon Avalanche Diodes," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 65, no. 3, pp. 970–981, Mar. 2018.

Anhang

Dieser Abschnitt enthält ergänzende Betrachtungen aus dem Haupttext. Dazu gehören eine Analyse des Einflusses einer nichtidealen Signalform in den indirekten Messverfahren sowie eine Betrachtung der Sensoroptik bezüglich der Schärfentiefe und der lateralen Auflösung.

A Nichtideale Signalform im indirekten Verfahren

Die grundlegenden Berechnungsgleichungen der Distanz in den indirekten Verfahren mit PM und CW auf Basis der in mehreren Zeitfenstern gezählten Ereignissen setzen eine ideale Signalform des empfangenen Lasersignals voraus. Da eine unendliche Flankensteilheit bzw. eine ideale Sinusform in der Realität nicht erreichbar sind, wird in den folgenden Betrachtungen der Einfluss einer nichtidealen Signalform auf die Distanzberechnung untersucht. Der Fokus liegt dabei auf dem systematischen Fehler.

A.1 Pulsmodulation

Beim indirekten Verfahren zur Distanzbestimmung auf Basis der Laufzeit mit PM hat neben der Totzeit auch die Pulsform einen Einfluss auf die Messung. Bei einem nichtidealen Signalpuls ergibt sich eine nichtlineare Übertragungskennlinie zwischen Laufzeit und Distanz. Dies soll hier am Beispiel eines trapezförmigen Pulses, d. h. eines Pulses endlicher Flankensteilheit, gezeigt werden.

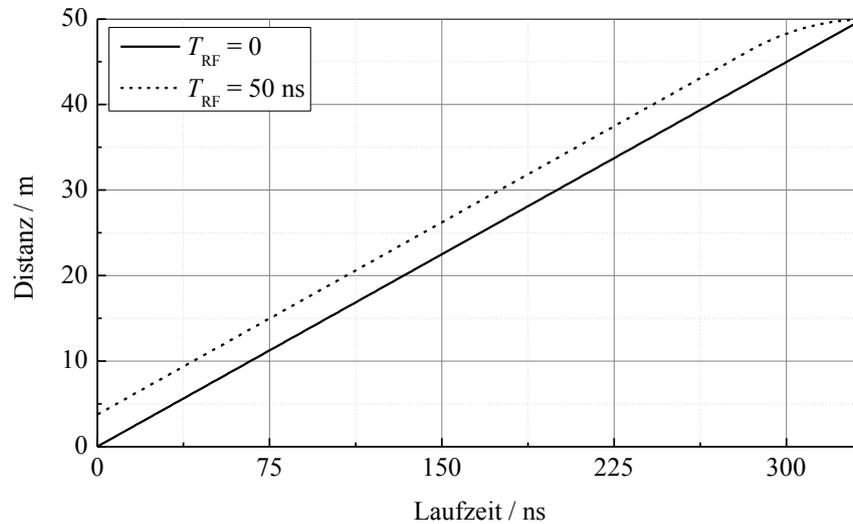


Abbildung A.1: Einfluss endlicher Flankensteilheit im indirekten PM-Verfahren

In Abbildung A.1 sind die aus der Berechnungsgleichung (3.6) resultierenden Übertragungskennlinien bei Annahme eines idealen Rechteckpulses sowie eines Pulses mit endlicher Flankensteilheit dargestellt. Für die Berechnung wurde eine Pulsdauer von 333 ns zu Grunde gelegt. Beide Pulse steigen bei $t = T_{\text{TOF}}$ an und fallen bei $t = T_{\text{TOF}} + T_{\text{P}}$ ab. Die Anstiegs- bzw. Abfallzeit T_{RF} beträgt im Falle des nichtidealen Pulses zur Verdeutlichung des Effektes 50 ns, was von realen Systemen deutlich unterboten wird. Auf das dritte Zeitfenster zur Subtraktion des Hintergrundes wird verzichtet, da dies auch bei nichtidealem Puls durch Verschiebung korrekte Resultate liefern kann. Die Berechnung zeigt, dass sich durch die endliche Flankensteilheit die Kurve verschiebt, da sich bereits bei einer Laufzeit von $T_{\text{TOF}} = 0$ ein Teil des Pulses im zweiten Fenster befindet. Weiterhin sinkt die Steigung für T_{TOF} gegen T_{P} , da durch die endliche Flankensteilheit weniger Photonen in das zweite Fenster übergehen.

Da die Pulsform im Allgemeinen konstant ist, kann dieser Effekt durch entsprechende Kalibration des Messsystems kompensiert werden. Im realen System ist zudem durch die im Vergleich zur Flanke lange Pulsdauer der Einfluss auf die Messvarianz gering. Um den Effekt der abnehmenden Kurvensteilheit zu eliminieren, kann der Puls länger als die maximale Messdistanz gewählt werden, sodass dieser Teil nicht betrachtet werden muss.

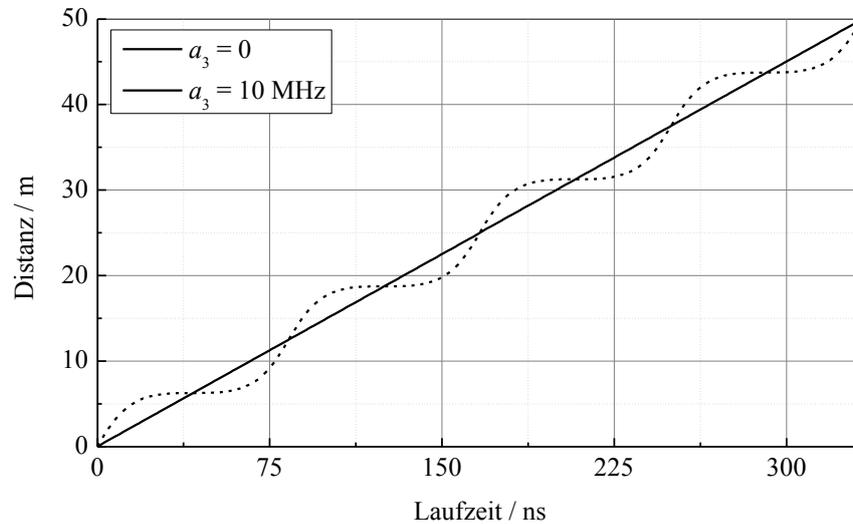


Abbildung A.2: Einfluss von Oberschwingungen im indirekten CW-Verfahren

A.2 Sinusmodulation

Auch beim indirekten Messverfahren mit CW verfälscht eine nichtideale Sinusform das Messergebnis, sofern keine Kalibration durchgeführt wird. Nichtlineare Übertragungsglieder wie die Laserquelle sorgen dafür, dass das Ausgangssignal S_{OUT} zusätzliche Harmonische enthält. Allgemein gilt

$$S_{\text{OUT}}(t) = a_0 + \sum_{i=1}^{\infty} a_i \sin(i\omega(t - \varphi)) \quad (\text{A.1})$$

mit der Amplitude der i -ten Harmonischen a_i , der Kreisfrequenz ω und der beliebigen Phasenverschiebung φ . Betrachtet man die Berechnung der Distanz für das Verfahren (3.12) zeigt sich, dass nur ungerade Harmonische einen Einfluss auf die berechnete Distanz haben. Gerade Harmonische werden durch die Subtraktion der um jeweils π verschobenen Zeitfenster eliminiert. Abbildung A.2 zeigt die Kennlinie bei einem idealen Sinussignal mit einer Amplitude von 10 MHz sowie mit zusätzlich der dritten Harmonischen mit selber Amplitude. Im letztgenannten Fall zeigt sich eine sinusförmige Abweichung der Kennlinie

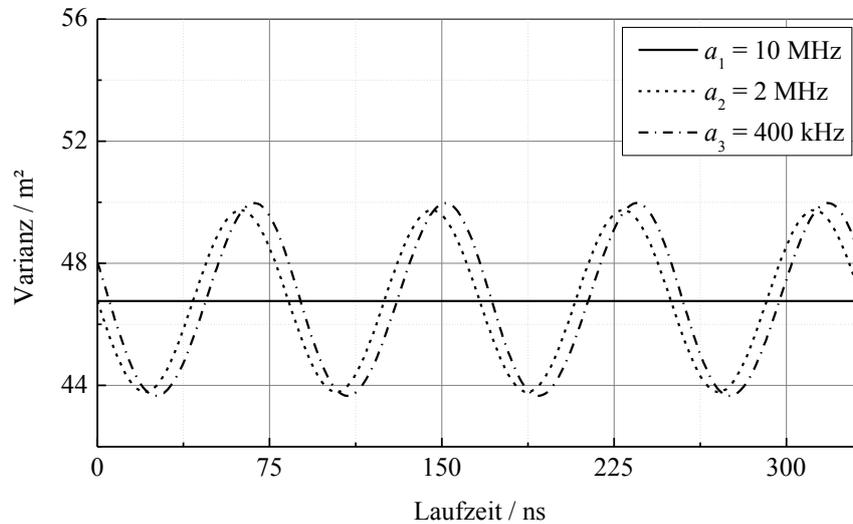


Abbildung A.3: Einfluss von Oberschwingungen im indirekten CW-Verfahren

vom Sollwert mit der vierfachen Frequenz der Grundschwingung, welche sich aufgrund der vier Messfenster ergibt und daher stets identisch ist.

Betrachtet man die Varianz des Messverfahrens, haben auch geradzahlige Harmonische einen Einfluss, da sich die Zählwerte in den einzelnen Zeitfenstern und folglich deren Varianz ändert. Abbildung A.3 zeigt die analog zu (4.48) berechnete Varianz der Distanzmessung, wobei die Varianzen der vier Zeitfenster über das Integral der nichtidealen Sinusfunktion nach (A.1) entsprechend

$$\sigma_i^2 = \int_{(i-1)T_{CW}/4 - T_W/2}^{(i-1)T_{CW}/4 + T_W/2} S_{OUT}(t) dt \quad (\text{A.2})$$

mit der Periodendauer der Grundschwingung T_{CW} und der Breite der Zeitfenster T_W bestimmt werden. Ohne vorhandene Oberschwingungen und mit einer Amplitude von 10 MHz zeigt sich analog zu Abschnitt 4.2.2.2 eine laufzeitunabhängige Varianz. Kommt eine zweite Harmonische mit einer Amplitude von 2 MHz hinzu, zeigt sich eine sinusförmige Variation über der Laufzeit. Bei einer zusätzlichen dritten Harmonischen zeigen sich nur eine geringe Verschiebung des Verlaufs sowie eine Amplitudenänderung.

Allgemein kann durch eine entsprechende Kalibration der systematische Fehler behoben werden, während der zufällige Fehler nicht entfernt werden kann. Liegen die Amplituden der Oberschwingungen aber deutlich unterhalb der Grundschwingung, ist entsprechend dem gezeigten Beispiel der Einfluss auf die Varianz der Distanzmessung gering. Da ein idealer Sinusverlauf nicht erzeugt werden kann, ist beim indirekten CW-basierten Verfahren die Kalibration schwieriger und wichtiger als bei der Messung mit PM.

B Schärfentiefe und Auflösungsvermögen des Sensors

In diesem Abschnitt werden auf Basis des Sensorlayouts und der verwendeten Optik die Schärfentiefe und das Auflösungsvermögen des Sensors betrachtet.

B.1 Schärfentiefe

Die Schärfentiefe gibt den Abstandsbereich vor der Kamera an, innerhalb dessen Objekte scharf abgebildet werden. Eine Abbildung ist scharf, wenn ein Objektpunkt durch die Optik auf ein einzelnes Pixel abgebildet wird [84]. Die hyperfokale Distanz bezeichnet die Entfernung, auf welche fokussiert werden muss, damit Objekte in unendlicher Entfernung gerade noch scharf abgebildet werden. Bei dieser Fokussierung ist die Schärfentiefe maximal und erstreckt sich von der halben hyperfokalen Distanz bis ins Unendliche. Sie berechnet sich

$$d_h = \frac{f^2}{f_\# D_z} + f \quad (\text{B.1})$$

mit der Brennweite des Objektivs f , der Blendenzahl $f_\#$ und dem Durchmesser des Zerstreuungskreises D_z [85]. Wird D_z gleich dem Pixelpitch von $40,56 \mu\text{m}$, die Brennweite zu 12 mm und die Blendenzahl zu $1,3$ gewählt, ergibt sich die hyperfokale Distanz zu $2,743 \text{ m}$. Wird auf diese Distanz fokussiert, werden Objekte ab einer Entfernung von $1,37 \text{ m}$ scharf abgebildet. Um die Schärfentiefe zu steigern, kann die Blende reduziert werden. Sollen bspw. Objekte ab einer Distanz von 1 m scharf sein, muss die Blendenzahl auf $1,8$ angehoben werden. Da der Sensor allerdings nur eine maximale Distanz von 192 m erfassen kann, ist

es nicht notwendig Objekte im unendlichen scharf abzubilden. Der Fernpunkt, d. h. die Entfernung bis zu welcher Objekte scharf sind, berechnet sich zu

$$d_f = \frac{g(d_h - f)}{(d_h - f) + (f - g)} \quad (\text{B.2})$$

mit der Fokussierung auf die Distanz g . Wird der Fernpunkt auf 192 m gelegt, erhält man einen Fokuspunkt von 2,7 m, was nahezu der hyperfokalen Distanz entspricht. Es zeigt sich damit, dass je nach Anwendung des Sensors der Fokus angepasst werden muss. Wird der Sensor für Notbremssysteme im Automobil eingesetzt, ist es ausreichend, Objekte erst ab einer Entfernung von 1,3 m scharf abzubilden, da nähere Objekte trotz einer gewissen Unschärfe erkannt und ein Bremsvorgang eingeleitet werden kann. Ferne Objekte hingegen müssen scharf abgebildet werden, da eine präzise Ermittlung der Position hilft Bewegungsabläufe vorherzusagen und Unfälle zu verhindern. In anderen Anwendungen, wie Gestenerkennung oder Einparkhilfen, muss der Fokus auf kurze Distanzen eingestellt werden. Hierbei sind ferne Objekte nicht von Interesse.

B.2 Auflösungsvermögen

In klassischen Kamerasystemen beschreibt das horizontale Auflösungsvermögen mit welchem Kontrast eine Struktur in der Szene erfasst werden kann [84]. Sie ist durch den Pitch und den Füllfaktor des Sensors bestimmt. Zur Beurteilung des Auflösungsvermögens dient die Modulationsübertragungsfunktion (MTF). Sie beschreibt das Verhältnis zwischen dem Kontrast der Bildaufnahme und der Szene. Abbildung B.4 illustriert die Herleitung der eindimensionalen MTF: Bei geringen Ortsfrequenzen liegt der Abtastbereich des Sensors, d. h. der Bereich, welchen die Detektoren des Sensors erfassen (grau hinterlegter Bereich), nahezu vollständig im Minimum bzw. Maximum der Sinuswelle. Folglich kann die Frequenz mit einem Kontrast nahe 100 % erfasst werden. Bei höheren Frequenzen hingegen deckt der Abtastbereich einen größeren Bereich der Sinuswelle ab [86]. Da der Detektor den Mittelwert des erfassten Bildbereichs wiedergibt, liegen dieser nun deutlich über dem Minimum bzw. Maximum der

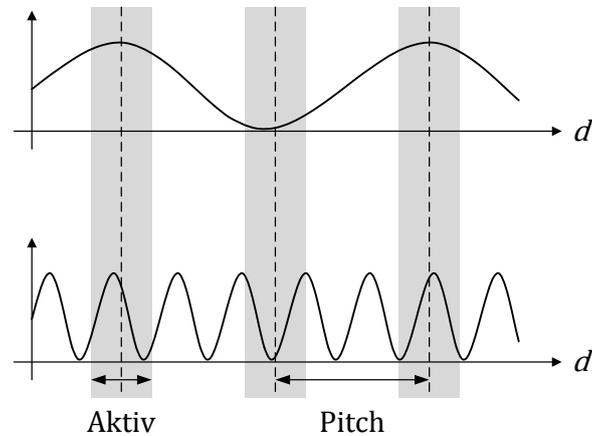


Abbildung B.4: Illustration der Modulationsübertragungsfunktion

Sinuswelle und der Kontrast in demzufolge geringer. Betrachtet man das sich ergebende Maximum bei Abtastung der Sinuswelle erhält man

$$k_{\max} = \frac{1}{2a} \int_{-\frac{a}{2}}^{\frac{a}{2}} 1 + \cos(2\pi f_x x) dx = \frac{1}{2a} \left(a + \frac{\sin(\pi f_x a)}{\pi f_x} \right) \quad (\text{B.3})$$

mit der Detektorbreite a und der Ortsfrequenz f_x . Für das Minimum gilt analog

$$k_{\min} = \frac{1}{2a} \int_{-\frac{a}{2}}^{\frac{a}{2}} 1 - \cos(2\pi f_x x) dx = \frac{1}{2a} \left(a - \frac{\sin(\pi f_x a)}{\pi f_x} \right). \quad (\text{B.4})$$

Die MTF ist nun durch die Differenz der beiden Funktionen dividiert durch den Kontrast der Sinuswelle, welcher 1 beträgt, gegeben zu

$$MTF = k_{\max} - k_{\min} = \frac{\sin(\pi f_x a)}{\pi f_x a}. \quad (\text{B.5})$$

Die Funktion hat ihre erste Nullstelle bei $1/a$. Um Aliasing, d.h. die Fehlerinterpretation von Strukturen in der Szene, vollständig zu eliminieren, muss die MTF ab der halben Abtastfrequenz (Nyquist-Kriterium), welche durch den Pitch p zu

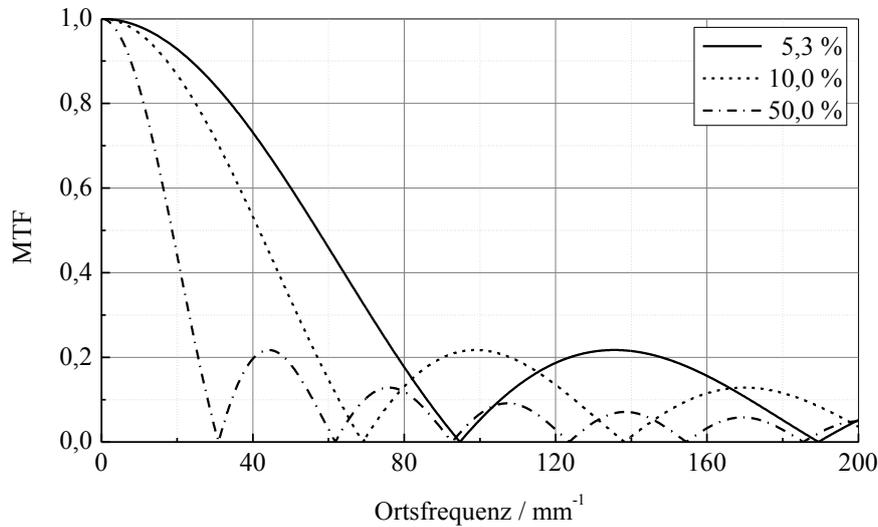


Abbildung B.5: MTF für einen Pitch von 40,56 μm und verschiedene Füllfaktoren

$1/2p$ gegeben ist, null sein. Um den Füllfaktor mit in die Gleichung der MTF zu bringen, kann bei einem runden Detektor dessen Breite in der Form

$$a = p \sqrt{\frac{4}{\pi} \eta_{\text{FF}}} \quad (\text{B.6})$$

ausgedrückt werden. Aus (B.5) folgt damit

$$MTF = \frac{\sin\left(\pi f_x p \sqrt{\frac{4}{\pi} \eta_{\text{FF}}}\right)}{\pi f_x p \sqrt{\frac{4}{\pi} \eta_{\text{FF}}}}. \quad (\text{B.7})$$

Abbildung B.5 zeigt die MTF nach (B.7) für den horizontalen Pitch des Sensors von 40,56 μm und den Füllfaktor von 5,32 %. Zur Veranschaulichung des Effektes eines höheren Füllfaktors sind zudem die Kurven für Werte von 10 % und 50 % dargestellt. Die Nyquist-Frequenz des Sensors liegt bei 12,32 mm^{-1} , welche von allen Kurven deutlich überschritten wird. Dennoch zeigt sich, dass sich die erste Nullstelle mit steigendem Füllfaktor zu niedrigeren Frequenzen verschiebt und so der Aliaseffekt reduziert wird. Bei einem Füllfaktor von 100 % entspricht die Nullstelle der doppelten Nyquist-Frequenz.

Es stellt sich die Frage, inwiefern sich die MTF auf einen Sensor zur Tiefenerfassung der Szene übertragen lässt. Der im Rahmen der Arbeit entworfene Sensor arbeitet nach der direkten Methode zur Laufzeitmessung. Hierbei wird die Zeitmessung mit dem ersten detektierten Ereignis nach dem Aussenden des Laserpulses gestoppt. Liegen im FOV des Pixels mehrere Objekte in unterschiedlicher Distanz, wird nur die kürzeste Distanz gemessen, da das hier reflektierte Licht den Sensor zuerst erreicht und die Zeitmessung stoppt. Ist der Reflexionsfaktor jedoch gering oder die Distanz sehr hoch, kann es aufgrund der geringen reflektierten Leistung dazu kommen, dass die Zeitmessung durch eines der ferneren Objekte angehalten wird. Bei der Akkumulation mehrerer Laserpulse im Histogramm führt dies zu einer Verbreiterung des Pulses, wobei der Pulsanfang immer noch bei der geringsten Distanz liegt. D. h. befinden sich mehrere Objekte vor einem Pixel, kann es unter bestimmten Bedingungen durchaus zu einer Verschiebung der gemessenen Distanz führen, allgemein lässt sich die MTF aber nicht übertragen, da in der direkten Laufzeitmessung keine Mittelung der im FOV des Pixels befindlichen Distanzen erfolgt. Bei einer ausreichend hohen Reflektanz der Zielobjekte und Laserleistung wird immer das Objekt in kürzester Entfernung erkannt. Diese Eigenschaft des direkten Verfahrens ist für die Zielanwendungen wie Kollisionsschutz im Automobil ein Vorteil gegenüber dem indirekten Verfahren, bei welchem eine Mittelung der Distanz durch die Integration der einfallenden Lichtmenge innerhalb der Zeitfenster stattfindet.