Control Algorithm and Evaluation Strategies for Single-Photon Avalanche Diode-Based Direct Time of Flight Systems

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

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaft

genehmigte Dissertation

von
Andre Henschke
aus
Duisburg

Datum der Einreichung: 29.03.2022

Referent: Prof. Dr. Anton Grabmaier
Koreferent: Prof. Dr. Andreas Stöhr

Tag der mündlichen Prüfung: 31.10.2022

For Lili.

Versicherung an Eides Statt

Ich versichere durch meine Unterschrift, dass ich die vorstehende Arbeit - mit Ausnahme der Anleitung durch die Betreuer - selbständig und ohne fremde Hilfe angefertigt und alle Stellen, die ich wörtlich oder dem Sinne nach aus Veröffentlichungen entnommen habe, als solche kenntlich gemacht habe, mich auch keiner anderen als der angegebenen Literatur oder sonstiger Hilfsmittel bedient habe.

Ich versichere an Eides Statt, dass ich die vorgenannten Angaben nach bestem Wissen und Gewissen gemacht habe und dass die Angaben der Wahrheit entsprechen und ich nichts verschwiegen habe.

Die Strafbarkeit einer falschen eidesstattlichen Versicherung ist mir bekannt, namentlich die Strafandrohung gemäß § 156 StGB bis zu drei Jahren Freiheitsstrafe oder Geldstrafe bei vorsätzlicher Begehung der Tat bzw. gemäß § 163 Abs.1 StGB bis zu einem Jahr Freiheitsstrafe oder einer Geldstrafe bei fahrlässiger Begehung.

Duisburg, den 29.03.2022

Ort, Datum	Andre Henschke

This thesis was written during my time as a research assistant at the Fraunhofer Institute for Microelectronic Circuits and Systems in Duisburg between 2019 and 2022. I would like to thank my direct supervisors Werner Brockherde, Prof. Anton Grabmaier, Manuel Ligges and Jennifer Ruskowski for their professional support and guidance in my work.

Further thanks go to my colleagues, including, but not limited to, Roman Burkard, Sara Grollius, Jan F. Haase, Olaf M. Schrey, Denis Piechaczek, Simon Grosse, Andreas J. Gisen, Charles Thattil, Yoon-Cue Lee and Maik Beer, whose scientific input has refined my findings time and again and for making work fun.

Special thanks go to Prof. Bedrich Hosticka for the countless technical discussions and his infectious enthusiasm for research.

Apart from that I would like to thank my parents, family and friends for their emotional support, and for their patience with me.

Abstract

Single-Photon Avalanche Diode-based Direct Time of Flight systems are currently attracting much attention due to the need for optical systems with high temporal and spatial resolution, high frame rates, range, and high ambient light rejection for autonomous driving applications. Within the scope of this thesis, methods aiming to detect targets under high ambient light power are evaluated. Different pixel structures are considered, and the Signal-to-Noise Ratio derived as a measure of their performance.

On this basis, a novel comparison between commercially available Avalanche Photodiode and Single-Photon Avalanche Diodes is conducted. It shows the latter's advantages when only small return signals are present. The theoretical predictions are then investigated with numerical simulation and verified using available research prototype systems.

With high frame rate and long range come small signal photon budgets. A rank-order filter-based signal processing chain for detecting faint return signal is investigated and compared to proven correlation-based approaches. It is shown that while the rank-order filters show excellent background light rejection, the signal is attenuated aggressively, limiting their usefulness in long-range ranging applications.

Strong return signals distort the recorded pulse distribution and limit accuracy. An approach for computationally correcting this accuracy deterioration from the measurement result for a given temporal laser pulse form is designed with a view to hardware integration. It can be corrected for signal peak event rates of up to 0.2 GHz.

The investigations result in a developed sensor control algorithm based on analysis of the recorded signal strength which yields high frame rates in low ambient light scenarios and high range in high ambient light scenarios. A range

increase of up to 20 % over systems without an active control algorithm can be shown, increasing to up to 700 % for systems without active ambient light suppression.

Zusammenfassung

Single-Photon Avalanche Diode-basierte Direct Time of Flight-Systeme genießen derzeit große Aufmerksamkeit, da für Anwendungen im Bereich des autonomen Fahrens optische Systeme mit hoher zeitlicher und räumlicher Auflösung, hohen Bildraten, Reichweite und Fremdlichtunterdrückung benötigt werden. Im Rahmen dieser Arbeit werden Methoden zur Objekterfassung bei hoher eingestrahlter optischer Fremdlichtleistung evaluiert. Es werden verschiedene Pixelstrukturen betrachtet und der Signal-Rauschabstand als Maß für ihre Leistungsfähigkeit berechnet.

Auf dieser Grundlage wird ein neuartiger Vergleich zwischen kommerziell erhältlichen Avalanche Photodioden und Single-Photon Avalanche Dioden durchgeführt. Letztere sind vorteilhaft, wenn nur eine geringe rückgestrahlte optische Leistung zur Verfügung steht. Die theoretischen Vorhersagen werden anschließend mit numerischen Simulationen untersucht und mit verfügbaren Forschungsprototypen verifiziert.

Durch die hohen Anforderungen an Reichweite und Bildrate stehen den Systemen oftmals nur eine niedrige Zahl an Signalphotonen zur Verfügung. Eine auf Rangordnungsfiltern basierende Signalverarbeitungskette wird untersucht und mit bewährten korrelationsbasierten Ansätzen verglichen. Es zeigt sich, dass die Rangordnungsfilter zwar zu einer hervorragenden Unterdrückung des Hintergrundlichts führen, das Signal jedoch stark abgeschwächt wird, was ihre Nützlichkeit für Anwendungen im Fernbereich einschränkt.

Eine hohe Signalstärke verzerrt das aufgezeichnete Messergebnis und verringert die Messgenauigkeit. Ein Ansatz zur rechnerischen Korrektur dieses Genauigkeitsfehlers aus dem Messergebnis für eine gegebene Laserpulsform wird im Hinblick auf ihre Hardware-Integrierbarkeit untersucht und entwickelt. Mit ihr kann die Messwertverschiebung für Ereignisraten von bis zu 0.2 GHz korrigiert werden.

Die Untersuchungen münden in einem entwickelten Algorithmus, der auf der Analyse der aufgenommenen Signalstärke basiert. Er ermöglicht hohe Bildraten bei schwachem Umgebungslicht und eine hohe Reichweite bei starkem Umgebungslicht. Es kann eine Reichweitensteigerung von bis zu 20 % gegenüber Systemen ohne aktiven Kontrollalgorithmus gezeigt werden, die sich bei Systemen ohne Fremdlichtunterdrückung auf bis zu 700 % erhöht.

Table of Contents

List of Figures

Glossary

Abbreviations

AEL	Accessible Emission Limit
AMCW	Amplitude Modulated Continuous Wave
APD	Avalanche Photodiode
AQR	Active Quenching and Reset
AWGN	Additive White Gaussian Noise
BBT	Band-to-Band Tunneling
BSI	Backside Illumination
CD	Coincidence Depth
CDF	Cumulative Distribution Function
CIS	CMOS Imaging Sensor
CMOS	Complementary Metal-Oxide-Semiconductor
DCR	Dark Count Rate
DLL	Delay-Locked Loop
DToF	Direct Time of Flight
DUT	Device Under Test
FF	Fill Factor
FIR	Finite Impulse Response
FLIM	Fluorescence Lifetime Imaging
FMCW	Frequency Modulated Continuous Wave
FoV	Field of View
FPS	Frames per Second
FSI	Front-Side Illuminated
FWHM	Full Width Half Maximum

HP	Homogeneous Poisson Process
IHP	Inhomogeneous Poisson Process
IIR	Infinite Impulse Response
IRF	Instrument Response Function
IToF	Indirect Time of Flight
ITS	Inverse Transform Sampling
LED	Light Emitting Diode
LiDAR	Light Detection and Ranging
MCP	Multi-Channel Plate
ME	Multi-Event
MEMS	Micro Electro Mechanical System
NDF	Neutral Density Filter
NIR	Near-Infrared Regime
PDE	Photon Detection Efficiency
PDF	Probability Density Function
PDP	Photon Detection Probability
PLL	Phase-Locked Loop
PMT	Photomultiplier Tube
PRF	Pulse Repetition Frequency
PS	Pulse Shaper
RADAR	Radio Detection and Ranging
RS	Rejection Sampling
SCR	Space Charge Region
SPC	Single Photon Counting
SNR	Signal-to-Noise Ratio
SPAD	Single-Photon Avalanche Diode
STI	Shallow Trench Isolation
TAC	Time-to-Amplitude Converter
TCSPC	Time-Correlated Single Photon Counting
TDC	Time-to-Digital Converter
TIA	Transimpedance Amplifier
ToF	Time of Flight

VCO	Voltage-Controlled Oscillator

Symbols

A_{illum}	Area illuminated by laser
A_{Lam}	Area of Lambertian half-sphere
A_{Pix}	Pixel Area
A_{Rec}	Receiver Aperture Area
α_{H}	Horizontal Source Opening Angle
α_{V}	Vertical Source Opening Angle
c	Speed of Light
c_0	Speed of Light in Vacuum
d	Target Distance
d_{Pix}	Pixel Diameter
D	Aperture Diameter
e	Elementary Charge
E_{e}	AM1.5G Irradiance in Wavelength Range at Target
E_{e0}	Total AM1.5G Irradiance at Target
E_λ	Spectral Solar Irradiance
E_{V}	Illuminance in Wavelength Range at Target
E_{V0}	Total Illuminance at Target
η_{A}	One-way Atmosphere Transmission Efficiency
η_{FF}	Fill factor
η_{Opt}	Optics Transmission Efficiency
η_{PDP}	Photon Detection Probability
η_{PDE}	Photon Detection Efficiency
η_{S}	Sensor System Efficiency
f	Focal Length
$f_\#$	f-Number
f_{FPS}	Frame Rate
f_{n}	Probability Density Function of the n-th Arrival
f_{PR}	Pulse Repetition Frequency
f_{r}	Probability Density Function of Rectangular Pulse
f_{S}	Probability Density Function

F_{S}	Cumulative Distribution Function
$F_{\mathrm{S,R}}$	Cumulative Distribution Function of Background and Rectangular Pulse
$F_{\mathrm{S,S}}$	Cumulative Distribution Function of Background and Sech2 Pulse
$F_{\mathrm{S,B}}$	Cumulative Distribution Function under Background Influence
H	Histogram
h	Planck Constant
i_{Max}	Index of Maximum Bin
K	Luminous Efficacy
k	Events of Poisson Process
k_{SN}	Histogram Signal-to-Noise Ratio
λ	Wavelength
λ_0	Operating Wavelength
μ	Expected Value
$\mu_{i,\mathrm{B}}$	Expected Value of Bernoulli Process
$\mu_{i,\mathrm{P}}$	Expected Value of Poisson Process
μ_{P}	Expected Number of Arrivals
$\mu_{\mathrm{P,L}}$	Expected Number of Arrivals from Laser
$\mu_{\mathrm{P,B}}$	Expected Number of Arrivals from Background
$m(t)$	Expected Number of Arrivals
N	Number of Bernoulli Trials
N_{e}	Number of Generated Electrons
N_{M}	Number of Measurements
N_{O}	Number of Incident Photons
N_{Ph}	Number of Photon Arrivals
N_{S}	Number of Detectors in Pixel
n	Bin Count Value
p_i	Probability for Successful Outcome in Bernoulli-Trial
p_{B}	Bin Triggering Probability through Background
p_{L}	Bin Triggering Probability through Laser
$p_{\mathrm{L,R}}$	Bin Triggering Probability through Laser (Rectangular Pulse Model)

$p_{\mathrm{L,S}}$	Bin Triggering Probability through Laser (Sech2 Pulse Model)
P	Probability
P_{A}	Optical Ambient Light Power on Detector
P_{AP}	Afterpulsing Probability
P_{C}	Probability for k Events during Coincidence Time
P_{CT}	Crosstalk Probability
$P_{i,\mathrm{B}}$	Probability in Bernoulli Trial for Count Value in Bin
P_{L}	Optical Power on Detector
$P_{\mathrm{P}}(k)$	Probability for k Poisson distributed arrivals
P_{T}	Peak Optical Output Power
q_i	Probability for Unsuccessful Outcome in Bernoulli-Trial
r	Event rate
$r(t)$	Instantaneous Time-dependent Rate
r'	Event Rate per SPAD
r_{A}	Ambient Event Rate
r_{B}	Undesired Background Event Rate
r_{Count}	Counting Rate of Diode
r_{DC}	Dark Count Rate
$r_{\mathrm{L,0}}$	Peak Rate of Sech2 Pulse
$r_{\mathrm{L,R}}$	Peak Rate of Rectangular Pulse
r_{R}	Rate Function of Rectangular Pulse and Ambient
r_{S}	Event Rate per Diode
r_{SS}	Steady-state Count Rate
ρ	Target Reflectivity Coefficient
σ	Standard Deviation
σ_{c}	Optical Cross Section
$\sigma^2_{i,\mathrm{B}}$	Standard Deviation of Bernoulli Process
$\sigma^2_{i,\mathrm{P}}$	Standard Deviation of Poisson Process
σ_{t}	Standard Deviation of Quantization
t_{Bin}	Bin Size
$\overline{t_{\mathrm{C}}}$	Average Time between Counts
t_{C}	Coincidence Time
t_{D}	Dead Time

t_{Hist}	Observation Time of Histogram
$\overline{t_1}$	Mean Time to First Detection out of Multiple Detectors
$\overline{t_2}$	Mean Time for k Detections out of Multiple Detectors
t_{P}	Pulse Width of Rectangular Pulse
τ_{P}	Scaling Parameter of Sech2 Pulse
t_{ToF}	Time of Flight
Θ	Heaviside Function
θ_{H}	Horizontal Detector Opening Angle
θ_{V}	Vertical Detector Opening Angle
$V(\lambda)$	Luminous Efficiency
W_{Ph}	Photon Energy

Chapter 1

Introduction

This chapter places Single-Photon Avalanche Diode (SPAD) devices and Time of Flight (ToF) systems in their historical context. Then, contributions to the state of literature given by this thesis are summarized.

1.1 Historical Context

A distance can be measured directly with a physical reference like measuring tape or indirectly by capturing an auxiliary quantity like the time it takes for an object with defined velocity to transit the given distance. The latter principle is called ToF measurement. Light is attractive for distance determination because of its well-defined and high velocity.

In optical ToF systems, photons are sent out at a defined reference time, transmitted through the air, possibly scattered, and reflected by objects in the scene. A detector capable of resolving the arrival times or the number of impinging photons can derive the object distances. Many optical detectors achieve this by having the incident photons generate electron-hole pairs via the photoelectric effect and detecting them as a current between the diode's anode and cathode.

However, optical detection can also prove challenging. The optical power budget varies as remitted laser power is distance-dependent. Ambient lighting is the most important source of interference and uncontrollable. Object characteristics can range from big and highly reflective to small and obscure. The systems thus have to be able to detect over a wide range of power budgets. Without modulation, it is impossible to distinguish whether a single incident photon corresponds

to the system's signal or stems from ambient light.

The first ToF camera, capable of simultaneous distance measurement of different points in space, was based on regular p-n photodiodes. Photodiodes are comparatively easy to build and require simple analog readout electronics. They also have inherent disadvantages. Compared to other detectors, their sensitivity is limited. Optical intensity can only be resolved if the number of generated carriers exceeds the analog noise floor of the circuit. Noise sources include thermal generation of carriers in the diode, photon shot noise, 1/f noise, and the general noise behavior of the readout electronics.

There have been considerable efforts to increase the performance of photodiodes. Besides technological improvements on the general p-n diode structure, regions with high electrical fields were introduced to let electron-hole pairs generate other carriers leading to an internal gain. Few photons can trigger many electron-hole pairs increasing the diode's output current, which can be high enough to be detected over other circuit noise. This type of device is called an Avalanche Photodiode (APD), and it has low noise and high gain characteristics. Today APDs use in optical ToF cameras is often limited to systems with time-dependent optical axes to achieve lateral resolution. Flash systems, recording the whole scene in each frame, are hard to realize with APDs because of their limited array-capability, with the largest scale implementation to date being 320×256 pixels [1].

Long-range and compact ToF systems have to estimate the distance based on low photon numbers. Emitted energy must comply with eye safety regulations that aim to protect humans' and animals' well-being in the proximity of laser-based systems. Thus, emitted photon counts are limited, making sensitivity a critical parameter. Diffuse targets and the need for compact systems with small apertures further amplify the need to efficiently use the available photon budget.

While APD provide good sensitivity and conversion efficiency, they cannot resolve single photons at room temperature. One advancement is to extend the APD working principle via even higher biasing voltages. If the space-charge region contains a sufficiently low number of defects, the diode can be carefully biased above its standard breakdown voltage, resulting in a meta-stable state. Incident photons now 'trigger' the diode, sparking a saturation current called avalanche breakdown which can be time-resolved via external circuitry. This time can be resolved finely, giving a precise value for the photon arrival time and making

SPADs attractive candidates for the design of optical ToF systems.

Implementations of SPAD based ToF systems have been demonstrated by a variety of research groups. In recent developments on the technological side, SPAD manufacturers aim to further increase SPAD Photon Detection Efficiency (PDE) to reach values which are comparable to the high sensitivity values of APDs. Higher sensitivity opens up the perspective of achieving an even higher range.

On the system side, research groups are striving for optimal readout circuits and evaluation algorithms for achieving high range while also balancing requirements on data throughput rates, readout circuit, algorithmic complexity, and production cost. This is where this work picks up.

1.2 Structure of the Work and Contributions

This work develops evaluation schemes for SPAD-based, pulsed ToF systems. After introducing the theoretical fundamentals of ToF measurements and SPAD detector technology in Chapter 2, multi-pixel structures and differences to APD receivers are discussed via a novel analytical approach in Chapter 3. It can be shown that SPAD detectors are advantageous in low power regimes while they lack the background light rejection of APDs and how readout circuitry of the SPAD detector improves ambient light suppression. A simulation engine capable of estimating receiver performance more flexibly than analytical approaches is introduced in Chapter 4. In Chapter 5, processing schemes are discussed. The performance of digital Finite Impulse Response (FIR) and rank-order filters is assessed in Section 5.4 and compared to a developed matched filter. Trade-offs and advantages in different operating conditions are shown, and their complexity in hardware implementations is briefly discussed. For resolving intensity-related errors, an approach using a real pulse intensity model is developed in Section 5.5, and its applicability and limitations are presented.

The work closes with Chapter 6 and a discussion of an evaluation scheme for setting optimal parameters of smart pixels capable of coincident and multi-event detection. The approach is validated with measurements on a real system. All considered sensor architectures, pre-processing, and evaluation schemes aim to achieve maximum range in adverse sunlight conditions.

Chapter 2

Theoretical Fundamentals

This chapter introduces theoretical prerequisites for understanding the developed concepts later in this work. After discussing the fundamental operating principle of ToF-based optical sensing, detection behavior and noise influences of SPADs are discussed. Finally, the relationship between optical intensity on the detector and statistics of generated electrical signals are presented, focusing on the influence of different operational modes.

2.1 Time of Flight Measurement Principle

A ToF measurement aims to determine a distance d by measuring the time t_{ToF} between emission and detection of a signal with known velocity. Optical ToF systems, in which emitter and detector are placed approximately on the same optical axis, measure the round-trip delay with a velocity equivalent to the speed of light c. Without disturbances like rain or fog, the speed of light in air can be approximated by the speed of light in vacuum c_0.

$$d = \frac{1}{2} t_{\mathrm{ToF}} c_0 \tag{2.1}$$

A basic ToF measurement setup based on SPADs as the optical detector is shown in Figure 2.1. A laser emits a signal pulse at a certain point in time. At the same time, a digital stopwatch generates a timestamp. The pulse travels through the channel, scatters and reflects from objects, and partially enters the receiving optics. The SPAD detector is triggered, and a second timestamp is generated. The difference between the two timestamps gives an estimate for t_{ToF}.

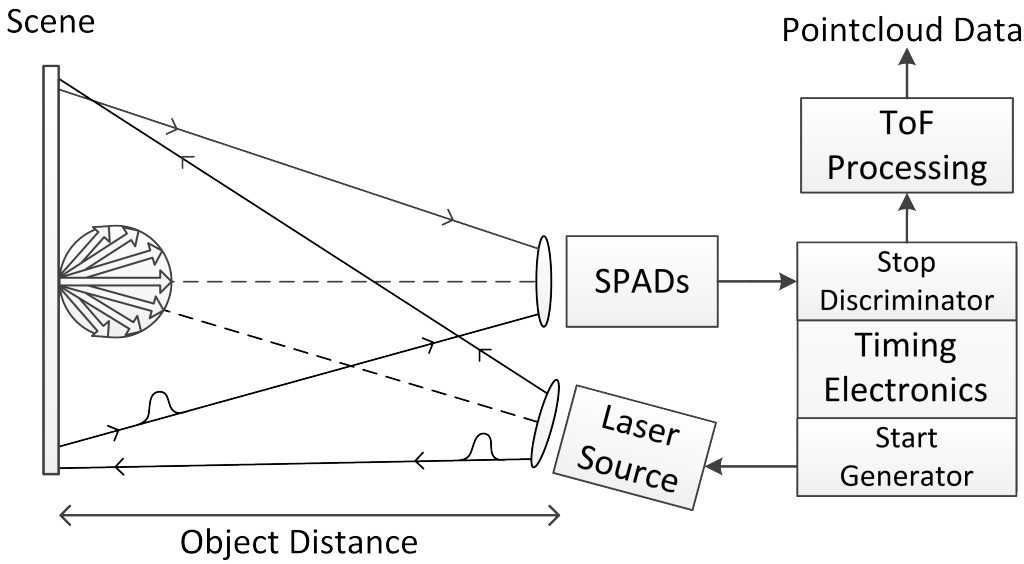


Figure 2.1: Setup for a SPAD-based DToF measurement.

A depth map of the scene is acquired by repeating this measurement for different points in space.

The statistical nature of photon detection and finite detection efficiencies makes depth estimation based on a single photon return unfeasible. High photon counts per pulse are required to operate in statistically favorable conditions, meaning high returned laser power. Additionally, multiple iterations of photon emission and detection can be conducted to differentiate between incoming ambient and signal photons through statistics.

There is a variety of different detection principles that can be used to measure ToF optically without resorting to stereoscopic approaches, some of which are based on sensing in the time domain, others use information from the frequency domain. These systems can be summarized as Light Detection and Ranging (LiDAR) systems as they all have in common that they use reflected light to measure distance.

An overview of the measurement principles, resolution, and range of realized systems from the last years is given in Figure 2.2. Frequency Modulated Continuous Wave (FMCW) LiDAR [3] is a modulated measurement principle that achieves high resolution but remains limited to short-range applications. Amplitude Modulated Continuous Wave (AMCW) LiDAR systems [4] can cover ranges of a few meters with high precision but are incapable of achieving high

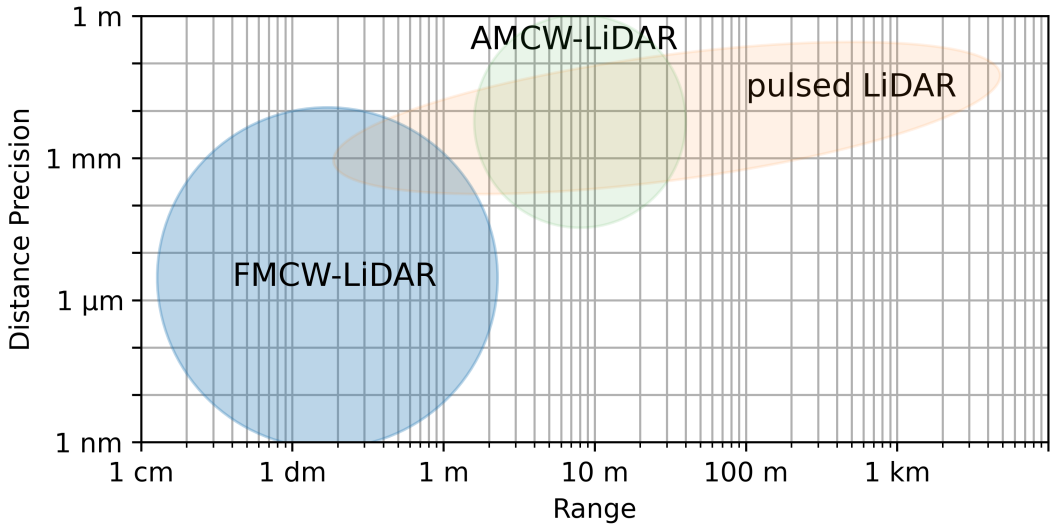


Figure 2.2: Comparison of achieved range and precision for realized systems with different measurement principles, after [2].

range [2]. Triangulation-based sensors are also precise but limited to short-range applications [5].

Additionally, SPAD performance in applications based on amplitude resolving like FMCW or AMCW LiDAR is limited. Acquiring the original pulse form with SPADs is difficult because the dead time makes them miss detections, and their triggering behavior is insensitive to phase relations. This work focuses on the Direct Time of Flight (DToF) method. It is a pulsed method called 'direct' because arrival times of photons are recorded instead of resolving their number during a time interval. Ideally, only the laser signal is detected in a DToF measurement. [6]

In DToF, laser Pulse Repetition Frequency (PRF) is connected to the system's maximum range. Converted to a distance via equation (2.1), the inverse of the PRF determines the so-called unambiguous range of the system. If objects at a higher distance than the unambiguous range reflect laser pulses, the optical path is so long that when the first pulse arrives, the next laser pulse is already sent out. Without further data, it can not be concluded whether the return is from an object far away, illuminated by the first laser pulse, or a close object illuminated by the second laser pulse. The range information provided is thus ambiguous. The unambiguous range limiting laser PRF has an indirect effect on achievable frame rates assuming that an image needs a fixed number of illuminations.

DToF is advantageous to modulated approaches for achieving high range because it relies on short pulse durations. Thus, high pulse peak powers are admissible following eye-safe laser operation guidelines. The temporal length of optical emission is longer in modulated approaches, so the maximum output power is more limited. The signal modulation could be beneficial for ambient light rejection, however. When comparing realized systems, DToF achieves a higher maximum range than Indirect Time of Flight (IToF) or modulated approaches. [2]

2.2 Single-Photon Avalanche Diodes

The detector's timing response and sensitivity are essential factors for achieving high accuracy in ToF measurements. Photomultiplier Tubes (PMTs) and Multi-Channel Plates (MCPs) that were used historically provided high time resolution and sufficient detection efficiency but are fragile and require high voltages and vacuum tubes for their operation [7].

Technologically developed from APDs but in functionality closer to triggering devices like PMTs, SPADs can provide sub-nanosecond timing accuracy while compatible with Complementary Metal-Oxide-Semiconductor (CMOS) processes, are array-capable, and not sensitive to magnetic fields. The first CMOS-based array implementation was demonstrated in 2004 [8], enabling integrated readout-circuitry and opening up new markets through dropping device costs. In this section, basic operating principles of SPAD devices are shown, and external circuitry for processing their output signals is discussed.

2.2.1 Structure

Two types of SPAD devices differ in the width of their absorption layer: thin and thick junction. Thick junction SPADs have the advantage that, following Beer's law, their extended absorption zone provides high photon conversion efficiency. However, they exhibit a poor timing response, making them less attractive in high precision ranging applications.

Thin junction SPADs are the more common type to date. They exhibit lower detection probability but typically have picosecond time resolution [7]. The structure of a thin junction SPAD in its typical implementation as Front-Side Illuminated (FSI) is shown in Figure 2.3. Independently from the chosen integra-

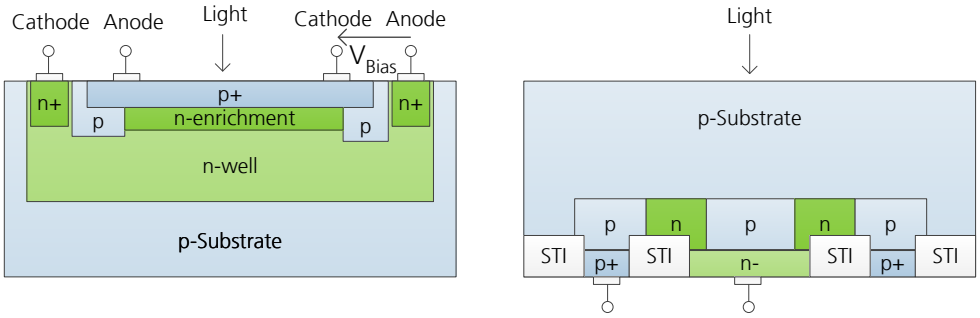


Figure 2.3: Schematic structures of planar thin SPADs. Left: FSI SPAD adapted from [7] and [9]. Right: Implementation as a BSI SPAD after [10].

tion type, the diode is surrounded by guarding structures to prevent premature edge breakdown and suppress electrical and optical interference between devices. Shallow Trench Isolation (STI) can improve on this suppression by separating differently doped zones. As in digital cameras, Backside Illumination (BSI) sensors [11] can be used to achieve higher electronics integration densities without sacrificing fill factor. They are called BSI sensors because, in contrast to FSI sensors, the optically active area is their substrate. It is thinned down to a few micrometers during manufacturing to let photons transmit into optically active areas. Wafer-bonding techniques make integration densities solely limited by SPAD geometry possible. The digital readout electronics are then integrated on a second wafer and bonded to the optically active wafer via bonding. This readout wafer can be designed in a smaller technology node than the wafer holding the SPADs. Furthermore, with BSI, the thickness of the multiplication region can be designed to achieve high sensitivity in the near-infrared regime, which is especially attractive for ranging applications. Because of the aforementioned reasons, BSI SPADs are currently researched intensively and are increasingly commonly used.

2.2.2 Operating Principle

The operating principle of the SPAD is based on the internal photoelectric effect. Incident photons generate electron-hole pairs. An external electric field applied over the p-n-junction accelerates these carriers to the diode's leads. In this regard, they do not differ from regular photodiodes. They are unique in that their

Space Charge Region (SCR) region is depleted from free carriers and designed to contain minimal defects. An incident photon generating an electron-hole pair in the SCR can now trigger the device, resulting in a measurable current. The generated carriers are accelerated with sufficient energy to generate other carriers through impact ionization, leading to this detectable saturation current. So, in contrast to analog photon detectors, the output current of the SPAD is not directly proportional to the incident optical power. The current flow means that the diode is insensitive to more incoming photons. Momentary power dissipation during avalanche breakdown can reach Watts and would lead to the thermal destruction of the diode. The external voltage at the SPAD terminals is reduced below breakdown voltage to quench the current.

As soon as the SCR is freed from carriers, the external biasing voltage can be increased again, bringing the SPAD back to initial conditions. Depending on the quenching circuit, SPADs have a characteristic dead time behavior. Dead time is either extended when more primary carriers are generated, called paralyzable, or remain constant, called non-paralyzable. Non-paralyzable dead time limits the performance of SPAD detectors in high flux conditions because the dead time extension increases the amount of missed incident photons.

Passive Quenching Circuits

One example of a paralyzable quenching circuit is passive quenching with a resistor. It is a self-quenching mechanism. Incident photons will lead to an avalanche breakdown and thus a current flow. This current flow leads to a voltage drop-off, bringing the diode out of Geiger mode. Passive quenching can be realized by connecting a high value (typically around 100 kΩ) series resistance at the SPAD terminal. Only limited current can flow through the high resistance to recharge the present capacitances, leading to prolonged time constants. The diode's internal resistance formed by the combined resistance of the SCR and surrounding neutral semiconductor material is added to the series resistance. They depend on the diode structure and range between several and a few thousand Ohm. The low-pass behavior of this series resistance and the stray and junction capacitances, together consisting of a few picofarads, limit the quenching speed of the device. [12]

If more photons are incident during recharge, biasing voltage build-up can be

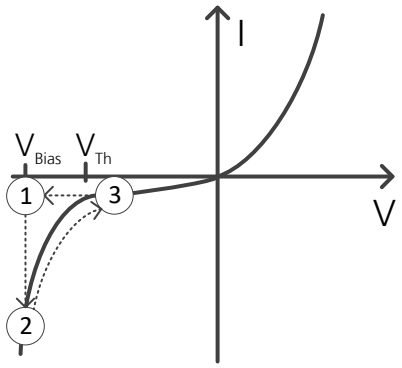


Figure 2.4: A schematic I-V curve of a SPAD and quenching operation.

reset. A generated electron-hole pair can still be detected at the output but is recorded with different sensitivity because of the lowered overvoltage. The lowered overvoltage also means that a generated analog output pulse is smaller. This height scaling can lead to timing deviations when the pulse is digitized with a comparator employing a constant threshold voltage. Furthermore, the timing response of the detector will exhibit a wider Full Width Half Maximum (FWHM) depending on the average count rate, so the output pulse is not only scaled but also has a different, intensity-dependent pulse shape. [12]

Active Quenching Circuits

Active quenching circuits aim to circumvent this limitation in achievable timing accuracy and reduce quenching times using a feedback loop, actively reacting to the triggering of an avalanche. A schematic view of the basic operating principle is shown in Figure 2.4. To put the diode into state (1), biasing voltage is ramped up in a controlled way. This is defined as the recharge phase. When an avalanche is triggered, a large current flows (2). Unlike in passive quenching circuits, the voltage drop-off is controlled actively by a quenching transistor. Because the permitted current flowing through this quenching transistor is higher than through the high resistance passive quenching resistor, the SCR can be cleared in a shorter time frame. An example high-performance, active quenching circuit can be seen in Figure 2.5. The recharge transistor supplies the SPAD with a controlled current during recharging. The quenching path contains two transistors, one to limit the current and one to disable the diode for a duration of a dead time when a hit was detected. The signals for hold-off and recharge are controlled via mono-

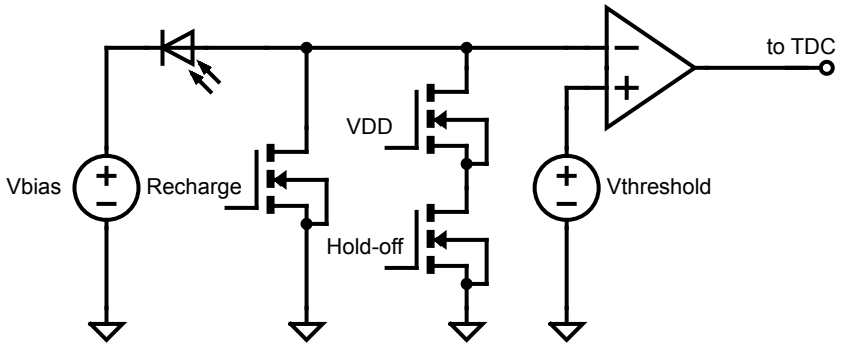


Figure 2.5: Example of an active/passive quenching circuit with low dead time from [13].

stable elements. A comparator generates a digital signal when the SPAD voltage crosses the threshold voltage. [13]

The quenching circuits employed in sensors throughout this work are all active quenching circuits with a widely non-paralyzable dead time and active reset (Active Quenching and Reset (AQR)). A mono-stable circuit is usually integrated to control dead time and set it to a constant value [12]. Placing the quenching circuitry on-chip next to the SPADs decreases the stray capacitance compared to an external quenching circuit and is beneficial to timing accuracy [14].

Detection Efficiency

Not all incident photons lead to the creation of carriers. The device's quantum efficiency is the probability of an incident photon generating an electron-hole pair. Another hurdle has to be overcome for these carriers to be detected: generating an avalanche breakdown. The primary carriers must be accelerated enough to produce other free carriers by impact ionization before recombining. This can be described as a statistical process with a certain probability. These factors are summarized in a wavelength-dependent magnitude called Photon Detection Probability (PDP). Also, there are optically active and inactive parts in the diode. Metal leads required for the electrical control reduce the exposed semiconductor surface. Guard ring structures used for electrical isolation further reduce the optically active area. The geometrical losses are summarized as the Fill Factor

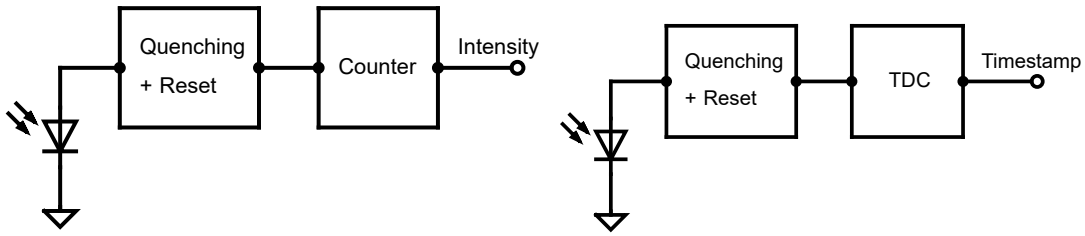


Figure 2.6: Methods for digital avalanche detection. Counting (left) and timing applications (right), after [17].

(FF). These efficiency factors can be again be summarized in a common efficiency magnitude called PDE.

2.2.3 Avalanche Breakdown Evaluation

A generated avalanche breakdown has to be detected electrically. Either the number of generated avalanches or their occurrence time can be evaluated.

The direct approach is to resolve the analog current pulses formed by the quenching circuit. This can be done by connecting the SPAD with a sampling oscilloscope to resolve timing or number of arrivals or using a Time-to-Amplitude Converter (TAC). A simple example implementation of a TAC is using a voltage ramp which is sampled by a capacitor whenever an avalanche occurs [15]. These techniques are capable of achieving higher accuracy than digital evaluation schemes. While this is a feasible approach, digitizing the signal right at the SPAD facilitates high-speed readout and scaling to big detector arrays [16]. On the digital evaluation side, two schemes are typical, as shown in Figure 2.6. Resolving the number of generated pulses per time interval via a counter forms a value proportional to the mean rate value at its output. This technique is called Single Photon Counting (SPC) [18] as the arrival count of single photons is recorded repeatedly. This scheme can acquire intensity-based images, similar to digital cameras. Optical filters like Bayer filters can be used to retrieve colored images.

The other possibility is digitizing the time of occurrence of the avalanche. That is called Time-Correlated Single Photon Counting (TCSPC) and can be realized using a Time-to-Digital Converter (TDC) structure. In TDCs, time information is stored digitally. Widespread implementations use tapped or vernier delay lines,

for which an example can be found in [19]. A 'high' bit is shifted through the delay line over time so that its position signals the current time. As soon as an avalanche is detected, the delay line is stopped. The state of the line is equivalent to the time of occurrence of the avalanche. The coded time information can then be read out as a timestamp. Local clock generation with a Phase-Locked Loop (PLL) or Delay-Locked Loop (DLL) structure enables the TDC to run at way higher frequencies than the rest of the chip. With different area requirements, clock speeds, calibration schemes, and TDC types, a precision of few picoseconds was achieved [20] even for array implementations. This matches the excellent timing behavior of SPAD detectors.

2.2.4 Noise Sources

In SPADs, noise sources can be modeled with regards to their influence on counting statistics. On the one hand, thermally generated carriers can trigger avalanche breakdowns, just as carriers related to photon absorption. On the other hand, carriers can either remain in the SCR of the diode itself or be injected from neighboring diodes via electrical or optical means, triggering avalanches when biasing voltage is ramped up again. Also, the randomness of the arrival process of photons manifests itself as speckle and photon shot noise and provides a limit on detectability. In the following section, effects contributing to SPAD noise are discussed. Later in Section 2.4, their influence on the counting statistics of the device is explained.

Dark Count Rate

Dark counts are the primary non-photon-correlated noise effect in SPADs. These are avalanche breakdowns that occur without a photo-generated electron-hole pair. A measure for dark counts is usually given by the Dark Count Rate (DCR), which has a unit of counts per second and unit area at a given temperature.

There are different causes for dark counts. Thermally generated carriers in the diode's SCR constitute the main contribution, similar to dark current in regular photodiodes [21]. They are caused by the trap-assisted generation of electron-hole pairs [22]. This effect shows a strong temperature dependence and can be controlled via cooling the device. Another less temperature-dependent

effect contributing to DCR is Band-to-Band Tunneling (BBT) of carriers from the surrounding substrate. At room temperature, SCR dominates over BBT. However, BBT can become the dominating effect at low temperatures and limit the additional performance increase through cooling the device. [14]

An effect exclusive to detectors with an internal gain is DCR dependence on biasing voltage. Increasing biasing voltage also increases the likelihood of an avalanche occurring in general, raising the probability that a thermally generated carrier will lead to an avalanche. Dark counts are rare, independent events and can thus be modeled with a Poisson distribution. The DCR constitutes the lower detection limit for SPAD detectors.

Afterpulsing

During an avalanche breakdown, many carriers pass the space charge region, a few of which remain the trapping centers of the semiconductor. These trapped carriers leave their trapped state over time with a certain probability and can cause another, non-photon-correlated avalanche. This effect is called afterpulsing. Carriers leaving their trapped state during the reset phase of an active quenching circuit cannot cause an afterpulse. Afterpulsing can thus be controlled using a long enough dead time t_{D}.

Unfortunately, the de-trapping time scales with device temperature and is longer at lower temperatures, also necessitating longer hold-off times and thus limiting the maximum count rate of the SPAD [23]. A typical magnitude for afterpulsing in modern devices is tens of nanoseconds. During this work, devices with long hold-off times are used which limit afterpulsing probability to $P_{\mathrm{AP}} \leqslant 1\%$. Additionally, measurements ended after first photon detection suffer from afterpulsing to a much lesser extent. The effects of afterpulsing are thus excluded from modeling. If afterpulsing is in the non-negligible regime, its timing behavior is hard to describe with a unified model which holds for SPAD devices from different vendors [24], although efforts are made to extend arrival time statistics to include afterpulsing [25][26].

Crosstalk

Besides avalanche breakdowns caused by carriers trapped in the SCR of the diode itself, another source of correlated noise is the injection of carriers by the neigh-

boring diodes. These effects are summarized in a magnitude called crosstalk. Crosstalk is influenced by the isolation of different diodes on electrical and optical paths [27]. It can be reduced by integrating trenches or by separating the detectors further geometrically. The diodes measured in this work possess conservative guard ring structures, trench isolation, and relatively low fill factors, pushing crosstalk into the negligible regime $P_{\mathrm{CT}} \leqslant 1\%$.

Speckle Noise

Speckle patterns occur because of the coherent nature of laser light [28]. Extensive mathematical modeling was done by Goodman in 1976 [29]. The term speckle noise describes intensity fluctuations of the reflected laser beam caused by interference [30]. In some applications, speckle patterns can be used to gain insight into the structure of materials reflecting laser photons. Though in simple narrow-band DToF, it introduces uncertainty in measured intensity and arrival time [31]. Fortunately, speckle noise can be averaged out of the ranging result by acquiring multiple range frames in short succession because the phase value of the laser radiation is not correlated between successive frames. Using a high number of laser pulses to build histograms thus makes negligence of speckle noise a valid option and is employed in all simulations and measurements in this work.

Photon Shot Noise

The most significant contributor of noise in SPAD DToF systems is photon shot noise. It is caused by the statistical nature of photon arrivals and describes the randomness of the number of photon arrivals during a time interval. Even when it is possible to describe the expected number of photon counts, uncertainty about the actual number of photon arrivals is still present. For low triggering probability, the standard deviation of counts is the square root of the expected counts, and the Poisson distribution is a well-matching distribution [32]. For laser communications and ToF applications, photon shot noise has two implications. On the one hand, it makes sending out single photons unfeasible for reliable communication. The lower the expected number of photons, the bigger the standard deviation of the signal in comparison to its mean. On the other hand, it implies that a signal has to overcome the statistical influence of present background photons.

2.3 Photon Budget

The number of collected photons during a measurement is crucial. How many incident photons can be detected on the detector's surface depends on the system parameters, mainly the laser source, the surrounding channel in which the measurement takes place, and the target object's properties.

In the following section, a model for the optical power on the detector's surface P_L is introduced. From there, the rate parameters for signal r_L and background photons r_B are derived, which determine the output signal of the DToF measurement.

2.3.1 Active Illumination

The received laser power on the detector's surface P_L is a fraction of the peak optical output power P_T in equation (2.2).

$$P_\mathrm{L} = P_\mathrm{T} \frac{\sigma_\mathrm{C}}{A_\mathrm{illum}} \frac{A_\mathrm{Rec}}{A_\mathrm{Lam}} \eta_\mathrm{A}^2 \eta_\mathrm{sys}, \tag{2.2}$$

where σ_c is the optical cross section of the system, A_illum is the area illuminated by the laser source, A_Rec the receiving optics aperture area and A_Lam the effective half-sphere area illuminated at distance d by the Lambertian scattering. The efficiency factors η_A and η_sys summarize the one-way atmospheric and system transmission efficiencies. The impact of η_A is squared because the atmosphere is passed through on the way from and to the detector. The relationships are illustrated in Figure 2.7.

The ratio between optical cross section σ_c and illuminated target area A_illum determines how much of the laser energy in the target plane is being projected into the pixel. The value for σ_c is determined by the target distance d, the respective vertical α_V and horizontal opening angle α_H, and the target reflectivity coefficient ρ to

$$\sigma_\mathrm{c} = 4\rho d^2 \tan\left(\frac{\alpha_\mathrm{H}}{2}\right) \tan\left(\frac{\alpha_\mathrm{V}}{2}\right). \tag{2.3}$$

Optical cross section is constant for flash systems in which illumination covers an equal or bigger area than the receiving optics Field of View (FoV). If the laser transmits into a pyramid stump solid angle with respective horizontal and vertical

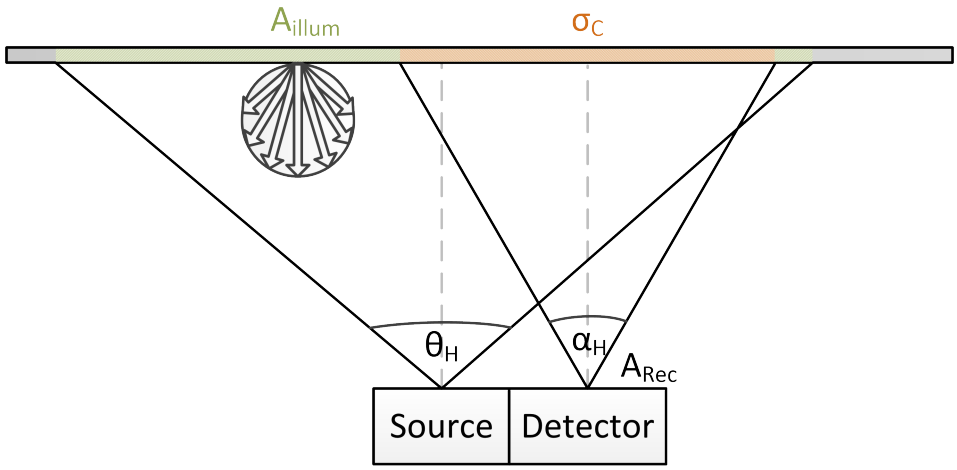


Figure 2.7: Geometrical definitions for the flash DToF setup. Example for horizontal dimension given. The vertical dimension is defined analogously.

opening angles θ_{H} and θ_{V}, the illuminated target area at distance d follows

$$A_{\mathrm{illum}} = 4d^2 \arcsin\left(\sin\left(\frac{\theta_{\mathrm{H}}}{2}\right)\sin\left(\frac{\theta_{\mathrm{V}}}{2}\right)\right). \tag{2.4}$$

The second area ratio is between the receiving optic's aperture A_{Rec} and the effective area illuminated by Lambertian reflection in the sensor plane. A circular aperture with aperture diameter D has an area of $A_{\mathrm{Rec}} = \pi\frac{D^2}{4}$. An object with Lambertian properties covers a projected solid angle of π, relating the Lambertian cosine law with a maximum reflected angle of 2π, the solid angle of the hemisphere. The illuminated area is thus $A_{\mathrm{Lam}} = \pi d^2$ [6]. From the ratio of D and target distance d, one can see that only a small fraction of emitted power P_{T} reaches the detector's aperture [33].

Using the relationship between pixel diameter d_{Pix}, respective full vertical or horizontal observation angle α and focal length $f = \tan(\alpha_{\mathrm{V}}/2) \cdot d_{\mathrm{Pix}}$, one can express the LiDAR equation for flash systems in dependence on the f-number $f_{\#} = \frac{f}{D}$ and the pixel area for square pixels $A_{\mathrm{Pix}} = d_{\mathrm{Pix}}^2$ [30].

$$P_{\mathrm{L}} = P_{\mathrm{T}}\rho\frac{1}{d^2}\frac{A_{\mathrm{Pix}}}{A_{\mathrm{illum}}}\frac{1}{4f_{\#}^2}\eta_{\mathrm{A}}^2\eta_{\mathrm{S}} \tag{2.5}$$

The efficiency factor η_{S} includes the PDP η_{PDP} of the detector and a factor modeling the geometrical efficiency of fill factor η_{FF}, model losses in the diode structure itself. Those two factors are commonly summarized as Photon De-

tection Efficiency (PDE) $\eta_{\mathrm{PDE}} = \eta_{\mathrm{PDP}}\eta_{\mathrm{FF}}$. Furthermore, a factor for loss in optics η_{Opt} can be included, which summarizes losses in detection optics and the passband attenuation of the optical bandpass filter. Together, these values are summarized as

$$\eta_{\mathrm{S}} = \eta_{\mathrm{PDP}}\eta_{\mathrm{FF}}\eta_{\mathrm{Opt}}. \tag{2.6}$$

Assuming monochromatic emission, the total rate of source photon triggered avalanches per second r_{L} is retrieved by division of P_{L} by a single photon's energy $W_{\mathrm{Ph}} = hc_0/\lambda$ at the given wavelength λ, where c_0 is the speed of light in vacuum and h the Planck constant.

$$r_{\mathrm{L}} = \frac{P_{\mathrm{L}}}{W_{\mathrm{Ph}}} = \frac{P_{\mathrm{L}}\lambda}{\mathrm{hc}_0} \tag{2.7}$$

Typical values for laser sources with emission durations in the low nanosecond regime are a few MHz at long ranges. Following the d^{-2}-dependence from equation (2.7), laser event rate at low distance can exceed values of a few GHz which is above the typically countable rate regime. The lower the laser event rate, the longer a system has to accumulate measurements for a successful evaluation. Higher event rates result in more easily evaluable signals.

2.3.2 Passive Illumination

Besides the available signal photon budget, the present ambient light is the second significant contributor to ranging performance. The most important ambient light source is the sun because of its high intensity and broadband emission spectrum.

The sun's emission can be modeled as black body radiation at temperature $T = 5800\,\mathrm{K}$. Emitted photons travel through space and are then partially absorbed by the contents of the earth's atmosphere. The resulting spectral solar irradiance on the earth E_λ is standardized in the AM1.5G spectrum and visualized in Figure 2.8. Additionally, the extraterrestrial spectrum is given for space applications. The absorption through the earth's atmosphere depends on distance traveled and thus incidence angle between sun and earth. Ranges in the spectrum which exhibit especially low irradiance are called atmospheric windows.

Two operating points at 905 nm and 1550 nm are shown, which are typical operating wavelengths λ_0 for CMOS and composite semiconductor-based DToF LiDAR systems, respectively. The operating point at 905 nm can be used by

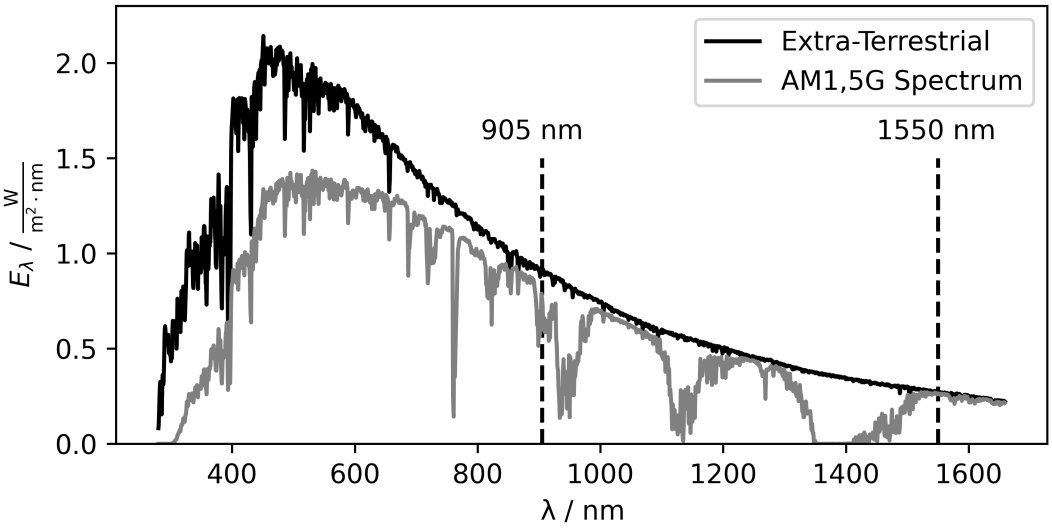


Figure 2.8: Extra-terrestrial and AM1.5G solar irradiance spectra. Operational points for 905 nm and 1550 nm are given as dashed lines. After [34].

cheaper CMOS-based systems. These systems can also be directly integrated with their readout and quenching electronics, enabling short, low-capacity connections. The low capacitance also improves SPAD timing response, compare Section 2.2.2. One can see that there is an atmospheric window at slightly higher wavelengths than 905 nm. Here, systems using a center wavelength of 940 nm operate. Still, 905 nm is often preferred over 940 nm because of the low sensitivity of silicon-based CMOS sensors for farther Near-Infrared Regime (NIR) wavelengths. Future technological improvements of CMOS NIR sensitivity could incentivize a shift to this atmospheric window. At 1550 nm, a higher optical power P_L is safe for the human eye because in contrast to 905 nm, the eye is incapable of focusing this wavelength [35]. Generally, longer wavelengths are less prone to Rayleigh scattering. The present solar irradiance for 1550 nm is lower but the system is possibly performing worse in adverse environmental conditions like fog or rain because the emission wavelength is more easily absorbed by water.

The calculation for ambient light radiant flux is similar to laser radiant flux. Integrating the AM1.5G spectrum over a wavelength range yields irradiance values for sunlight. AM1.5G is specified at an elevation angle of 48.2°. It integrates

over wavelength to a total irradiance E_{e0} of

$$E_{e0} = \int_0^\infty E_\lambda(\lambda)\mathrm{d}\lambda = 1000.4\,\mathrm{W\,m^{-2}}. \tag{2.8}$$

Integrating over an optical bandpass filter passband $\lambda = [\lambda_0 - \Delta_{BP}/2, \lambda_0 + \Delta_{BP}/2]$ and assuming normal incidence gives the irradiance in the target plane E_e which is impacting a ToF ranging system.

$$E_e = \int_{\lambda_0-\Delta\lambda_{BP}/2}^{\lambda_0+\Delta\lambda_{BP}/2} E_\lambda(\lambda)\mathrm{d}\lambda \tag{2.9}$$

Sensor requirements are often formulated using photometric quantities. The total irradiance of the AM1.5G spectrum E_e corresponds to an illuminance via integration of the photoscopic luminosity function [36]

$$E_{V0} = K \int_{360\,\mathrm{nm}}^{830\,\mathrm{nm}} E_{e0} V(\lambda)\mathrm{d}\lambda = 109\,\mathrm{klx}, \tag{2.10}$$

where $V(\lambda)$ is a weighing function, in this case representing CIE illuminance data [37]. K is the luminous efficacy, a constant by which the illuminance data is normed.

Depending on the used weighing function, slightly different values are retrieved so this approach remains an approximation. For the given photoscopic luminous efficiency data $V(\lambda)$ from CIE 1924, the luminous efficacy is found to $K = 683\,\mathrm{lm\,W^{-1}}$ and a conversion factor between Watt and Lux can be found [36]. To express how sunny it is during a DToF measurement, a fraction E_V of the total illuminance with unit Lux can then be given. Summarizing the previous considerations, the total ambient power incident on the detector surface P_A expressed via photometric quantities equates to

$$P_A = E_e \frac{E_V}{E_{V0}} A_{Pix} \frac{1}{4f_\#^2} \eta_A^2 \eta_{sys}. \tag{2.11}$$

The value of ambient rate r_A is again retrieved through the energy of a single photon W_{Ph} as

$$r_A = \frac{P_A}{W_{Ph}}. \tag{2.12}$$

Additionally to the ambient event rate r_A, the dark count rate r_{DC} affects the total

count rate. As the addition of two independent Poisson processes with constant rate parameters leads to a new Poisson process with added rate parameter, the expression for this total undesired background rate r_{B} on the pixel is

$$r_{\mathrm{B}} = r_{\mathrm{DC}} + r_{\mathrm{A}}. \tag{2.13}$$

The dynamic range for occurring ambient event rate is high. It ranges from virtually no ambient light in controlled environments or during new moon nights and exceeds rates of 100 MHz in sunlight-dominated environments. Compared with typical laser event rates, calculated after equation (2.7), one can see that they can exceed the ambient event rates by far. However, the opposite case is also possible for low reflectivity objects at high distances.

2.4 Photon Statistics and Histograms

Understanding the statistics governing SPAD arrival time distributions helps in accurately measuring distance. This section introduces a statistical model for photon counting. Also, common laser pulse distributions and histogram building are discussed.

Triggered SPADs create integer-valued, countable sums of 'events' in an observation time interval. Commonly, discrete random variables can be used to model counting processes. Especially for rare events, the Poisson process is one of the most important counting processes.

2.4.1 Arrival Process

In the Poisson process, the probability for an occurrence of exactly k events during time interval t is following the Poisson distribution in equation (2.14)

$$P_{\mathrm{P}}(k) = \frac{\mu_{\mathrm{P}}^k}{k!} \exp(-\mu_{\mathrm{P}}), \tag{2.14}$$

where μ_{P} is the mean number of arrivals. For time-dependent signals $r(t)$ it corresponds the integral over time $m(t)$

$$m(t) = \int_0^t r(\tau)\mathrm{d}\tau \tag{2.15}$$

When the measurement aims is to determine ToF, the time between subsequent events is of great interest. The arrival distribution of the next photon is called inter-arrival time and can be derived from the Poisson distribution using the probability $P(k = 0)$ to detect zero photons during a time interval. The probabiity to detect the next photon during the time is then the complementary probability.

$$P(T_1 \leqslant t) = P(k \geqslant 1) = 1 - P(k = 0) = 1 - \exp(-m(t)) \tag{2.16}$$

For a homogeneous Poisson process, the inter-arrival times are thus exponentially distributed.

2.4.2 Photon Counting

In low photon counting rate scenarios, SPADs behave like linear counters. After each detection, a dead time has to pass before the next detection. Photon arrivals are independent so, besides being shifted by a dead time, the counting distribution of the 'next' photon is not influenced by the previous detection. Since dead times of active quenching circuits are in the nanosecond scale, 'counting' photons with high incidence rates is possible even using a single diode for detection. Besides directly recording intensity images, measured ambient light intensity can also be useful as an additional input for controlling sensor properties and evaluating timed measurements.

The average combined inter-arrival time $\overline{t_\mathrm{C}}$ at constant illumination assuming non-paralyzable dead time thus follows

$$\overline{t_\mathrm{C}} = t_\mathrm{D} + \frac{1}{r_\mathrm{B}}. \tag{2.17}$$

The resulting count rate follows from the inverse of the average inter-arrival time as

$$r_\mathrm{Count} = r_\mathrm{DC} + \frac{1}{\overline{t_\mathrm{C}}} = r_\mathrm{DC} + \frac{r_\mathrm{B}}{1 + t_\mathrm{D} r_\mathrm{B}}. \tag{2.18}$$

This expression can be used to calculate r_B in high flux scenarios where a substantial amount of detections are missed due to dead time [38] to

$$r_\mathrm{B} = \frac{r_\mathrm{Count}}{1 - t_\mathrm{D} r_\mathrm{Count}}. \tag{2.19}$$

At high flux, r_{Count} converges to the inverse of t_{D}. This introduces a big uncertainty to the measurement when calculating r_{B} from r_{Count}. The authors of [39] suggest recording intensity values using timing based measurement schemes in addition to arrival counts to improve background light estimation. The necessity to record pulse-based timing information in this work requires ambient light to be limited, so the more simple dead time compensation method is used to capture intensity information. A visualization of the output counting range for exemplary values is given in Figure 2.9. The output in Figure 2.9 a) is like a linear detector between its limits constituted by DCR r_{DC} on the lower end and dead time t_{D} on the upper end. At these limits and without further processing, the detector is introducing an error of 100 %. In orange, the behavior of an ideal linear detector is shown. The error count rates of the real detector equal

$$r_{\mathrm{Count}} \overset{r_{\mathrm{B}}=r_{\mathrm{DC}}}{=} 2 \cdot r_{\mathrm{DC}} \tag{2.20}$$

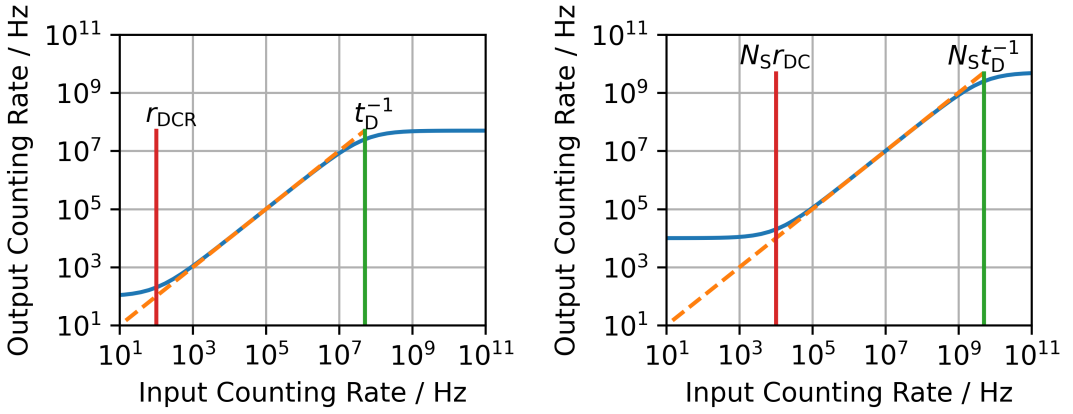
for an input rate equal to DCR and

$$r_{\mathrm{Count}} \overset{r_{\mathrm{B}}=t_{\mathrm{D}}^{-1}}{=} r_{\mathrm{DC}} + \frac{1}{2t_{\mathrm{D}}} \overset{t_{\mathrm{D}}^{-1} \gg r_{\mathrm{DC}}}{\approx} \frac{1}{2} t_{\mathrm{D}}^{-1} \tag{2.21}$$

for an input rate equal to the inverse of the dead time. In Figure 2.9 b), the output counting rates of multiple detectors are shown using a mutual counting structure, the upper count can be extended above the output count rate limit of t_{D}^{-1} to $N_{\mathrm{S}} t_{\mathrm{D}}^{-1}$. In summary, SPAD detectors counts like linear detectors over a large range of input signals. In low input regimes, counting output is limited by DCR and in high input regimes it is limited by dead time. Incident photons could trigger an avalanche, but the diode being in dead time prevents that, lowering count value. This lowered count rate can partially be compensated for, depending on whether the dead time is paralyzable or non-paralyzable. Counting incoming photons yields an intensity image of the recorded frame. The typically low DCR makes SPADs detectors well fit to record images even at very low intensities.

2.4.3 Photon Timing

The exponential inter-arrival time in equation (2.16) and detector dead time imply that incident photons rapidly saturate the sensor. This saturation limits



(a) Single SPAD.

(b) Multi SPAD pixel with $N_S = 100$ diodes.

Figure 2.9: Counting rates depending on noise and dead time. Graph generated with $r_{DC} = 100\,\text{Hz}$ and $t_D = 20\,\text{ns}$.

the range of SPAD DToF systems when ambient light is present.

First-Photon Detection

Historically, SPAD-based detectors were built with passive quenching circuits. To suppress afterpulsing efficiently, dead times spanning multiple microseconds were necessary. When ending a measurement cycle after detecting the first incoming photon, detector dead time requirements are relaxed. It is sufficient for the detector to be reset when the next laser pulse is emitted. This time, determined by the inverse of the laser's PRF, is usually in the microsecond regime.

First-photon detection produces distributions highly influenced by the magnitude of intensity impinging on the sensor. Arrival times are distributed according to equation (2.16). The derivative of this distribution function with respect to time leads to the Probability Density Function (PDF) of first-photon arrival, or strictly speaking of arrival of one or more photons, which depends both on instantaneous rate $r(t)$ and mean number of arrivals $m(t)$.

$$f_S(k \geqslant 1) = \frac{\mathrm{d}}{\mathrm{d}t} P(k \geqslant 1) = r(t) \exp(-m(t)) \tag{2.22}$$

With equation (2.22), first-photon distributions of arbitrary time-dependent rate functions can be predicted. In the following, two pulse models are presented.

The simplest temporal laser pulse form approximation is a rectangular pulse.

The rate function for a rectangular pulse with peak rate $r_{\mathrm{L,R}}$ and temporal width t_{P}, returning at time t_{ToF}, and superimposed onto a constant ambient illumination of rate r_{B}, is defined in equation (2.23).

$$r_{\mathrm{R}}(t) = r_{\mathrm{B}} + r_{\mathrm{L,R}}[\Theta(t - t_{\mathrm{ToF}}) - \Theta(t - (t_{\mathrm{ToF}} + t_{\mathrm{P}}))] \tag{2.23}$$

Inserting this into equation (2.22) yields the first-photon PDF for a rectangular pulse with ambient light present.

$$\begin{aligned} f_{\mathrm{R}}(t) = {} & r_{\mathrm{B}} \exp(-r_{\mathrm{B}} t)[\Theta(t) - \Theta(t - t_{\mathrm{ToF}})] \\ & + (r_{\mathrm{L}} + r_{\mathrm{B}}) \exp(-(r_{\mathrm{L}} + r_{\mathrm{B}})t + r_{\mathrm{L}} t_{\mathrm{ToF}})[\Theta(t - t_{\mathrm{ToF}}) - \Theta(t - (t_{\mathrm{ToF}} + t_{\mathrm{P}}))] \\ & + r_{\mathrm{B}} \exp(-r_{\mathrm{B}} t - r_{\mathrm{L}} t_{\mathrm{P}}) \cdot \Theta(t - (t_{\mathrm{ToF}} + t_{\mathrm{P}})) \end{aligned} \tag{2.24}$$

The rectangular pulse model is a simplification. It is well-suited to model the mean photon count impinging on the sensor but exhibits infinitely short rise times. Also, the theoretical maximum of the pulse PDF in equation (2.24) for $r_{\mathrm{B}} = 0$ is always at time t_{ToF}.

The real temporal pulse form of lasers is often modeled with a Gaussian distribution. Depending on the physical properties of the laser, the intensity distribution can match more closely with a hyperbolic sechans function (called sech2 in the following) in equation (2.25). The sech2 pulse profile can be assumed by mode-locked lasers, especially with passive mode-locking through saturable absorbers [40].

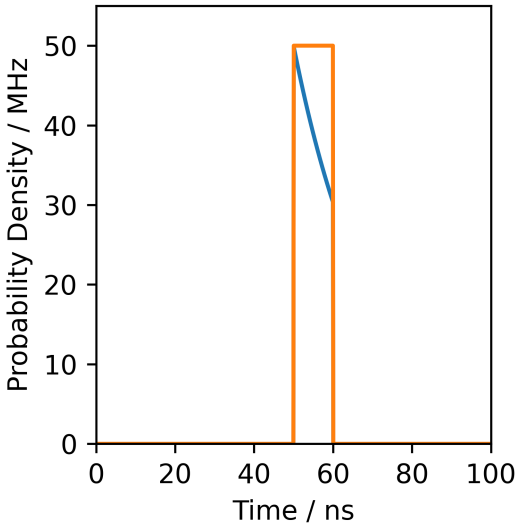
$$r_{\mathrm{Sech2}}(t) = r_{\mathrm{B}} + \frac{r_{\mathrm{L},0}}{\cosh^2\left(\frac{t - t_{\mathrm{ToF}}}{\tau_{\mathrm{P}}}\right)} \tag{2.25}$$

The temporal pulse width is scaled with the parameter τ_{P}. Again combining with equation (2.22) gives the first photon PDF for sech2-distributed impulses as

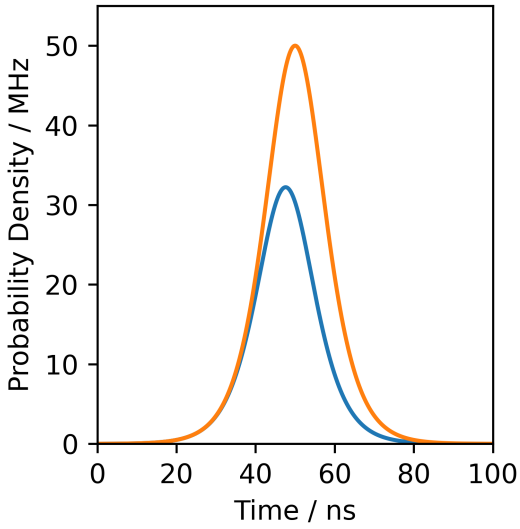
$$f_{\mathrm{Sech2}}(t) = r_{\mathrm{Sech2}}(t) \exp\left(-r_{\mathrm{B}} t - r_{\mathrm{L},0} \tau_{\mathrm{P}} \left[\tanh\left(\frac{t - t_{\mathrm{ToF}}}{\tau_{\mathrm{P}}}\right) + \tanh\left(\frac{t_{\mathrm{ToF}}}{\tau_{\mathrm{P}}}\right)\right]\right). \tag{2.26}$$

It provides a better understanding of the temporal signature, especially the laser pulse's slopes, and not only the mean number of generated photons.

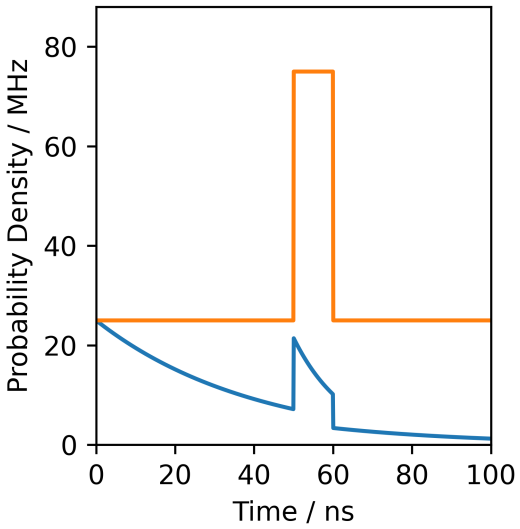
An example for such distributions following equations 2.24 and 2.26 is given in Figure 2.10. The difference between intensity function and PDF for both pulse models is clearly visible. With rising event rate, the maximum of the PDF is



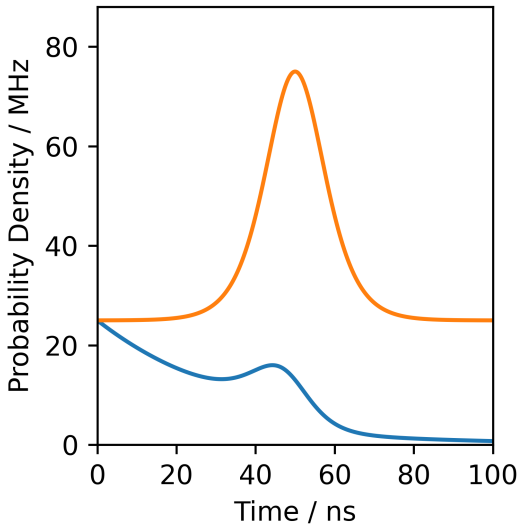
(a) Rectangular pulse model.



(b) Sech2 pulse model.



(c) Rectangular pulse model with $r_\mathrm{B} = 25\,\mathrm{MHz}$.



(d) Sech2 pulse model with $r_\mathrm{B} = 25\,\mathrm{MHz}$.

Figure 2.10: Intensity (orange) and PDF (blue) for different temporal laser pulse profiles. $t_\mathrm{ToF} = 50\,\mathrm{ns}$, $t_\mathrm{P} = \tau_\mathrm{P} = 10\,\mathrm{ns}$ and $r_\mathrm{L} = 50\,\mathrm{MHz}$.

shifted to earlier times. This effect is emphasized under the influence of ambient light. Under ambient influence, the maximum value is assumed at the beginning of the PDF. Using a pure peak detector would derive the wrong ToF value and further processing is required to estimate the right value.

Binning

A histogram is a scaled, noisy realization of the arrival processes' PDF. It is built by measuring arrival times for a number of laser pulses and counting all arrival times during the same discretization step or bin. The continuous values of the PDF are discretized by the finite resolution of the TDC. The amount of noise depends on the number of samples and is discussed in more detail in Section 3.1. Here, the connection between timing and counting statistics in SPADs is shown. If ambient light is present, multiple measurements can be accumulated in a histogram and analyzed to differentiate signal and noise contributions. Arrivals of photons stemming from ambient light are randomly distributed over the histogram, while arrivals from the laser source accumulate during the laser pulse return. The total distribution of arrivals is finally processed to derive ToF.

Integration of the arrival time PDF $f_{\mathrm{S}}(t)$ over time leads to the Cumulative Distribution Function (CDF) of the distribution $F_{\mathrm{S}}(t)$ as

$$F_{\mathrm{S}}(t) = \int_0^t f_{\mathrm{S}}(\tau)\mathrm{d}\tau. \tag{2.27}$$

It will be used to determine the triggering probabilities for each bin p_i by evaluating $F_{\mathrm{S}}(t)$ over the histogram $t_{\mathrm{i}} = [t_{\mathrm{Bin}}, 2 \cdot t_{\mathrm{Bin}}, ..., n_{\mathrm{Bin}} \cdot t_{\mathrm{Bin}}]$ as expressed in equation (2.28).

$$p_i = \int_{i \cdot t_{\mathrm{Bin}}}^{(i+1) \cdot t_{\mathrm{Bin}}} f_{\mathrm{S}}(t)\mathrm{d}t = F_{\mathrm{S}}((i+1) \cdot t_{\mathrm{Bin}}) - F_{\mathrm{S}}(i \cdot t_{\mathrm{Bin}}) \tag{2.28}$$

This problem can thus be modeled as a Bernoulli trial and with probability p_i, the value of a single ToF measurement lies within the bin and with probability $1 - p_i$ it falls into a different bin or outside the histogram. This measurement is repeated a total of N times. The resulting count value follows a binomial distribution with probability $P_{i,\mathrm{B}}$ for k out of N detections in equation (2.29).

$$P_{i,\mathrm{B}} = \binom{N}{k} p_i^k (1-p_i)^{N-k} \tag{2.29}$$

The expected value $\mu_{i,\mathrm{B}}$ and standard deviation $\sigma_{i,\mathrm{B}}$ of bin count values are then

$$\mu_{i,\mathrm{B}} = \mathrm{E}(P_{i,\mathrm{B}}) = Np_i \tag{2.30}$$

and

$$\sigma_{i,\mathrm{B}}^2 = \mathrm{Var}_{i,\mathrm{B}} = Np_i(1-p_i). \tag{2.31}$$

For ease of modeling, this distribution will be approximated for low p ($\ll 10\%$) by the simpler Poisson distribution in equation (2.14) with an expected value of

$$\mu_{i,\mathrm{P}} = Np_i \tag{2.32}$$

and variance of

$$\sigma_{i,\mathrm{P}}^2 = \mathrm{Var}_{i,\mathrm{P}} = Np_i. \tag{2.33}$$

The histogram H is then a sequence of realizations n_i of this random process with an expected value $\mu_{i,\mathrm{P}}$

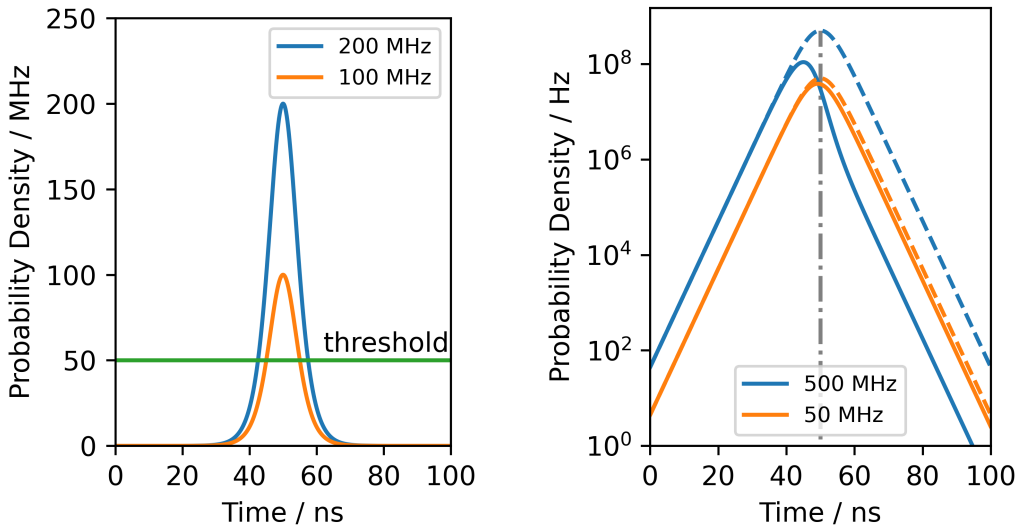
$$H = \{n_1, n_2, ..., n_{\mathrm{N}}\}. \tag{2.34}$$

As PDFs are always positive, increasing values for TDC quantization time t_{Bin} also lead to higher mean count values. Standard deviation also increases with the root of count value but evaluability improves overall. When using a peak detection algorithm and without further processing, t_{Bin} constitutes the time resolution of the system. Increasing t_{Bin} will consequently also decrease time resolution.

For example one can consider the limiting case with only a single time slot. It will always catch the maximum number of incoming signal photons, but time resolution is not present anymore and signal detection is limited to 'signal was present during measurement' or 'no return detected'. In low signal regimes, an operating point between evaluability and time resolution has to be found.

2.5 Intensity-dependent Histogram Signatures

Besides the obvious case that the returned optical power is so low that no photons were recorded after several laser shots, there are also distortions of the histogram



(a) Intensity-dependent range-walk for peak rates $r_{\text{Sech2}} = 100\,\text{MHz}$ (orange) and $r_{\text{Sech2}} = 200\,\text{MHz}$ (blue).

(b) Visualization of pile-up present in dead time affected detectors. Intensity function and PDF shown for different peak intensities.

Figure 2.11: Intensity-dependent ranging errors: range-walk and pile-up.

signature caused by high signal intensity.

2.5.1 Range-Walk

A signal rate-dependent effect also present in analog detectors is called range-walk and is shown in Figure 2.11 a).

In essence, the determination of ToF with a fixed threshold leads to a shift of measured ToF to earlier times for higher values of optical return power. While the temporal pulse form of incident light stays the same, an increase in intensity leads to crossing a constant threshold at different times. In analog detectors, this effect can be mitigated using adaptive thresholds like in constant fraction discrimination [41].

2.5.2 Pile-Up

Besides range-walk, detectors with a characteristic dead time after each detection show an additional intensity-dependent behavior called pile-up. It describes the overemphasis of photons arriving at the leading edge of the returned laser pulse because of first-photon detection. The higher the intensity, the higher the

probability of detecting earlier photons because of their shorter inter-arrival time. High intensity also shifts the maximum position of the PDF in comparison to the intensity function. Figure 2.11 b) shows this effect for two different event rates. The position of the pulse maximum shifts to earlier times for higher rates. One problem of assuming pulses of infinite duration can be seen on a logarithmic scale. Even way before the actual pulse return, non-zero PDF values are assumed.

Different study groups have tackled the problem of pile-up correction from measurement data. An early but still influential contribution is the estimator published in [42] that corrects the pile-up in the histogram by correcting the reduced count values inside the other histogram bins. Recent contributions extending this approach and aiming to reduce the noise emphasis in later bins is found in [43] and [44]. A different approach is manipulating the acquisition of photon arrivals so that the pile-up effect is attenuated beforehand and does not affect the histogram. This work takes a deeper look at the signal-processing based correction on pile-up affected histograms in Section 5.5. The pile-up effect is only detrimental when determining ToF from a wide range of possible laser return rates because it constitutes a rate-dependent shift. Determining t_{ToF} from a more narrow distribution generally means that the distance variance is smaller because deviation of possible outcomes is also smaller, leading to higher ranging accuracy as shown in [45]. During the course of this work, a contribution regarding the influence of bin size and laser pulse rate has been published [46].

2.5.3 Quantization Errors

When the returned laser power is so high that inter-arrival times are smaller than a bin size t_{Bin} and the temporal Instrument Response Function (IRF) of the system is narrow enough to prevent spreading arrival times into different bins, all generated events can fall into a single time bin. In this case, interpolation or center-of-mass evaluation schemes can no longer increase depth resolution using the bin values of neighboring bins. This means an effective loss of depth resolution when high return power is present.

To suppress quantization errors, the system can thus be operated in power regimes where the quantization effect is not dominant. If that is not possible, the arrival times can be artificially delayed in time to fall into different time slots, which is called dithering. An implementation as subtractive dithering is

found in [47]. It should be emphasized that we are looking at two different kinds of variance: variance of count value and variance of arrival time. This can also be made plausible by looking at the variance of arrival time introduced by binning. As it represents a form of quantization where all outcome states have the same probability, it can be modeled as a uniform distribution with standard deviation σ_t.

$$\sigma_\mathrm{t} = \frac{1}{\sqrt{12}} t_\mathrm{Bin}. \tag{2.35}$$

Increasing bin size thus is advantageous for reliably finding the signal but decreases timing accuracy and an application-dependent trade-off has to be found. One example of using this to the user's advantage is in two-step histogramming. First, pulse position is roughly identified. Then, this region of interest is recorded using a finer quantization. As bin size is also proportional to generated data output, this approach also reduces data throughput.

2.6 Ambient Mitigation

With the high sensitivity of SPADs, detecting smaller return intensities stemming from objects far away or with low reflectivity is possible. Not only detecting signal photons but also detecting undesired ambient photons is getting more probable. There are many approaches to mitigate ambient light influence, a few of which are discussed in this section. The focus lies on methods implementable on a digital circuit level.

2.6.1 Multi-Event Detection

Detectors with first-photon detection are 'blind' after detecting the first incident photon or event. This makes their sensitivity drop off exponentially caused by the inter-arrival time of background photons following equation (2.16). By re-enabling the device after detection of the first photon, sensitivity deterioration can be reduced even if dead time has to pass after each arrival.

If the sensor capabilities allow for the detection of multiple incoming photons per measurement cycle [48], their respective arrival time distributions are of interest. The arrival time PDF for multi-photon detection follows an Erlang

distribution. It can be derived using the fact that the detection probability of the k-th photon depends on the detection of $k-1$ photons up to the time t. The probability of seeing a wait time smaller than t for the k-th photon S_k describes the same condition that there were more or equal than k detections up to point t [49] and is

$$F_S(t) = P\{S_k \leqslant t\} = P\{N(t) \geqslant k\} = \sum_{j=k}^{\infty} \exp(-m(t)) \frac{m(t)^j}{j!}. \tag{2.36}$$

The corresponding probability density function can then be determined via differentiation to

$$f_S(t) = r(t) \exp(-m(t)) \frac{m(t)^{k-1}}{(k-1)!}. \tag{2.37}$$

Detector dead time is another influence on the arrival time of subsequent events. The detection of the second photon depends on the first-photon detection and the dead time, which has to pass after each first-photon detection so that the detector is sensitive again. The detection mechanism differs from that of a purely Poisson process because its events depend not only on the observation interval t but also on the history of the detection process.

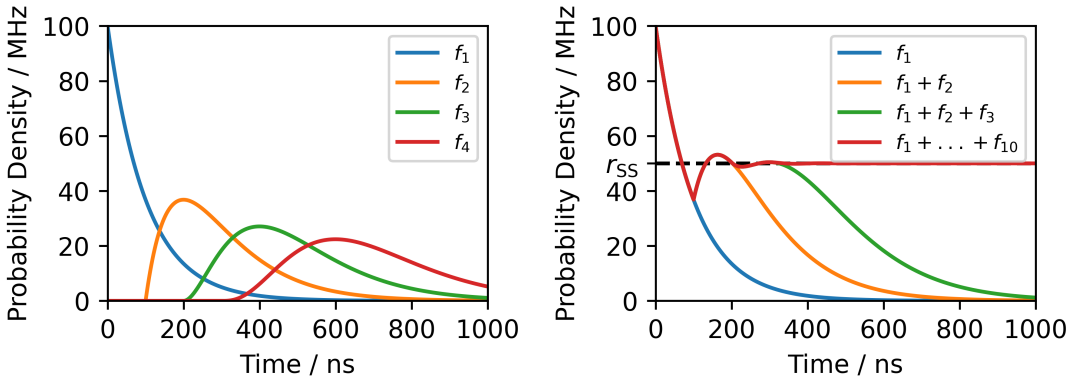
The authors of [50] give an expression for the dead time dependency of the distribution. It stems from signal theory that systems with a present dead time t_D have their PDF shifted by t_D. The PDF $f_{\mathrm{k,IHP}}$ of respective single event levels k of an Inhomogeneous Poisson Process (IHP) can be calculated recursively following

$$f_{\mathrm{k,IHP}}(t) = \Theta(t-(k-1)t_\mathrm{D}) r(t) \int_{kt_\mathrm{D}}^{t-t_\mathrm{D}} f_{\mathrm{k-1}}(\tau) \exp\left(-\int_{\tau+t_\mathrm{D}}^{t} r(t')\mathrm{d}t'\right) \mathrm{d}\tau \tag{2.38}$$

This can be simplified for a constant rate $r(t) = r$ and forms a Homogeneous Poisson Process (HP). The result is shown in equation (2.39) and will be used for modeling the detection process in Section 3.2. How well the distribution fits real measurement data is assessed in Chapter 6. A visualization of the theoretic distribution is given in Figure 2.12.

$$f_{\mathrm{k,HP}}(t) = \Theta(t-(k-1)t_\mathrm{D}) \frac{r^k (t-(k-1)t_\mathrm{D})^{k-1}}{(k-1)!} \exp(-r(t-(k-1)t_\mathrm{D})) \tag{2.39}$$

In completely free-running systems, a steady-state rate r_SS will be reached under



(a) Single-photon arrival distributions. (b) Accumulated photon arrival distributions.

Figure 2.12: Multi-Photon PDFs f_i for constant event rate $r_\mathrm{B} = 100\,\mathrm{MHz}$, dead time $t_\mathrm{D} = 100\,\mathrm{ns}$ and different event levels. The more event levels are implemented, the longer sensitivity can be upheld over time.

constant illumination after dead time effects of the first few events have settled. Its dependence on dead time t_D and actual incoming photon rate r follows

$$r_\mathrm{SS} = \frac{r}{1 + r t_\mathrm{D}}. \tag{2.40}$$

This value is indicated in Figure 2.12 for ten accumulated photon arrivals. It is approximately the same rate that is assumed under constant ambient light illumination in equation (2.18). For systems with finite multi-event resolving capabilities, this steady-state will only be present for parts of the resulting histogram. Sensitivity for incoming photons will eventually drop off over the histogram. The next detection can only occur whenever a dead time t_D has passed. If an observation interval of time t_Hist is considered, the maximum number of possible detections N_Ph is following

$$N_\mathrm{Ph} = \left\lceil \frac{t_\mathrm{Hist}}{t_\mathrm{D}} \right\rceil. \tag{2.41}$$

This is also the limiting number of detections possible in a free-running implementation. The number of implemented multi-event levels will be lower than that because of the hardware requirements and the fact that it is a limiting case at infinitely high event rates.

Considering a DToF system that should cover a range of about 200 m, the value of t_Hist should cover at least 1335 ns following equation (2.1). Modern active

quenching circuitry makes low dead times with low afterpulsing possible. With an example dead time in the nanosecond regime e.g., 50 ns (it depends on technology and even lower ones were reported), up to $N_{\mathrm{Ph}} = 27$ events could be recorded. Comparing this to microsecond dead times of passive quenching circuitry shows that passive quenching systems are limited to first-photon detection in DToF applications for meters rather than kilometers. In all cases, using the value for N_{Ph} and the PDF for constant rates in equation (2.39), one can find the arming probability P_{A} at time t as

$$P_{\mathrm{A}}(t) = \sum_{j=1}^{N_{\mathrm{Ph}}} \frac{f_{\mathrm{j,HP}}(t)}{r}. \tag{2.42}$$

The model assumes the detector to be unblocked at the start of the measurement and not influence the distribution of the first detected photon. Suppose however, that assumption is not correct and there is a possibility that the detector is in dead time. In that case, the first-photon arrival time PDF will either be completely shifted or reduced in magnitude by the probability to detect before the start of the measurement. The longer the dead time of the system, the worse the mean sensitivity to incoming arrivals.

2.6.2 Time Gating

The gate defines the active phase of the SPAD. It is the time during which external biasing voltage is high enough to put the SPAD into Geiger-mode.

As described previously, one limiting factor for DToF systems is saturation through early background photons. A way to prevent these early photons from deteriorating the signal is to open the gate just before the signal photons arrive. Choosing the right activation time is crucial because opening the gate too late will exclude the target return signal from the detection histogram.

The choice of this time instance is not trivial for DToF because the ideal time is the ToF which is also the quantity to be measured. It is more readily applicable to measurements where a rough idea of the signal return is present, like in Fluorescence Lifetime Imaging (FLIM) or Raman spectroscopy [51]. Rough estimation for ToF could be delivered by sensor fusion with different measurement devices, possibly with lower depth resolution.

Another possibility is using target distance agnostic time gating schemes e.g., [52]. Their drawback is that they make histogram analysis more complex and are hard to combine with other ambient mitigation strategies evaluated in this work and are thus not discussed further.

2.6.3 Coincidence Detection

A major issue of DToF systems is that each photon produces a high-resolution timestamp which constitutes multiple bit of data. This information is generated per SPAD whenever a photon is incident and triggers an avalanche. These timestamps have to be read out, leading to large amounts of data, especially in big detector arrays.

Counting coincidence detections means only triggering the generation of timing information when multiple detectors are triggered simultaneously. Photons never truly arrive at the same point in time. Also, SPAD detectors exhibit certain conversion times and uncertainties in time measurement. Coincidence detection in SPAD arrays is thus implemented using finite resolving times called coincidence time t_{C} and a number of required triggered SPADs n_{C} called coincidence depth. Ambient photons lead to constant intensity over the acquisition time. For typical background event rates, arrival times are distributed sparsely. The laser photons arrive temporally correlated. Coincidence detection can reduce the number of timestamps from ambient photons, using a temporal and spatial correlation between multiple detectors [53]. In some conditions, the ratio between unwanted ambient photon detections and signal contributing laser photon detections is improved [54]. Using an acquisition frame roughly matched to the temporal width of the laser pulse or even shorter enables the use of this difference in arrival time densities.

The attenuation of coincidence detection can be set via n_{C} and t_{D}. A more detailed mathematical model is formulated in Section 3.1.3. Hit probabilities in DToF LiDAR are often generally kept low, so having multiple detectors fire to generate a timestamp can lead to only two or three counts accumulating in the same range bin after sending hundreds of laser pulses [6]. Still, coincidence detection is an effective way of suppressing ambient light, provided the necessary laser power for strong attenuation is available.

Chapter 3

Pixel Structures

The following chapter deals with pixels built from multiple SPADs. Multiple SPADs per pixel can circumvent a single device's dead time behavior and enable coincidence detection. Also, technological advances open the possibility for implementing smaller SPAD diodes. Pixel miniaturization is a decade long trend in the CMOS Imaging Sensor (CIS) industry. There are several factors that make a decrease in pixel size attractive. Spatial resolution is increased by dividing the FoV up into smaller sub-areas, but chip cost decreases because the same number of smaller pixels will occupy less area. The smaller pixel diameter can also increase temporal resolution because smaller distances must be bridged. DCR can also be reduced as smaller diodes are less prone to contain semiconductor impurities. The downside on a device level is that smaller pixels and smaller pixel pitches emphasize electrical and optical interference between neighboring pixels. On a circuit level, the quenching circuitry and TDC use roughly the same space. The area of readout circuitry required is the same for a small pixel as for a large pixel. In FSI-implementations and as diode diameter decreases and the area of pixel-individual readout electronics stays the same, fill factor is lowered. Also, power dissipation increases with the number of pixels.

These effects can be reduced by sharing electronics. Quenching electronics are best kept close to the SPAD to reduce stray capacitance and propagation delays. TDC structures can more easily be shared between different diodes. The limiting case is an array evaluated with a single TDC. As the timing electronics have a finite conversion time, simultaneously occurring events can be missed. With an increasing number of diodes and activity levels, this becomes more probable and

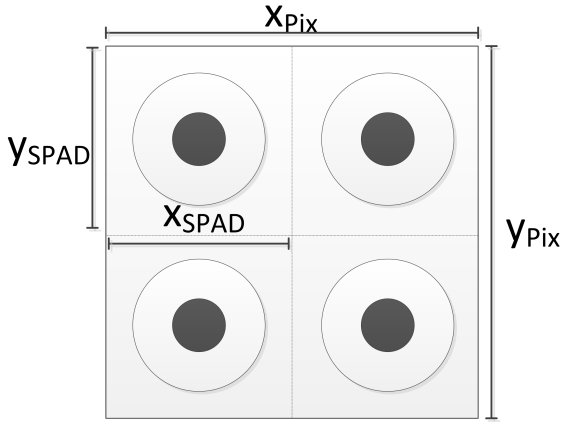


Figure 3.1: Definitions for SPAD and pixel diameters in x and y-direction. Black areas portray the optically active area. Circular structures are guard rings.

ranging performance is lower than the theoretical limit [55]. A good compromise between SPAD-individual and column or array-wide TDC sharing is using structures similar to silicon photomultiplier cells. A TDC is shared between a small number of individual cells and placed either close to the pixel in FSI implementations or right below the pixel on a companion wafer in BSI.

This chapter discusses the relationship between chosen SPAD pixel architectures, counting behavior, and expected ranging performance using Signal-to-Noise Ratio (SNR). A focus lies on Multi-Event (ME) and coincident detection.

3.1 SPAD Pixel Architectures

Figure 3.1 shows the naming conventions for pixels and diodes. The name pixel is an abbreviation for picture element and refers to an image of a point in the scene. The incident power on a pixel is translated into an event rate r which is calculated according to equation (2.7) for laser emission or equation (2.13) for ambient light. For the following examinations, the rate of photons impinging on one out of N_S SPADs is of interest which will be defined as

$$r_\mathrm{S} = \frac{r}{N_\mathrm{S}}. \tag{3.1}$$

Equation (3.1) assumes illumination to be homogeneous over the whole array and is a simplification to show the general pixel behavior. This assumption will

only hold approximately in a real system and can be replaced by a more sophisticated model for diode individual power budgets for more accurate performance predictions.

Evaluability of chosen Multi-Pixel Structures

The intensity ratio between ambient and laser is not a good indicator for SPAD performance because it neglects the non-linear relationships of ambient light saturation and generally does not cover the time-dependent behavior of the systems. Still, it finds widespread use to access system performance e.g., [56]. In this work, the quantity of SNR will be used throughout this section to compare different pixel architectures to one another and compare SPAD ranging performance to APDs. Detecting a signal means identifying its presence over ambient and noise influences. In general, the SNR k_{SN} of a detector is defined as signal mean over total variance (3.2) [57].

$$k_{\mathrm{SN}} = \frac{\mu}{\sigma} \tag{3.2}$$

For a single time bin, counting values in case of pulse return and ambient light respectively can be retrieved from the arrival CDF. Using the identities from equations (2.32) and (2.33), an expression for a laser return with Poisson mean $\mu_{\mathrm{P,L}}$ during a single bin can be found to

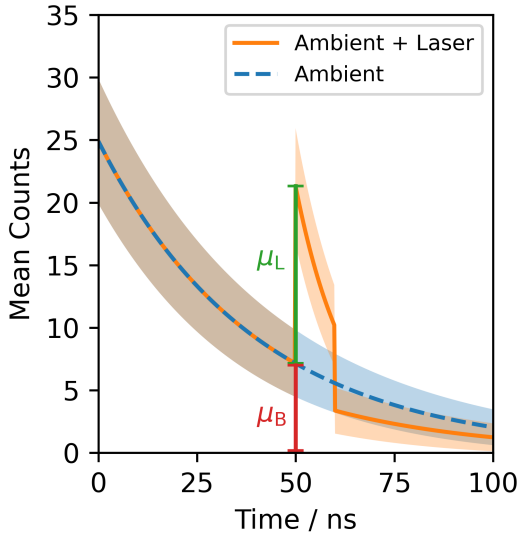
$$k_{\mathrm{SN\,H}} = \frac{\mu_{\mathrm{P,L}}}{\sigma_{\mathrm{P,L}}} = \frac{\mu_{\mathrm{P,L}}}{\sqrt{\mu_{\mathrm{P,L}}}} = \sqrt{N p_i} \tag{3.3}$$

For a mixed return of pulse and background and first-photon detection, it depends on the Poisson expected number of counts generated during laser return μ_{L} and expected ambient light count μ_{B}. The standard deviation of the total noise influence is the square sum of the standard deviation of all noise contributions.

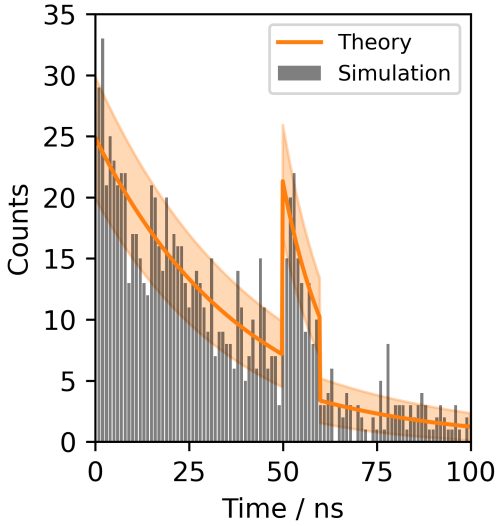
$$k_{\mathrm{SN\,H}} = \frac{\mu_{\mathrm{L}}}{\sqrt{\mu_{\mathrm{L}} + \mu_{\mathrm{B}}}} = \sqrt{N_{\mathrm{M}} P_{\mathrm{A}}(r, t_{\mathrm{ToF}}, 1)} \frac{p_{\mathrm{L}}}{\sqrt{p_{\mathrm{L}} + p_{\mathrm{B}}}} \tag{3.4}$$

Figure 3.2 shows expected value and variance for rectangular and sech2-distributed laser intensity functions. If the maximum bin SNR is of interest, the ambient level must be compared to the highest count value.

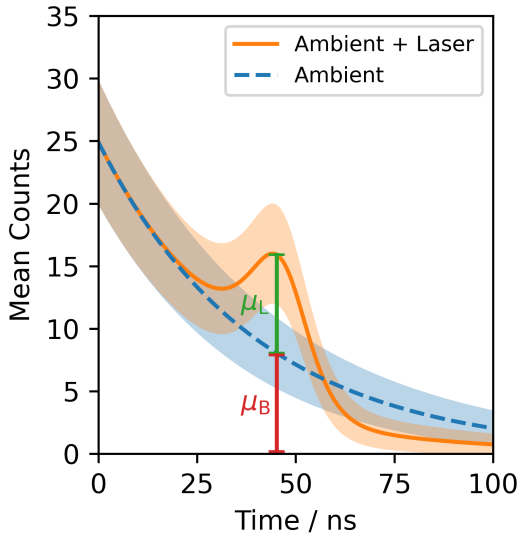
For a rectangular pulse, the maximum value of the PDF during pulse return lies at t_{ToF}. The maximum pulse-related triggering probability $p_{\mathrm{LB,R}}$ can thus be



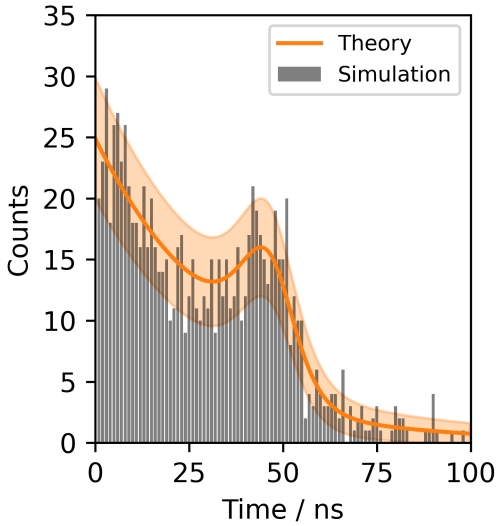
(a) Rectangular pulse, width $t_P = 10\,\text{ns}$.



(b) Rectangular pulse histogram.



(c) Sech2 pulse, width $\tau_P = 10\,\text{ns}$.



(d) Sech2 pulse histogram.

Figure 3.2: Histogram expected value μ (lines) and standard deviation σ (shaded area) for the two pulse models. $t_{ToF} = 50\,\text{ns}$, $r_L = 50\,\text{MHz}$, $r_B = 25\,\text{MHz}$, $t_{Bin} = 1\,\text{ns}$. Note that the transmitted energy of both pulses is not equal because for same signal heights, rectangular and sech2 shapes enclose different areas.

retrieved by evaluating the rectangular pulse CDF for the bin of the pulse return as

$$p_{\mathrm{LB,R}} = F_{\mathrm{S,R}}(t_{\mathrm{ToF}} + t_{\mathrm{Bin}}) - F_{\mathrm{S,R}}(t_{\mathrm{ToF}}). \tag{3.5}$$

It is an expression for the triggering probability through the combined influence of laser pulse return and ambient. If one is interested in the triggering probability increase caused by laser pulse return, as is the case when calculating SNR, expected ambient counts have to be subtracted, leading to

$$p_{\mathrm{L}} = p_{\mathrm{LB}} - p_{\mathrm{B}}. \tag{3.6}$$

The value of background triggering probability p_{B} can be calculated from the PDF under pure background influence as

$$p_{\mathrm{B}} = F_{\mathrm{S,B}}(t_{\mathrm{ToF}} + t_{\mathrm{Bin}}) - F_{\mathrm{S,B}}(t_{\mathrm{ToF}}). \tag{3.7}$$

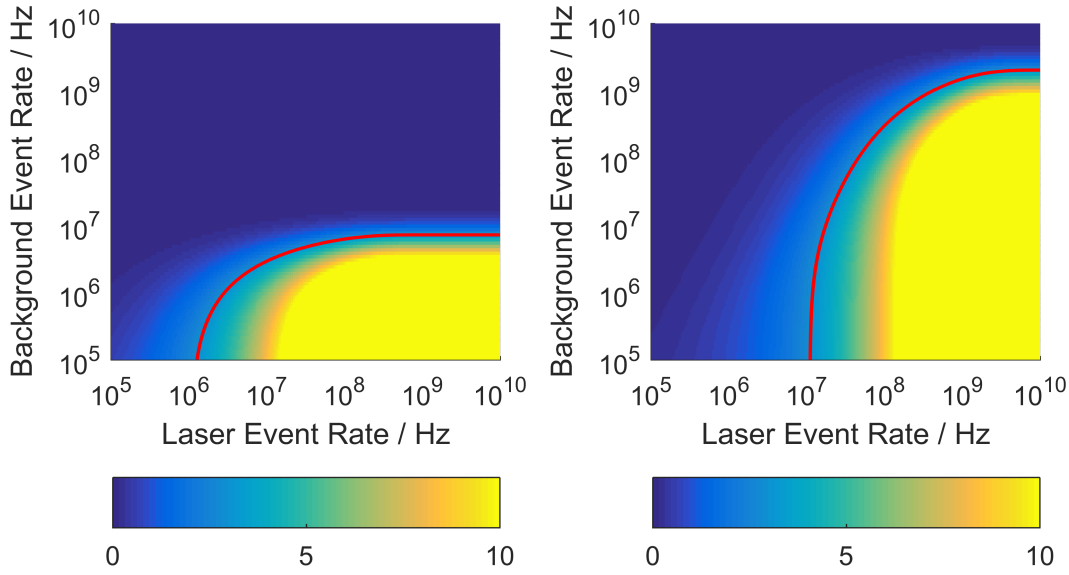
For more realistic pulse forms, the expression is more complicated because the function local maximum is not occurring during a fixed point in time but wanders depending on the absolute value of intensity. From the derivative of the sech2 PDF in equation (2.26), the local maximum at pulse return can be retrieved as

$$t_{\mathrm{Max}} = t_{\mathrm{ToF}} - \tau_{\mathrm{P}}/2 \cdot \operatorname{arcsinh}(r_{\mathrm{L},0}\tau_{\mathrm{P}}). \tag{3.8}$$

Integrating symmetrically around t_{Max} yields the approximate histogram signal contribution for the sech2 pulse form. It will later be used to correct for pile-up for sech2-distributed pulses.

$$p_{\mathrm{LB,S}} = F_{\mathrm{S,S}}(t_{\mathrm{Max}} + t_{\mathrm{Bin}}/2) - F_{\mathrm{S,S}}(t_{\mathrm{Max}} - t_{\mathrm{Bin}}/2) \tag{3.9}$$

Both pulse models will be used throughout this work. The following discussion of DToF SNR will be based on the rectangular pulse model because it is easier to handle than the more precise sech2 model. As already mentioned, this neglects the influence of the laser rising edge. The focus lies on low return power cases, where whether a pulse is still detectable is more important than the accuracy with which it can be detected. In later discussions of pile-up effects, the real



(a) Idealized SPAD with $\eta_{\mathrm{PDP}} = 1$ in first-photon mode.

(b) SPAD with $\eta_{\mathrm{PDP}} = 0.1$ in first-photon mode.

Figure 3.3: SNR contour lines over background and laser event rate. Contour for $k_{\mathrm{SN}} = 3$ is shown as a limit for evaluability.

pulse model is used.

An example of SNR of a single SPAD is shown in Figure 3.3 a). Higher values of r_{L} and lower values of r_{B} increase SNR. Signal height in the histogram increases when going further to the bottom right corner of the graph. For values greater than 10, the same color shading is chosen for good contrast. Notably, there is a value of laser event rate r_{L} under which detection becomes impossible. The same is true for values above a specific maximum background event rate r_{B}. Figure 3.3 b) shows the same system but with finite detection efficiency $\eta_{\mathrm{PDP}} = 0.1$. The same event rates are plotted over their respective axes, but the number of photons impinging on the diode per second is reduced. Here, higher laser event rates are required to reach the same SNR but higher background event rates can also be tolerated. The required laser event rate increases by factor 10, while the exponential relationship for background influence increases the tolerable background event rate by nearly three decades.

Now, arrival statistics of different combination schemes for pixels are discussed. From these expressions, multi-event and coincidence detection and their combination are also discussed.

3.1.1 First-Photon Detection with Multi-Pixel Structures

The first event of N_S diodes can be found according to equation (3.10) from the probability that one diode triggers $p(r_\mathrm{S}, t)$ while the others have not triggered yet [58].

$$\begin{aligned} p_\mathrm{N}(r,t) &= N_\mathrm{S} r_\mathrm{S} t \exp(-r_\mathrm{S} t) \left(1 - \int_0^t r_\mathrm{S} \exp(-r_\mathrm{S}\tau)\mathrm{d}\tau\right)^{N_\mathrm{S}-1} \\ &= rt\exp(-rt) \end{aligned} \tag{3.10}$$

This is a simplification in that it assumes homogeneous illumination over the diodes. The rate of first-photon detections of multiple detectors is then equal to the total incident rate r.

In contrast, evaluating all first-photon detections of N_S SPAD diodes leads to a detection probability following

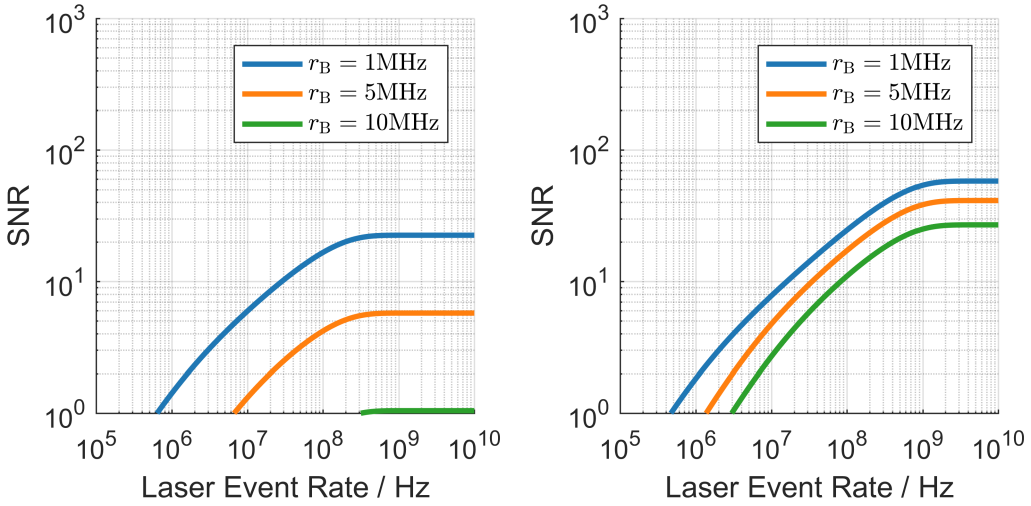
$$p_\mathrm{N}(r,t) = \sum_{j=1}^{N_\mathrm{S}} p(r_\mathrm{S}, t) = N_\mathrm{S} r_\mathrm{S} t \exp(-r_\mathrm{S} t). \tag{3.11}$$

The argument of the exponential function is smaller in this case which means that the function does not saturate as fast as when evaluating only the first event of a single SPAD. This also leads to improved SNR but will increase generated data by the number of detectors.

One can see that the structure of four diodes is more resistant to ambient light and systematically reaches higher SNR. In a real system, this performance gain has to be weighed with the increased generated data.

3.1.2 Multi-Event and Free-Running Detection

Besides SPAD dead time, ME capable TDC often require their own hold-off or dead time between recorded events. In systems where this hold-off time is much shorter than the individual SPAD dead time, SPAD dead time is the dominant factor for the resulting PDF's shape. While the exact relationship is more complicated and depends on a combination of SPAD-individual and TDC dead time, this work simplifies the dependency to be dominated by the TDC conversion time. The applicability of this simplification is further investigated in Chapter 6. For constant event rate this results in a PDF that follows equation (2.39) with an



(a) SNR of a single diode

(b) SNR of $N_S = 4$ diodes

Figure 3.4: SNR comparison between one and four diodes, assuming that incident power is homogeneous.

increased rate of $N_S r_S$, similar to equation (3.10).

From that we can find an expression for the multi-event SNR as

$$k_{\mathrm{SNR\,H}} = \frac{\mu_L}{\sqrt{\mu_L + \mu_B}} = \sqrt{N P_A(N_S r_S, t_{ToF}, N_{Ph})}\,\frac{p_L}{\sqrt{p_L + p_B}}. \tag{3.12}$$

The resulting SNR contour lines are displayed in Figure 3.5. With rising ME levels, high background tolerance can be reached. Each additional level shows diminishing returns in terms of their performance gain, though. Gained background tolerance decreases per level, but the required memory cells close to the TDC and possible data output continue to scale linearly. Furthermore, providing a large number of possible levels leads to most of them being unused in most scenarios.

3.1.3 Coincidence Detection

Now the influence of coincidence detection on arrival process statistics is modeled. The mathematical model for coincidence detection uses the description theorized in [58]. In this work, two cases are separated: coincidence time longer than dead time and shorter than dead time. Coincidence time is frequently limited to the temporal laser pulse width. While coincidence events between background and

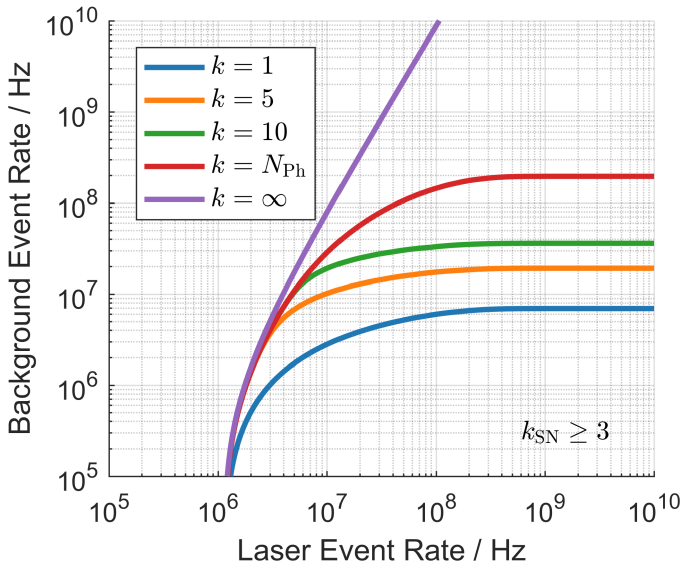


Figure 3.5: SNR contour lines of $k_{\mathrm{SN}} = 3$ for different ME levels and the limiting case of an asynchronous detector. $N_{\mathrm{S}} = 4$, $t_{\mathrm{Bin}} = 8\,\mathrm{ns}$, $N_{\mathrm{M}} = 1000$.

signal photons can be beneficial, the primary use of coincidence detection is to detect the high photon density during the laser pulse return. As laser pulses of modern systems are mostly shorter than current dead times, the more common case today is coincidence time shorter than dead time, and only it is considered further.

The model is now introduced. It aims to derive steady-state coincidence event rates by integrating the respective arrival densities. This is why dead time effects are not neglected in the description and the dead time t_{D} reduced steady-state rate r_{SS} from equation (2.40) is used. The first-photon PDF of a dead time-affected SPAD in steady-state is described by equation (3.13).

$$f_{1,1}(t) = \begin{cases} \dfrac{r_{\mathrm{S}}}{1 + r_{\mathrm{S}} t_{\mathrm{D}}}, & t < t_{\mathrm{D}} \\ \dfrac{r_{\mathrm{S}}}{1 + r_{\mathrm{S}} t_{\mathrm{D}}} \exp(-r(t - t_{\mathrm{D}})), & t \geqslant t_{\mathrm{D}} \end{cases} \tag{3.13}$$

When extending this expression to the first event of N_{S} diodes in a pixel, the detection PDF depends on one of the SPADs triggering while no other has triggered beforehand. The common PDF is thus found to

$$f_{1,N_{\mathrm{S}}}(t) = N_{\mathrm{S}} f_{1,1}(t) \left(1 - \int_0^t f_{1,1}(\tau)\mathrm{d}\tau\right)^{N_{\mathrm{S}}-1}. \tag{3.14}$$

Inserting equation (3.13) leads to

$$f_{1,N_\mathrm{S}}(t) = \begin{cases} N_\mathrm{S}\dfrac{r_\mathrm{S}}{1+r_\mathrm{S}t_\mathrm{D}}\left(1-\dfrac{r_\mathrm{S}}{1+r_\mathrm{S}t_\mathrm{D}}t\right)^{N_\mathrm{S}-1}, & t < t_\mathrm{D} \\ N_\mathrm{S}\dfrac{r_\mathrm{S}}{1+r_\mathrm{S}t_\mathrm{D}}\exp(-r(t-t_\mathrm{D}))\cdot & \\ \left(1-rt_\mathrm{D}+\dfrac{1}{1+r_\mathrm{S}t_\mathrm{D}}(\exp(-r(t-t_\mathrm{D}))-1)\right), & t \geqslant t_\mathrm{D}. \end{cases} \tag{3.15}$$

The inter-arrival time of the next detection is an exponential PDF, corresponding to equation (2.37) for $k = 1$, constant rate, and shifted temporally by a dead time.

$$f_{\mathrm{IA},1}(t) = \Theta(t-t_\mathrm{D})r_\mathrm{S}\exp(-r_\mathrm{S}(t-t_\mathrm{D})) \tag{3.16}$$

Similarly, the concept of the next arrival can be extended to N_S detectors. Because here we are interested in exactly $n_\mathrm{C}-1$ detections, with n_C being the coincidence depth, following the first one and dead times are assumed to be longer than typical coincidence times, only $N_\mathrm{S}-1$ detectors are ready for detection and the expression becomes

$$f_{\mathrm{IA},N_\mathrm{S}}(t) = (N_\mathrm{S}-1)f_{1,1}(t)\left(1-\int_0^t f_{\mathrm{IA},1}(\tau)\mathrm{d}\tau\right)\left(1-\int_0^t f_{1,1}(\tau)\mathrm{d}\tau\right)^{N_\mathrm{S}-2} \tag{3.17}$$

Again inserting the previous expressions leads to (3.18).

$$f_{\mathrm{IA},N_\mathrm{S}}(t) = \begin{cases} (N_\mathrm{S}-1)\dfrac{r_\mathrm{S}}{1+r_\mathrm{S}t_\mathrm{D}}(1-\dfrac{r_\mathrm{S}}{1+r_\mathrm{S}t_\mathrm{D}}t)^{N_\mathrm{S}-2}, & t < t_\mathrm{D} \\ (N_\mathrm{S}-1)\dfrac{r_\mathrm{S}}{1+r_\mathrm{S}t_\mathrm{D}}\exp(-2r(t-t_\mathrm{D}))\cdot & \\ \left(1-rt_\mathrm{D}+\dfrac{1}{1+r_\mathrm{S}t_\mathrm{D}}(\exp(-r(t-t_\mathrm{D}))-1)\right)^{N_\mathrm{S}-2}, & t \geqslant t_\mathrm{D} \end{cases} \tag{3.18}$$

Finally, the probability density for a successful coincidence event can be found. It requires $k = N_\mathrm{S}-1$ previous detections and another one during detector's observation time. The joint probability distribution can be found recursively via convolution of the arrival densities.

$$\begin{aligned} f_{k,N_\mathrm{S}}(t) &= f_{k-1,N_\mathrm{S}}(t) * f_{\mathrm{IA},N_\mathrm{S}}(t) \\ &= \int_0^\infty f_{k-1,N_\mathrm{S}}(\tau) f_{\mathrm{IA},N_\mathrm{S}}(t-\tau)\mathrm{d}\tau \end{aligned} \tag{3.19}$$

The condition for a successful coincidence detection is a combination of the aforementioned detection conditions. Firstly, a coincidence detection cannot occur without detecting the first photon inside an array of multiple SPADs f_{1,N_S}. The expected time for this event is summarized in a time $\overline{t_1}$.

$$\overline{t_1} = \int_0^\infty f_{1,N_\mathrm{S}}(\tau)\tau\mathrm{d}\tau \tag{3.20}$$

The next time is the mean time it takes for the remaining events to occur during the coincidence time. This expected value is normed because integration over coincidence time t_C leaves out finite probability outside of the integration interval.

$$\overline{t_2} = \frac{1}{\int_0^{t_\mathrm{C}} f_{k,N_\mathrm{S}}(\tau)\mathrm{d}\tau} \int_0^{t_\mathrm{C}} f_{k,N_\mathrm{S}}(\tau)\tau\mathrm{d}\tau \tag{3.21}$$

Two expected times are yielded. The time until first-photon detection always has to pass. The expected time for a coincident event is weighed with the probability that a coincidence event occurs at all. This probability is expressed as

$$P_\mathrm{C} = \int_0^{t_\mathrm{C}} f_{k,N_\mathrm{S}}(\tau)\mathrm{d}\tau. \tag{3.22}$$

The mean time that passes when no coincidence event is detected must be modeled in a similarly. If no coincidence event occurs, coincidence time t_C has to pass completely. Additionally, the mean time that it takes for the next coincidence event $\overline{t_\mathrm{C}} = r_\mathrm{C}^{-1}$ has to pass before the next coincidence event is triggered. These two times are added and weighed with the probability of an unsuccessful detection.

The mean time between coincidence events is then

$$r_\mathrm{C}^{-1} = \overline{t_1} + (t_\mathrm{C} + r_\mathrm{C}^{-1})(1 - P_\mathrm{C}) + \overline{t_2}P_\mathrm{C}. \tag{3.23}$$

Solving this equation gives a closed-form solution for the required coincidence event rate r_C as

$$r_\mathrm{C} = \frac{P_\mathrm{C}}{\overline{t_1} + t_\mathrm{C}(1 - P_\mathrm{C}) + \overline{t_2}P_\mathrm{C}}. \tag{3.24}$$

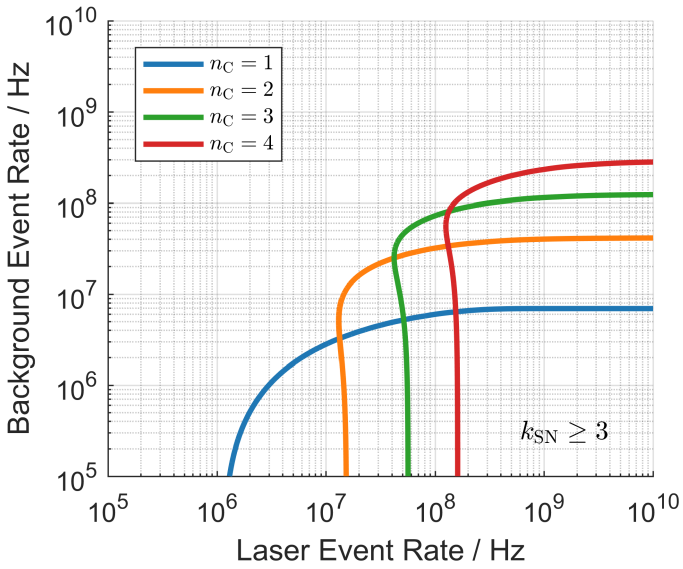


Figure 3.6: Coincidence detection SNR contour lines of $k_{\mathrm{SN}} = 3$ for different coincidence depths n_{C} and first-photon detection. $N_{\mathrm{S}} = 4$, $t_{\mathrm{Bin}} = 8\,\mathrm{ns}$, $N_{\mathrm{M}} = 1000$.

The form presented in equation (3.24) links single-detector photon rate and coincidence rate r_{C}. Again using the SNR model for piece-wise constant functions, the respective rates for background $r_{\mathrm{C,B}} = f(r_{\mathrm{B}})$ and background plus signal $r_{\mathrm{C,LB}} = f(r_{\mathrm{L}}, r_{\mathrm{B}})$.

$$k_{\mathrm{SN\,C}} = \sqrt{N t_{\mathrm{Bin}} \exp(-r_{\mathrm{C,B}} t_{\mathrm{ToF}})} \frac{r_{\mathrm{C,L}}}{\sqrt{r_{\mathrm{C,L}} + r_{\mathrm{C,B}}}} \tag{3.25}$$

The non-linear attenuation of coincidence detection benefits detectability because low event rates are more attenuated than high event rates. In scenarios where $r_{\mathrm{C,B}}$ approaches zero, coincidence detection can even be detrimental to successful detection.

$$k_{\mathrm{SN\,C}} = \sqrt{N t_{\mathrm{Bin}} r_{\mathrm{C,L}}} \leqq \sqrt{N t_{\mathrm{Bin}} r_{\mathrm{L}}} \tag{3.26}$$

Also, $r_{\mathrm{C,LB}}$ depends on background event rate. The resulting contour lines for coincidence detection are seen in Figure 3.6. It is again visible that higher coincidence levels require higher laser power for successful detection. At the same time, higher ambient light levels can be tolerated. This corresponds somewhat to a change in sensitivity. Counter-intuitively, the resulting output count rate during signal return can be lower when less ambient light is present. Careful choice of

coincident detection settings is thus required for best system performance.

3.2 Multi-Event Coincidence Detection

With multi-pixel structures, a combination of multi-event and coincident detection is possible. Both multi-event and coincident [53] detection provide advantages at high ambient illumination and have certain drawbacks.

Through coincidence detection, ambient noise can be suppressed and a better ratio of measurements containing ambient photons to measurements with signal photons can be achieved. At the same time, weak signals are being attenuated, requiring higher laser event rates than first-photon detection for successful measurement. Especially for fixed coincidence times, attenuation of the laser pulse can be as high as tolerable for the given returned laser power, but the suppression of background photons is not sufficient to prevent background detections before pulse return. Another drawback in using coincidence detection lies in its additional timing uncertainty. Longer coincidence times mean that later pulse photons can be registered in the histogram, and predicting their distribution is intensity-dependent and not trivial.

Additionally, the fact that higher coincidence levels require multiple detectors to be triggered affects distribution variance. Implementations that generate a timestamp from the last event introduce additional randomness compared to first-photon detection. This effect scales with the length of inter-arrival times and is thus especially severe in low-rate scenarios.

A combined expression for multi-event and coincidence detection is given by coincident counting rates and multi-event saturation behavior to

$$k_{\mathrm{SN\,ME,CD}} = \sqrt{N P_{\mathrm{A}}(r_{\mathrm{C,B}}, t_{\mathrm{ToF}}, N_{\mathrm{Ph}})}\frac{p_{\mathrm{C,L}}}{\sqrt{p_{\mathrm{C,L}} + p_{\mathrm{C,B}}}}. \tag{3.27}$$

A similar contour plot as previously is given in Figure 3.7. If the contours are interpreted as a limit of evaluability, best system performance is achieved when switching sensor properties every time the contours intersect. In comparison to Figure 3.6 one can see that ME detection increases the background light tolerance for each given coincidence level. This is especially interesting if laser power is limited, which it always is when complying with eye safety norms. Otherwise,

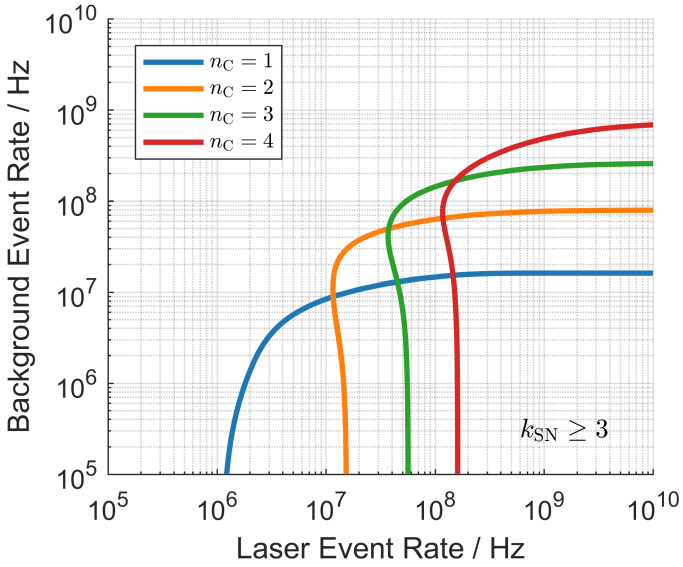


Figure 3.7: Multi-Event Coincidence SNR contour lines of $k_{SN} = 3$ for different coincidence depths n_C and $k = 4$. $N_S = 4$, $t_{Bin} = 8\,ns$, $N_M = 1000$.

coincidence was always a sufficient solution for the problem of high ambient illumination. Notably, the results imply that for certain laser event rates, ambient light is beneficial for detection.

The high dynamic range between ambient light levels and illuminated target objects in the scene means that choosing the right level for multi-event detection and coincidence requires an analysis of signal levels in the measurement result. Where one pixel observes a part of the scene that is comparatively dark, another one could see a highly reflective target, requiring different sensor settings.

The considerations in this section show the increase in ranging performance through multi-event and coincident detection. At the same distance, higher ambient light tolerance is achieved. In high ambient light scenarios, an increased range can be covered.

In the course of this work, a combination method for coincident and multi photon detection has been developed which is explained in Section 6.5.2. A schematic overview of favorable operating conditions for the respective combinations of ME and coincidence detection is given in Figure 3.8.

First-photon detection is sufficient when both laser and ambient power are low. Coincidence detection will provide the best performance if laser power is high but ambient is relatively low because both are attenuated. The case where laser is low but ambient light high is best covered by ME detection because laser

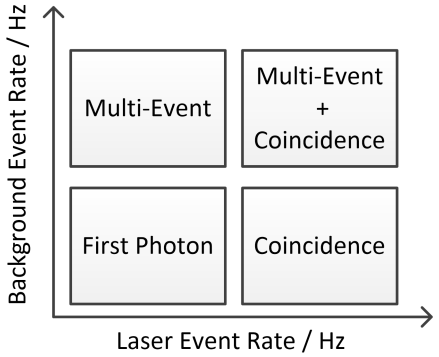


Figure 3.8: Summary of best signal conditions for the introduced sensor operating modes.

return power is not attenuated further. If both laser and ambient are high, ME detection combined with a low-level coincidence threshold is ideal, combining the non-linear attenuation of coincidence detection with the saturation preventing effects of ME detection.

Next, a comparison of the SPAD behavior with ideal settings choice and an APD with typical values for rangefinding is conducted.

3.3 Comparison with Avalanche Photodiodes

As mentioned previously, APDs and SPADs are related in nature. APDs use a negative biasing voltage below the breakdown voltage and proportionally amplify their input signal while SPADs react to impinging photons with a saturation current. To choose the suitable detector for the respective pulsed ToF application, both diode behavior and expected operating conditions have to be evaluated. The detector comparison was conducted previously to this work and published in [59]. The results are again presented here and extended by including combined multi-event and coincidence detection. For comparison, parameters of diodes currently available on the market are considered but an outlook for future technological development is also given.

3.3.1 Avalanche Photodiode Noise Model

Now, an expression for APD SNR is derived for comparison with the previously established expressions for SPADs. As APDs are analog devices, both device and

circuit noise must be considered. The signal and noise terms are modeled by their equivalent current densities.

Excess shot noise is one effect in both SPADs and APDs that influences both signal and noise. The internal APD gain is advantageous for detecting small return signals but the stochastic process of avalanche multiplication leads to excess noise.

A formulation of the excess noise factor F depending on the ionization coefficients for holes α_p and electrons α_n, is given by [60] as

$$F = M\left(\frac{\alpha_\mathrm{p}}{\alpha_\mathrm{n}}\right) + \left(2 - \frac{1}{M}\right)\left(1 - \frac{\alpha_\mathrm{p}}{\alpha_\mathrm{n}}\right). \tag{3.28}$$

As can be seen, excess noise depends highly on the amplification M of the diode. In silicon-based diodes, the electron ionization coefficient is higher than the ones for holes and for high amplification, the term becomes

$$F \approx 2 + M\left(\frac{\alpha_\mathrm{p}}{\alpha_\mathrm{n}}\right). \tag{3.29}$$

Typical values of excess noise factor range from around 2 to 3 in silicon-based APDs [61]. Noise currents amplified by the internal gain comprise the photon generated currents I_S and I_B for signal and background photons, respectively, and the parts of dark current flowing through the multiplication region in the bulk I_db. The total dark current also contains the surface dark current I_ds so that the total output current of the APD I is found to be

$$I = I_\mathrm{ds} + M(I_\mathrm{db} + I_\mathrm{S} + I_\mathrm{B}), \tag{3.30}$$

which leads to a shot noise current density $\langle I \rangle$ of

$$\langle I \rangle = \sqrt{2e(I_\mathrm{ds} + M^2 F(I_\mathrm{db} + I_\mathrm{S} + I_\mathrm{B}))}. \tag{3.31}$$

The proportionality factor between incident optical power P and generated current I is the responsivity $\mathbf{R}$ of the APD after amplification. When expressed without the amplification M, the symbol $\mathbf{R}'$ is used. The symbol e is the elementary charge constant.

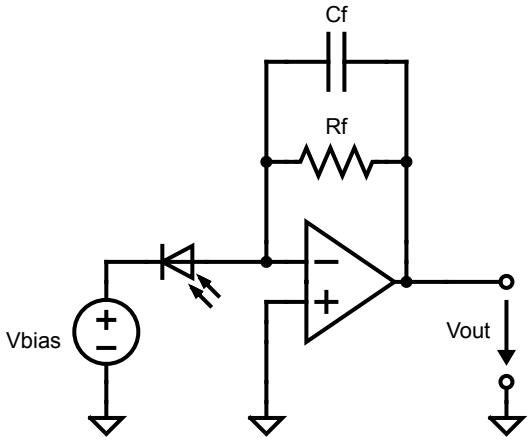


Figure 3.9: Typical readout circuit for an APD and TIA.

$$I = \mathbf{R}P = \mathbf{R}'MP \tag{3.32}$$

Besides internal noise contributors, external circuitry also introduces noise. As a typical example circuit, a Transimpedance Amplifier (TIA) is employed to convert the output current of the APD into a proportional output voltage. An overview of the diode plus circuitry is shown in Figure 3.9. The feedback resistor R_f produces thermal Johnson-Nyquist noise. The effective noise current follows

$$I_\mathrm{Rf} = \sqrt{B_\mathrm{N}}\langle I_\mathrm{Rf}\rangle = \sqrt{\frac{4k_\mathrm{B}TB_\mathrm{N}}{R_\mathrm{f}}}, \tag{3.33}$$

with the bandwidth B_N being the proportionality factor between current densities and mean current, k_B the Boltzmann constant and T the effective noise temperature. The capacitance C_f in the feedback loop compensates for a pole in the gain characteristic.

The operational amplifier also introduces noise. It is modeled as an additional current density at the output of the amplifier. The virtual ground at the input of the operational amplifier ensures that the output voltage V_out is equivalent to the voltage drop-off over R_f. At low frequencies, it can thus be expressed as a current in the feedback loop following Ohms law as

$$\langle I_\mathrm{amp}\rangle = \frac{\langle V_\mathrm{amp}\rangle}{R_\mathrm{f}} \tag{3.34}$$

Summarizing these noise current contributions yields an expression of APD

SNR following

$$k_{\mathrm{SN\,APD}} = \frac{I_{\mathrm{S}}}{\sqrt{2\mathrm{e}B_{\mathrm{N}}(I_{\mathrm{ds}} + (I_{\mathrm{S}} + I_{\mathrm{B}} + I_{\mathrm{db}})MF) + B_{\mathrm{N}}\left(\frac{4\mathrm{k_B}T}{R_{\mathrm{f}}} + \langle I_{\mathrm{amp}}^2 \rangle\right)}}. \tag{3.35}$$

Neglecting surface dark current I_{ds}, accounting for N individual signal acquisitions and finite fill factor η_{FF}, and relating the shot noise terms to incident optical powers P_{S} and P_{B} leads to the expression

$$k_{\mathrm{SN\,APD}} = \sqrt{\frac{N_{\mathrm{M}}\eta_{\mathrm{FF}}^2 \mathbf{R}'^2 P_{\mathrm{S}}^2}{2eB_{\mathrm{N}}(\mathbf{R}'\eta_{\mathrm{FF}}P_{\mathrm{S}} + \mathbf{R}'\eta_{\mathrm{FF}}P_{\mathrm{B}} + I_{\mathrm{db}})F + \frac{1}{M^2}B_{\mathrm{N}}\left(\frac{4\mathrm{k_B}T}{R_{\mathrm{f}}} + \frac{\langle V_{\mathrm{amp}}^2 \rangle}{R_{\mathrm{f}}^2}\right)}}. \tag{3.36}$$

where P_{S} and P_{B} are optical power on the diode from signal and background respectively, I_{db} the bulk dark current, T the effective noise temperature, V_{amp} the amplification voltage, e the elementary charge, R_{f} the value of the resistor in the TIA feedback loop and N_{M} is the number of accumulated measurements.

3.3.2 System Comparison

Now, two theoretical DToF sensor systems built with APD and SPAD detectors are compared in terms of their required signal and tolerable background optical powers. The same optics, time resolution and laser power are assumed because they are independent of the detector technology. At first, diodes of equal size and ideal photon conversion efficiency are compared. Then, systems with real parameters are considered.

Idealized Detectors

The whole pixel area is optically sensitive, resulting in $\eta_{\mathrm{FF}} = 1$. In both detectors, all incident photons generate primary electron-hole pairs and they are detected before they could recombine.

For the APD, this expression equates to a quantum efficiency of unity. The resulting value of $\mathbf{R}'$ depends on the operating wavelength λ_0 and is connected to quantum efficiency following equation (3.37).

$$\mathbf{R}' = \frac{I}{P} = \frac{N_\mathrm{e} e}{N_\mathrm{O} W_\mathrm{Ph}} = \eta_\mathrm{QE} \frac{\mathrm{e}\lambda_0}{\mathrm{hc}} \tag{3.37}$$

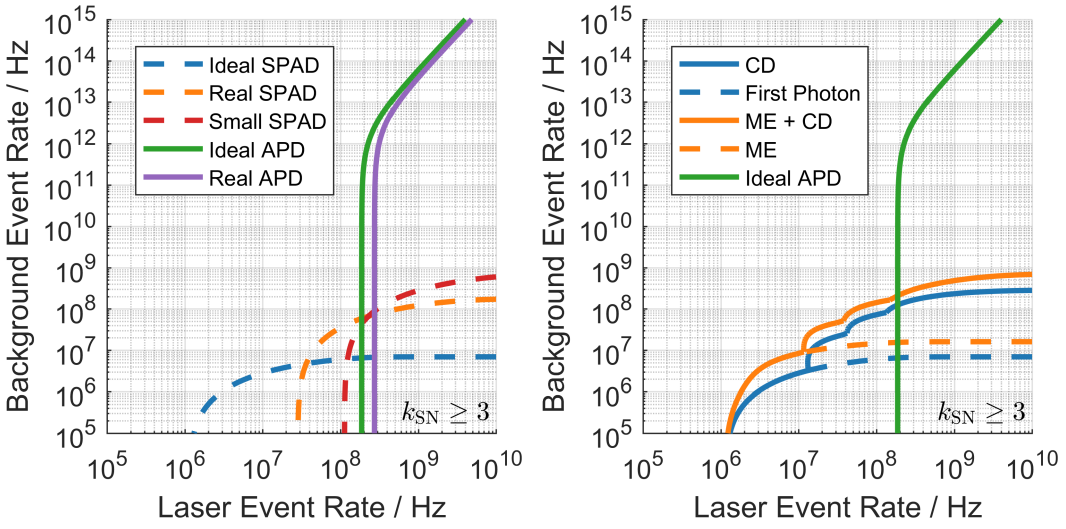
The current through the diode I is formed by a number of electrons N_e with charge e. On the optical side, incident optical power comprises a number incident photons N_O carrying energy W_Ph. Their conversion ratio is defined as the quantum efficiency η_QE of the APD. An ideal APD converts incoming photons with $\eta_\mathrm{QE} = 100\,\%$. The resulting value for the maximum responsivity at a wavelength of $\lambda_0 = 905\,\mathrm{nm}$ is

$$\mathbf{R}'_{905\,\mathrm{nm}} = \frac{\mathrm{e}\lambda_0}{\mathrm{hc}} = 0.7299. \tag{3.38}$$

The ideal APD does not suffer from dark current, hence $I_\mathrm{db} = 0$. The Johnson noise and amplifier noise terms are also set to zero. The only modeled noise factor is excess noise with $F = 2$ at amplification $M = 100$. The evaluation is conducted with a bandwidth of $B_\mathrm{N} = 450\,\mathrm{MHz}$. The resulting SNR is dominated by Poisson variance or photon shot noise.

Certain idealizations are used to compare with the limiting case of an ideal SPAD. To have the same power budget on both receivers, the laser event rate r_L incident on the SPAD is converted to an optical power on the APD via equation (2.7). The photon detection probability of the SPAD is set to $\eta_\mathrm{PDP} = 1$. Analogously to the zero dark current of the APD, the SPAD is assumed to have zero probability of detecting a dark count during a measurement interval expressed in $r_\mathrm{DC} = 0\,\mathrm{Hz}$. The SPAD is also shot noise limited. Both diodes are assumed to sum up $N_\mathrm{M} = 1000$ measurements and sense a target at $t_\mathrm{ToF} = 667\,\mathrm{ns}$, corresponding to a planar target at $100\,\mathrm{m}$.

The resulting SNR contour lines for the ideal APD and SPAD are given in Figure 3.10. Only the first-photon behavior of the SPAD is given. One can see that the APD requires more power for successful detection. If this minimum power can be provided, more unwanted background photons can be tolerated than with the SPAD detector. The SPAD can detect at much smaller power budgets. If however, ambient light is present, measures have to be taken to suppress it. The SNR contour lines for ME and coincidence detection is shown in Figure 3.10 b). The dynamic range of the diode can be extended significantly but cannot match the ambient light suppression of the APD. These results match the



(a) First-photon detection and finite detection efficiency.

(b) Multi-event and coincidence detection for the ideal SPAD.

Figure 3.10: Comparison of SPAD and APD receiver SNR.

general behavior of the simulations and measurement results identified by [62] for succesful first-photon, coincidence, and APD evaluation.

Realistic Detectors

Besides the results for idealized assumptions about the detectors, some for more realistic systems are given. The real APD is based on the SARP500X [63] from Laser Components. It exhibits a reduced sensitivity of $\mathbf{R}' = 0.5\,\mathrm{A\,W^{-1}}$ and a bulkt dark current of $I_{\mathrm{db}} = 0.5\,\mathrm{nA}$. A few assumptions about the associated read-out circuit are made. Amplification voltage density is $\langle V_{\mathrm{Amp}} \rangle = 1.414 \times 10^{-7}\,\mathrm{V}/\sqrt{\mathrm{Hz}}$, $R_{\mathrm{f}} = 1\,\mathrm{k\Omega}$, $T = 300\,\mathrm{K}$. The SNR contour shows that higher optical power is required for the same evaluability because of the reduced sensitivity. The real SPAD plot shows the resulting SNR for $r_{\mathrm{DC}} = 100\,\mathrm{Hz}$, a high effective fill factor of $\eta_{\mathrm{FF}} = 0.80$, which is a competitive value achievable by employing microlenses, and $\eta_{\mathrm{PDP}} = 0.05$ for incident photons with a wavelength of $\lambda_0 = 905\,\mathrm{nm}$. It also requires a higher minimum signal but can tolerate slightly higher ambient power.

One advantage of SPADs is that they are more easily implementable in array structures. The considered APD has a diameter of 500 µm. Typical SPAD size is way smaller than that. For the last comparison, SPAD diodes with a pixel pitch

of 20 µm are used. The area of one APD is filled with an equivalent number of SPAD detectors, with parameters following the previously discussed 'real SPAD', resulting in an array of 25×25 diodes. The resulting SNR is shown as the small SPAD in Figure 3.10.

One can see that the required minimum power is in a similar regime as for the APD. However, this detector brings inherent spatial resolution. The APD structure can only measure a single point in space. For spatial resolution, beam-steering is required and achievable by reflecting the laser beam off polygon or Micro Electro Mechanical System (MEMS)-based scanners. The SPAD system inherently provides spatial resolution. It can be used as a flash system and has the advantage of being completely solid-state, increasing ruggedness, decreasing the need for optics calibration, and attenuating outside mechanical influences like vibration.

Chapter 4

Photonic Simulator

Analytical calculation can provide a quick and closed-form way to predict a system's performance. However, system response behavior or influence of external circuitry can prevent finding analytical solutions. Numerical simulation can enable more insight into these more complex systems.

The photonic simulation can be characterized as a technique in which the temporal behavior of an optical system is imitated [64]. It is especially advantageous if multidimensional random variables are of interest, as with most optical systems. Statistical influences like source behavior, channel fluctuations, and diode response function on range, accuracy, and precision can be predicted. The drawback, however is that for covering a wide range of scenarios, generation of large numbers of random values is required.

The employed simulation techniques and system model are laid out in the following chapter. Afterward, they are used to simulate the behavior of three distinct applications.

4.1 Arrival Time Generation

The photonic simulator uses a pseudo-random number generator to generate photon arrival times. Depending on whether the CDF F_s of the arrival process is invertible, which means that a function F_s^{-1} exists, different sampling methods are possible as visualized in Figure 4.1. The computationally more efficient Inverse Transform Sampling (ITS) is beneficial for simpler arrival time distributions while the more general tool of Rejection Sampling (RS) can be used for

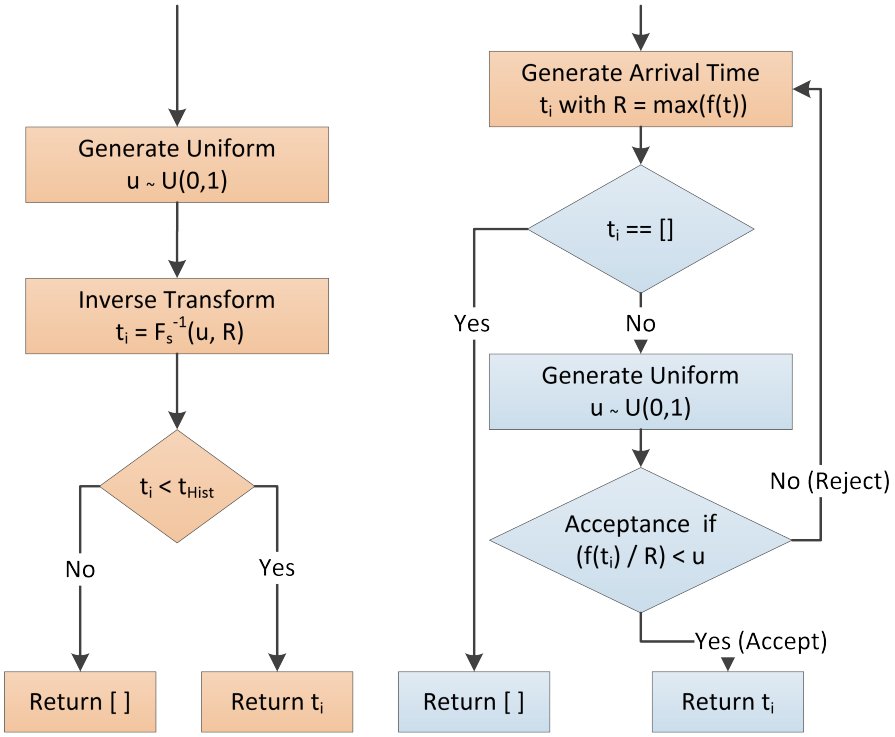


Figure 4.1: Two sampling modes for generating a single arrival time t_i. Left: ITS (shaded in orange). Right: RS. If applicable, ITS is preferable because it is computationally more efficient.

arbitrary distributions.

If the CDF is invertible, ITS is an optimal choice because it has the advantage of being computationally efficient. In this method, a sample from a uniformly distributed variable $u \sim U(0,1)$ is drawn and inserted into the inverted CDF F_{s}^{-1} to transform the uniform distribution into the distribution with the desired shape. Per generated arrival time, only a single random number is required. The generated arrival time is empty if it lies outside of the histogram observation interval t_{Hist}.

For simulating arbitrary time-dependent rates, which is interesting for generated arrival times of realistic temporal laser pulse shapes without analytically reversible CDF, a computationally more intensive algorithm has to be used. It is called Rejection Sampling and requires generating more random values than inverse transform sampling. The first step of RS is generating an arrival time from a HP using the maximum rate value R present in the PDF. If that generated value lies in the observation interval, another uniformly distributed random value u is generated to decide whether this arrival time is being accepted or rejected.

The condition for acceptance is that the instantaneous value of the PDF $f(t_i)$ divided by the maximum value R, is smaller than the generated value of u. If so, the arrival time is stored (accepted). Otherwise, the value is discarded (rejected) and the process starts over. [65]

4.2 Histogram Generation

The generated arrival times represent photons incident on a single detector. Generation of the complete DToF histogram of a pixel is carried out using the sequence in Figure 4.2.

The previously sketched arrival time generation is conducted for all SPADs inside a single pixel. Each generated arrival time is stored in a mutual data structure. Afterward, all arrival times occurring during SPAD dead time are deleted from the arrival time structure.

Once arrival times for all single SPADs are generated, they are put into another common array and sorted in ascending order. If coincidence settings are given, coincident arrivals are identified, and all others deleted. Then, for systems that have a circuit wide dead time either caused by TDC conversion time or a similar effect, another dead time cleansing is conducted.

Now, ME arrivals can be filtered out by deleting all arrivals after the first k. As the last step, discretization is emulated by rounding all arrival times to integral multiples of the bin size t_{Bin}. This procedure is carried out for each measurement or sent out laser pulse, resulting in N_{M} iterations. The resulting bin count values are summed up for each one, resulting in the simulated histogram.

Once the histogram is generated, the only step left is to determine ToF from the histogram. Pre-processing steps can be executed where ambient photons are compensated, or evaluation schemes like FIR filters are applied. From this processed histogram, ToF is derived using an appropriate evaluation technique. The next chapter gives more detailed information for the latter two steps.

4.3 Example Applications

Based on the simulation engine from the previous section, example systems can be modeled and evaluated. The system modeling makes it possible to estimate

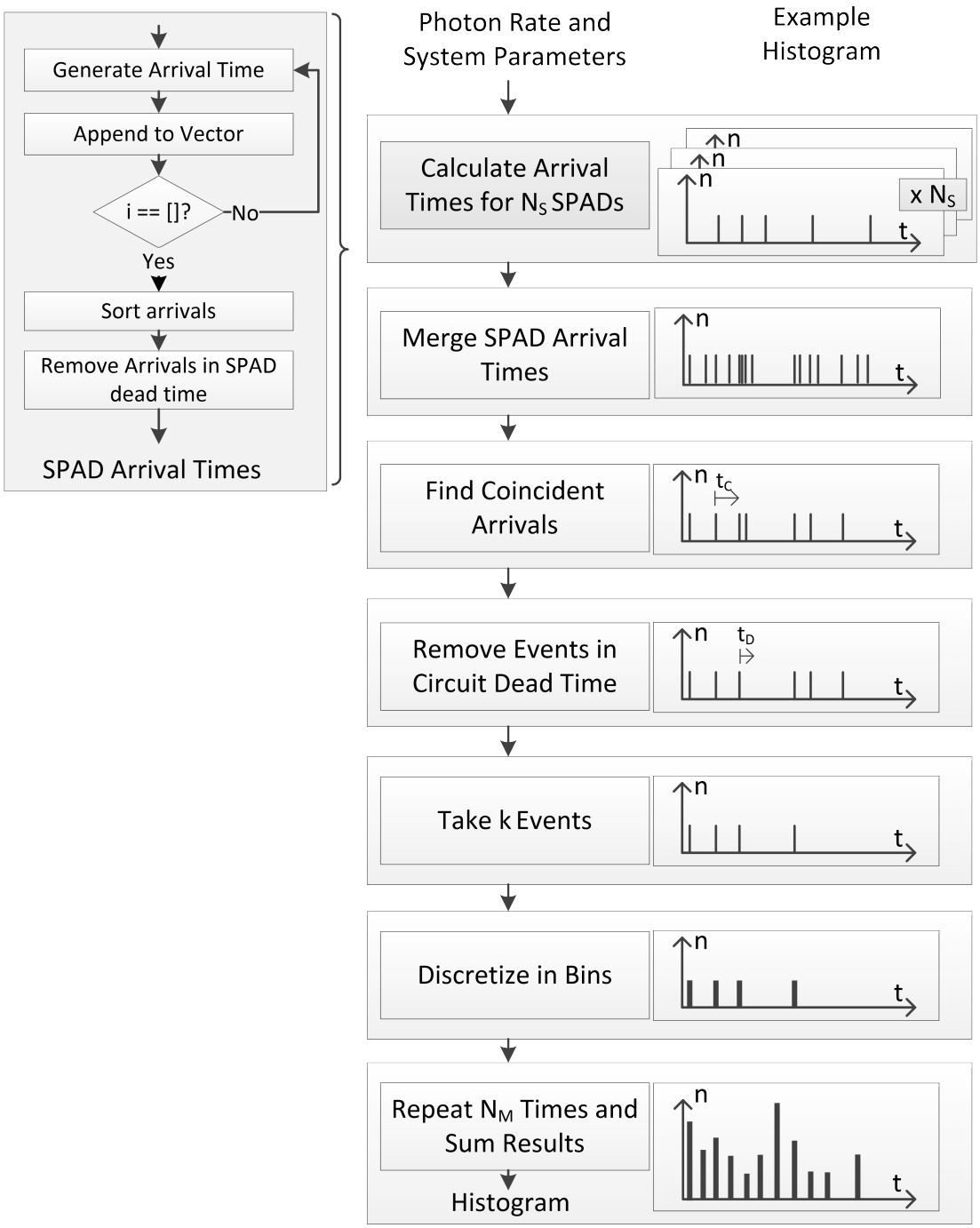
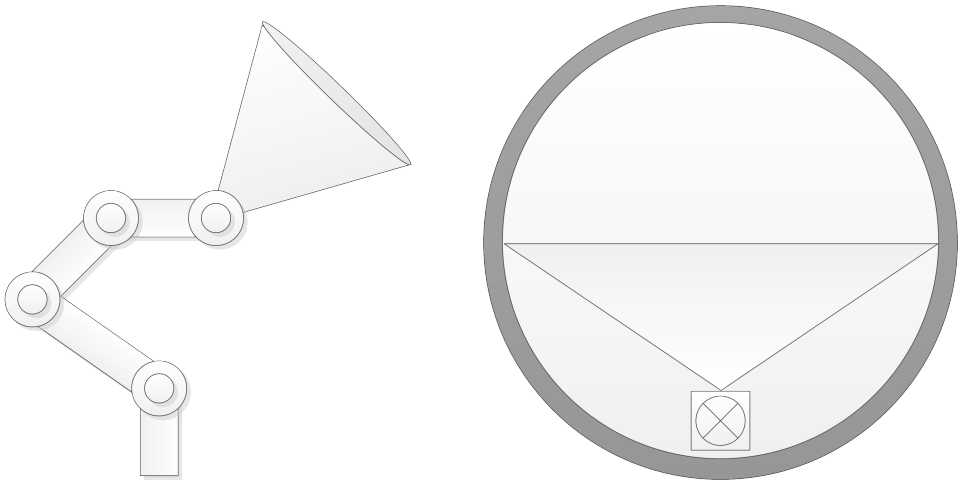


Figure 4.2: Sequence diagram illustrating histogram building. Arrivals are generated on a single diode basis, merged, arrivals in dead time removed, arrivals during coincidence time found, and k events saved in the histogram.



(a) Industrial robot. (b) Canal robot.

Figure 4.3: Visualizations for the industrial and canal robot.

range, precision, and accuracy. The main differences between applications are the amount of ambient illuminance present, the FoV to be covered, reflection characteristics of the target objects, required range, and frame rate. For further analysis, three different applications are defined and discussed in the following.

Industrial Robot

The first application is an industrial robot with a fixed operating area. Capturing 3D information can be of great value in the cooperation of human and artificial working force. While light curtains are a proven, sturdy, and flexible way to stop machine operation when human operators get in its vicinity, they can only detect the crossing of simple borders and lines. A device mounted on a robot arm, observing a solid angle spanned in space, can detect the characteristic features of the human body and navigate accordingly. A visualization of a possible system placement is given in Figure 4.3 a).

For the envisioned sensor in an industrial robot application, a required maximum range of $d = 5\,\mathrm{m}$ is more than sufficient. Expected target objects in this environment are non-cooperative, meaning they are not optimized for optical detection. That is modeled in a reflection characteristic of targets with $\rho = 10\,\%$ and $\rho = 80\,\%$, which are simulated as limiting cases. A wide FoV of $90° \times 90°$ should be covered to limit the number of necessary systems for detecting the whole space around the robot. For a full horizontal scene coverage, only four

systems are required. To guarantee fast reaction time and with that high possible movement speed of the system itself, a frame rate of $f_{\mathrm{FPS}} = 50\,\mathrm{Hz}$ or a time between frames of 20 ms is targeted.

While sunlight can be present by falling in through hall windows, it can be limited via shutters. The selective use of wavelengths other than the one used for the DToF system also helps in relaxing requirements on ambient light suppression in the histogram and only medium ambient illuminance $E_{\mathrm{V}} = 20\,\mathrm{klx}$ should be tolerated. If higher levels of automation find their way into fabrication processes, fabrication halls without any lighting and human operators could become reality. There, only sensor inherent noise effects limit ranging performance. While certainly realizable and advantageous for energy consumption, the dark factory currently still seems far out of reach [66].

Canal Robot

The second application is another robotics application, but this time for a controlled environment. The use of robots in photon-starved environments like inside pipes or canals comes close to the long-term development of the dark factory. For these systems, high accuracy is critical, while ambient light suppression plays a tangential role. Also, system velocity is limited and spatial correlation between subsequent frames can be assumed. Accumulation of large statistics to increase depth resolution and accuracy is viable. Ambient light will only be present because of artificial lighting on the robot itself like status Light Emitting Diodes (LEDs) or stray light of low-intensity coupling in from the surface into the canal system. The ambient scenario is expressed in a reduced ambient illuminance of $E_{\mathrm{V}} = 2\,\mathrm{klx}$. The robot traverses the canal slowly and senses it orthogonally to the direction of its movement. It thus requires only low Frames per Second (FPS) of $f_{\mathrm{FPS}} = 10\,\mathrm{Hz}$ and a small FoV in the direction of movement but a high one in the transversal direction. This requirement is addressed with a targeted FoV of $10° \times 120°$. Three of these solid-state systems or a time-dependent axis would be required for full coverage. Reflectance is assumed to be known to $\rho = 80\,\%$. The application inside a canal limits the maximum range to $d = 2\,\mathrm{m}$.

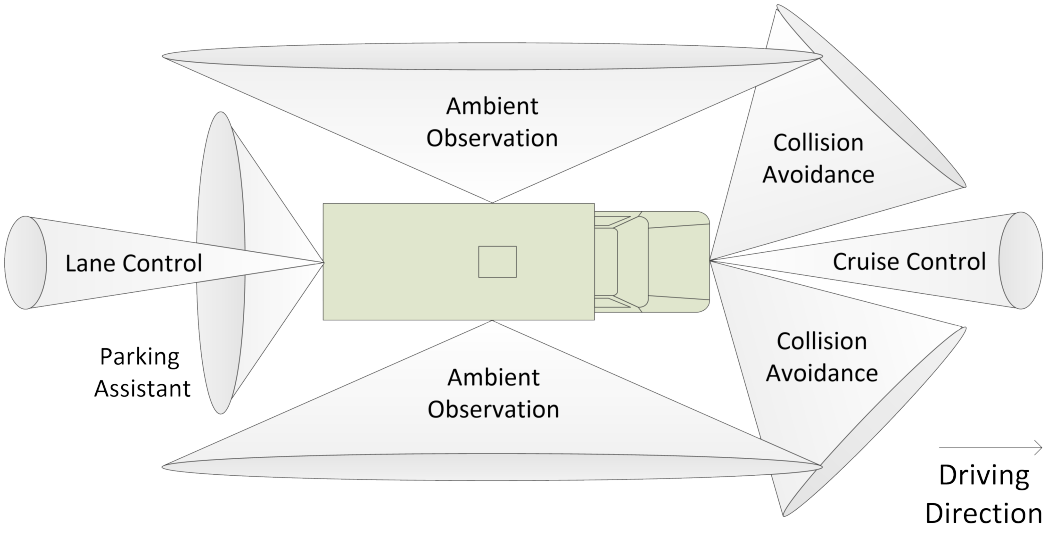


Figure 4.4: Example system placement for different tasks in autonomous driving inspired by [67].

Automotive Scenario

The third application is an automotive collision avoidance application with a medium FoV of 15° × 30°. LiDAR requirements for automotive applications change depending on the used observation concept for the car's environment and choice of sensors. Currently available sensors are often mounted on top of the vehicle. The disadvantage is that very short distances around the vehicle can not be covered. A possible concept for covering a wide area and the space around the vehicle is shown in Figure 4.4.

For very wide FoV, short range and high precision applications like parking assistance, microwave-based systems are currently employed and aided by conventional cameras. Other sensors like the indicated system for front collision avoidance can be placed inside the car's headlight, also covering small objects on the ground next to the vehicle. For small FoV and high range systems like indicated for lane and adaptive cruise control, scanning LiDAR or Radio Detection and Ranging (RADAR) systems seem ideally fit.

Systems for autonomous driving still provide a challenge for LiDAR solutions. One of the most pressing issues is maintaining high frame rates with bright sunlight present when sensing non-cooperative targets. Consequently, most commercial, automotive systems are still based on macroscopic, rotated mirrors that sequentially scan the scene. MEMS-based systems with advantages mainly in power

Parameter	Symbol	Value			Unit
		Auto	Industry	Canal	
Frame Rate	f_{FPS}	25	50	10	Hz
Opening Angle Horizontal	θ_{H}	30	90	120	°
Opening Angle Vertical	θ_{V}	15	90	10	°
Target reflectivity	ρ	10 to 80	10 to 80	80	%
Max. Ambient Illuminance	E_{V}	50	20	2	klx
Max. Distance	d	50	5	2	m

Table 4.1: Example DToF system requirements.

consumption and system dimensions approach market readiness. However, they still have issues because of their limited FoV or low frame rates when covering a wide FoV. A promising development that currently lacks market readiness is beam steering via photonic crystal waveguides. [68]

For wider FoV systems, the here considered flash LiDAR is a viable alternative. However, flash LiDAR systems have issues at high ambient light and require the acquisition of multiple frames for achieving sufficient SNR. As automobiles are used all around the globe in all kinds of environments, only few constraints about occurring objects and ambient light scenarios can be formulated. This leads to a high dynamic between occurring objects and a challenging scenario for ToF sensors, expressed in a high ambient illuminance of $E_{\mathrm{V}} = 50\,\mathrm{klx}$. Diffusely reflecting objects between $\rho = 10\,\%$ and $\rho = 80\,\%$ should be detected. While the reflected laser power can be higher than that in the case of e.g., street signs with directed reflection characteristics, these cases are generally easier to detect, not harder, and thus excluded for simplicity's sake. Available frame rates and range have to be comparatively high as moving objects should be tracked from an itself moving sensor system. In this example, they are set to $f_{\mathrm{FPS}} = 25\,\mathrm{Hz}$ and $d = 50\,\mathrm{m}$.

The assumed requirements for the three different applications are summarized in Table 4.1.

4.3.1 System Parameters

Now the chosen system parameters for covering the introduced requirements are presented. The automotive and industrial applications are operated in areas populated by humans. At 905 nm, the operating wavelength λ_0 is chosen outside of the wavelength range of human vision (400 to 700 nm) but not too high for silicon-based detectors. The canal robot can use a wavelength in the visible regime because safety issues regarding human operators do not limit it. A standard red laser wavelength is 650 nm. At this wavelength, silicon detector sensitivity is higher than in the near-infrared. The value for PDP at 650 nm is $\eta_{\mathrm{PDP}} = 20\,\%$ and for 905 nm only $\eta_{\mathrm{PDP}} = 10\,\%$, which is already a high value for this operating wavelength.

All source parameters are chosen at their estimated eye safety limit. A high PRF is beneficial because it will let the system accumulate many measurements during a short time. According to eye safety limitations [35], the PRF can limit emittable peak power. Shorter temporal laser pulse width is generally beneficial for the accuracy of the systems. Signal-correlated photons will occur during the whole temporal laser pulse width and will thus deviate further from ToF if the laser pulse width is long. Evaluating the measurement with pulse shape-dependent filters like matched filters can reduce the error introduced by long pulses. Shortening the temporal pulse width allows for higher optical peak power. Unfortunately, available optical peak powers are limited by available devices, so it can be beneficial to use longer pulses with the same peak power to increase transmitted energy overall.

For the industrial robot, a PRF of $f_{\mathrm{PR}} = 100\,\mathrm{kHz}$ is chosen. Combined with the desired frame rate of $f_{\mathrm{FPS}} = 20\,\mathrm{Hz}$, this results in a number of accumulated laser pulses per histogram of $N_{\mathrm{M}} = 2000$, which is a high number for ToF systems. Assuming a relatively long pulse of 10 ns puts the estimated eye-safe optical power at $P_{\mathrm{T}} = 815\,\mathrm{W}$. Because of the short maximum range of the canal robot, a high PRF of 200 kHz can be chosen at a temporal pulse width of $t_{\mathrm{P}} = 1\,\mathrm{ns}$. At this rate, Accessible Emission Limit (AEL) is limited through the mean optical power. The high PRF lets the system accumulate $N_{\mathrm{M}} = 20000$ laser pulses per histogram without compromising FPS requirements. The automotive system uses a low PRF of 25 kHz with temporal pulse width $t_{\mathrm{P}} = 5\,\mathrm{ns}$. In these conditions, 230 W can be safely emitted and a total of $N_{\mathrm{M}} = 1000$ laser shots can maximally be detected

per histogram without compromising frame rate requirements.

The scenarios shown here are discussed without the use of post-processing via filters. They are there to give the reader an idea of the influence of different system parameters on the results of ToF measurements. For all scenarios, the atmospheric transmission efficiency η_{A}^2 is assumed to be unity because of the small distances considered. All SPADs have a dead time of $t_{\mathrm{D}} = 20\,\mathrm{ns}$ and a DCR of $r_{\mathrm{DC}} = 10\,\mathrm{kHz}$. One aspect through which the application directly influences the system parameters is dictating allowed system cost. A system used in consumer applications will have a smaller budget and post-processing steps like integration of microlens arrays for improving effective fill factor [69] can thus be too expensive, resulting in lower sensitivity. Also, selecting devices with a low number of hot pixels or low DCR during production leads to detector performance scaling somewhat with device cost. Another measure increasing cost and reducing SPAD noise behavior is temperature stabilization or cooling. Cost can also have an impact on the employed optics. Plastic lenses can be manufactured more cheaply and also weigh less than glass lenses which can be interesting in applications where overall system weight is limited, like the system mounted to a robot arm. For the two robotics systems, plastic lenses are assumed with a transmission efficiency of $\eta_{\mathrm{Opt}} = 40\,\%$. The fill factor assumes a standard value of $\eta_{\mathrm{FF}} = 20\,\%$. The automotive scenario uses glass optics with $\eta_{\mathrm{Opt}} = 50\,\%$. A microlens array increases the effective fill factor to $\eta_{\mathrm{FF}} = 80\,\%$. The system for the industrial robot employs a typical optical bandpass filter passband of $\Delta_{\mathrm{BP}} = 40\,\mathrm{nm}$. For the canal robot, an even broader passband $\Delta_{\mathrm{BP}} = 60\,\mathrm{nm}$ and with that cheaper filter is sufficient. For the automotive scenario, ambient light suppression is key and a more expensive and optimized value of $\Delta_{\mathrm{BP}} = 10\,\mathrm{nm}$ is modeled.

Different bin sizes are employed. The broadest bins with $t_{\mathrm{Bin}} = 0.5\,\mathrm{ns}$ are used in the automotive scenario. The high range relaxes the accuracy requirements and the system can benefit from the broader integration time. The canal robot uses bins with $t_{\mathrm{Bin}} = 0.1\,\mathrm{ns}$ as the covered distance is small and the signal is usually high enough to be detected with shorter bins. In the middle between the two lies the industrial scenario with a bin width of $t_{\mathrm{Bin}} = 0.25\,\mathrm{ns}$, constituting a compromise between evaluability and precision.

To gain insight into the scene with a given spatial resolution, the FoV has to be

sampled using a sufficient number of pixels. Taking the canal robot as an example which has a given wide FoV of $120° \times 10°$ and a range of $d = 2\,\mathrm{m}$, the system images a square area of 6.92 m by 0.71 m. For a spatial resolution of one millimeter, a pixel matrix of 6920×710 would be necessary. Also, considering typical SPAD diameters, these geometrical relations can be too large for photomask sizes in production. This shows a scaling problem of high FoV, high-resolution systems. Challenges arise for routing, readout design, and pixel uniformity. To date, arrays with these pixel counts were not realized, with the highest pixel count being 1 Megapixel, reported by Canon in 2020 [70]. For the industrial robot, a SPAD diameter of $d_{\mathrm{Pix}} = 20\,\mu\mathrm{m}$ is used as a typical value. As the aim of the canal robot is high resolution, optimized smaller SPADs with a diameter of $d_{\mathrm{Pix}} = 10\,\mu\mathrm{m}$ are modeled. The high ambient illuminance in the scene of the automotive scenario requires additional sensor capabilities. For that, a multi-pixel structure of 2×2 SPADs with an individual diameter of $d_{\mathrm{Pix}} = 20\,\mu\mathrm{m}$ capable of coincidence and ME detection is considered. The total pixel diameter is then $d_{\mathrm{Pix}} = 40\,\mu\mathrm{m}$ for the automotive scenario. The automotive scenario pixel structure enables additional sensing modes. The pixel can record the first four ME levels and coincident events with up to coincidence depth four. The coincidence time window t_{C} is matched to the temporal laser pulse width of 5 ns.

A summary of all chosen system parameters for the respective simulation scenario is given in Table 4.2.

The resulting source event rates for each scenario are shown in Figure 4.5. The available power budget for the automotive application is the highest because of the big pixels and high optical fill factor stemming from the more costly system optimizations previously discussed. For the automotive and industrial scenario, an event rate region is formed between the high and low object reflectivity that objects of arbitrary diffuse reflectivity can assume. This event rate region is indicated by the shaded area between the respective applications' upper and lower event rate limits. The canal robot is an exception because it is optimized for a single target reflectivity and thus does not exhibit an event rate range. The ambient event rates are constant over distance for a given object and ambient illuminance. Generally, the higher the ambient influence, the smaller the range. The industrial robot is specified at a maximum ambient illuminance of $E_{\mathrm{V}} = 20\,\mathrm{klx}$ and for the target with $\rho = 80\,\%$ has an ambient event rate of $r_{\mathrm{B}} = 10.05\,\mathrm{MHz}$ impinging on

	Parameter	Symbol	Value			Unit
			Auto	Industry	Canal	
Laser	Peak Optical Power	P_{T}	230	815	240	W
	Pulse Width (FWHM)	t_{P}	5	10	1	ns
	PRF	f_{PR}	25	100	200	kHz
	Operating Wavelength	λ_0	905	905	650	nm
Detector	Fill Factor	η_{FF}	80	20	20	%
	PDP	η_{PDP}	10	10	20	%
	Bin Size	t_{Bin}	0.5	0.25	0.1	ns
	Pixel Area	A_{Pix}	40×40	20×20	10×10	$\mu\mathrm{m}^2$
	SPADs per Pixel	N_{S}	4	1	1	
Optics	Bandpass Filter Width	Δ_{BP}	10	40	60	nm
	F-Number	$f_{\#}$	1.3	1.3	1.3	
	Optics Transmission	η_{Opt}	50	40	40	%

Table 4.2: Example DToF system parameter sets.

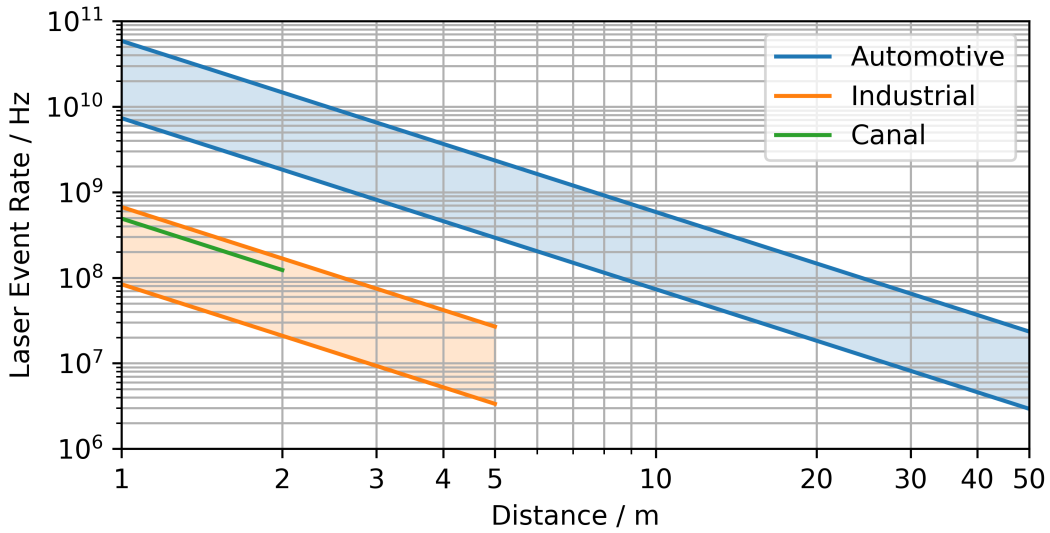
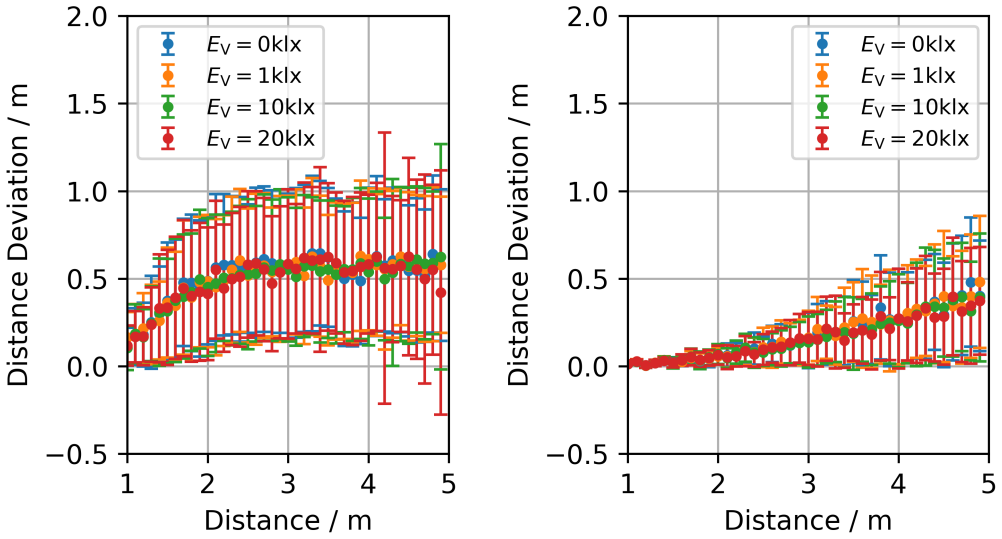


Figure 4.5: Laser event rates for the three applications. In the automotive and industrial applications, shaded areas indicate the area between event rates for $\rho = 10\,\%$ and $\rho = 80\,\%$. In the canal scenario, only objects with $\rho = 80\,\%$ are present.



(a) Results for low reflectivity $\rho = 10\,\%$.

(b) Results for high reflectivity $\rho = 80\,\%$.

Figure 4.6: Simulated accuracy (points) and precision (error bars) for the industrial scenario. High values are assumed because no post-processing is used.

the sensor surface. The strongest ambient influence is present in the automotive scenario. The total rate impinging on the pixel is also specified as reflected by a target with $\rho = 80\,\%$, but pixel size, specified illuminance of $E_V = 50\,\text{klx}$, and overall system sensitivity are higher. The total event rate impinging on the employed 2×2 SPAD structure is thus $r_B = 110.5\,\text{MHz}$. The lowest ambient event rate is present at the surface of the canal robot and equates to $r_B = 1.03\,\text{MHz}$.

4.3.2 Results

Now precision and accuracy for the different scenarios are estimated. They are only portrayed if the measurement was successful. A measurement is defined as 'successful' if the derived ToF is temporally close to the pulse return. This is the case if it is less than $\pm 5\,\%$ away from the ideal ToF or within a FWHM temporal laser pulse width.

Industrial Robot

The simulated precision and accuracy for the industrial robot are shown in Figure 4.6. Figure 4.6 a) displays the results for a diffusely reflecting e.g., black

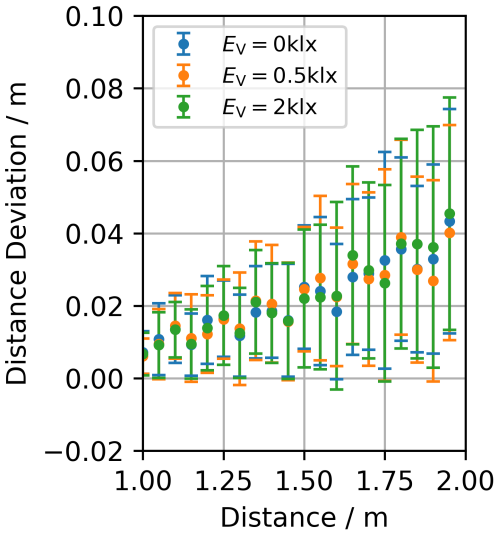


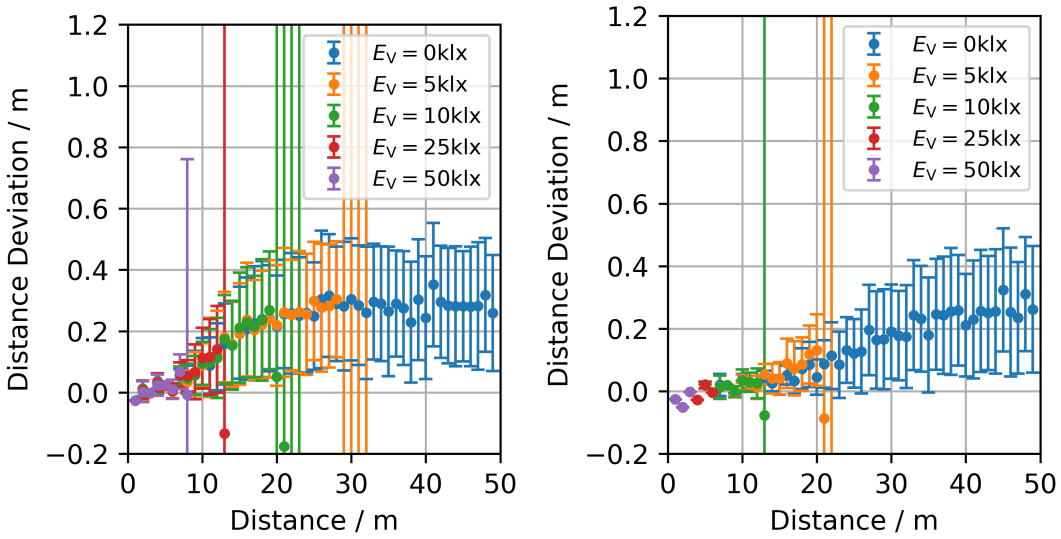
Figure 4.7: Simulated accuracy (points) and precision (error bars) for the canal robot.

paper, target with $\rho = 10\,\%$. Figure 4.6 b) shows the results for a diffusely reflecting target with $\rho = 80\,\%$. No post-processing is employed and the distance is derived from the maximum bin of the histogram. This leads to relatively high standard deviations because arrival times can fluctuate over the temporal pulse width. Error bars denote the standard deviation of the derived distance, which correlates mainly to this wandering over the temporal pulse width. One can see that the standard deviation is in the range of meters for both object types, with the white target exhibiting lower mean distance deviation and also lower standard deviation. At high rates, the pulse distribution in the histogram is more narrow and derived distance fluctuates less like laid out in Section 2.5.2.

This example is used to show how temporal laser pulse width influences ranging precision when not employing post-processing and the negative influence of pile-up effects regarding differences in object reflectance.

Canal Robot

The results for the canal robot are shown in Figure 4.7. Distance deviation is smaller than in the industrial robot case, achieving precision and accuracy in the centimeter regime. There are multiple reasons for this difference in accuracy between the two robots. One is the increased number of samples. Low require-



(a) Results for low reflectivity $\rho = 10\,\%$.

(b) Results for high reflectivity $\rho = 80\,\%$.

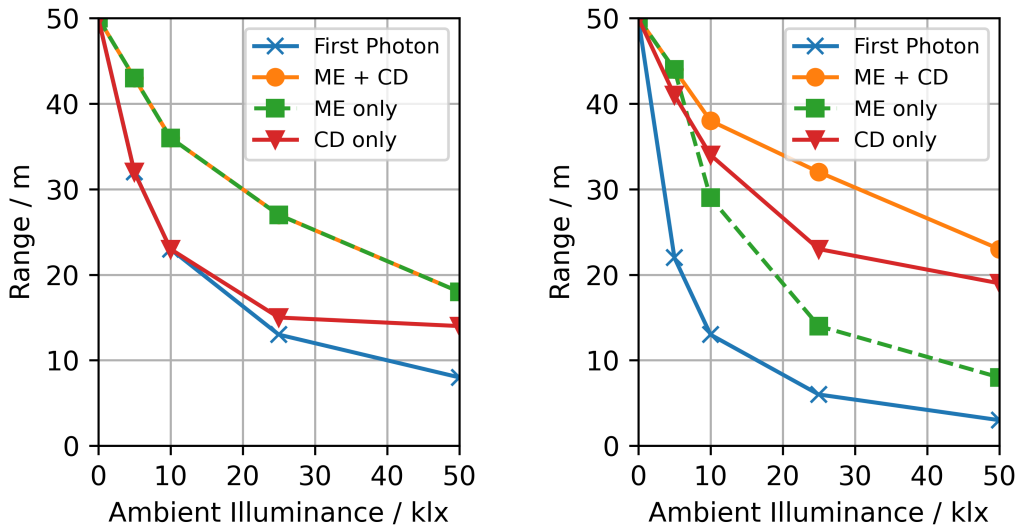
Figure 4.8: Simulated accuracy (points) and precision (error bars) for the automotive scenario.

ments on frame rate and the available high laser PRF enable the acquisition of $N_M = 20000$ samples per histogram. The reduced laser pulse width of $t_P = 1\,\text{ns}$ means that signal correlated photons will only accumulate over about 15 cm. While the overall power budget is a bit smaller because of the reduced pixel size and high PRF, the small maximum distance, and high PDP also lead to a higher photon budget than at the end of the industrial robots range. As only one target is specified for the system, calibration of distance-dependent measurement error is possible, potentially further increasing achievable precision. This scenario is used to illustrate the gain in precision reachable by limiting object characteristics, using short temporal laser pulse width and increasing the number of laser pulses per histogram.

Automotive Scenario

The achievable simulated precision and accuracy for two types of objects sensed with the automotive scenario system are shown in Figure 4.8.

The precision is determined mainly by pile-up effects. At small distances, returned laser power is high and the derived ToF is close to the ideal value. Also, accuracy is high because the first-photon PDF is narrow. At higher distances,



(a) Range gain for low reflectivity (10 %)

(b) Range gain for high reflectivity (80 %)

Figure 4.9: Achievable range in the best multi-pixel configurations depending on multi-event and coincidence level. The combination of ME and Coincidence Depth (CD) achieves the highest range.

the results for low and high reflectivity are similar. Precision is reduced and approaches 30 cm, which roughly corresponds to half a laser pulse width. Short before the respective scenarios are not successful anymore, it is possible that simulations in which ambient light was detected are included in the mean values so that higher values are assumed.

Figure 4.8 b) shows the achievable range for higher reflectivity objects. Without any ambient illumination, 50 m range is reached in all sensor configurations. At an ambient level equivalent to $E_V = 5$ klx, the measurement on the high reflectivity object is performing worse than the low reflectivity one. Through the use of the ambient mitigation sensor modes, the high reflectivity object outperforms the low one. The same is true for the higher ambient illumination levels. A big difference is that here the combination of coincidence and ME detection leads to best performance with coincidence alone also exceeding the performance gain through multi-event detection.

Figure 4.9 summarizes the gains through the best combination of coincidence and ME configurations. In blue, the first-photon response is shown. Here, rapid saturation through high ambient illumination can be seen and is expressed in

a reduction in the achievable range. This corresponds to the range from the previous Figure 4.8.

Coincidence detection can increase the achievable range, especially in high ambient scenarios. Without significant noise influence, the improvement is very marginal. For low reflectivity objects, shown in Figure 4.8 a), a higher gain is achieved by using ME detection. The additional employment of coincidence detection does not improve achievable range anymore. Figure 4.8 b) shows that both ME and coincidence detection increase achievable range significantly. The combination of the two can detect highest ranges and most sunlight dominated environments.

This example illustrates two key issues. On the one hand, achieving long range is limited by the photon budget and low reflectivity objects with low returned laser power, which could be returned by obscure, non-cooperative targets. In these operational conditions, the combination of coincidence and ME detection does not provide any advantage and when choosing between the two, multi-event detection provides higher gains. On the other hand, this combined gain can be shown for high reflectivity and thus higher laser power scenarios. For the worst first-photon scenarios, range can be increased by up to 700 %.

Chapter 5

Evaluation Schemes

The previous chapter shows that even if the laser pulse can successfully be identified from the histogram, deriving ToF from the bin with the highest count value has limited precision. This chapter introduces different digital evaluation schemes for DToF histograms and discusses how to calculate them from a given histogram. It is organized in multiple sections. In Section 5.1 two methods to derive ToF from the measurement histogram are laid out and compared. Then in Section 5.2 and Section 5.3, pre-processing steps to increase histogram SNR with correlation filters and a novel rank-order filter-based approach are discussed. Their performance is assessed with simulation in Section 5.4. Finally, in Section 5.5, a method for correcting pile-up effects on ToF is laid out and evaluated through photonic simulation.

5.1 Time of Flight Estimation

Estimating the ToF means identifying a pulse signature in the histogram. With that information, a time corresponding to the distance of the object during the recording of the histogram can be derived. In the following, two basic techniques for ToF estimation are laid out and their differences are discussed.

5.1.1 Maximum Detection

One common approach for histogram evaluation is to derive ToF from the bin accumulating the maximum number of successful detections. Benefiting from the fact that during the laser pulse return, even with background present, both laser

and background intensity are summed up [71]. In this part of the histogram, the highest optical intensity is present.

The exponential behavior of first-photon detection in equation (2.22) invalidates this strategy when significant background light is present. There, high ambient and dead time lead to the maximum intensity caused by laser and background not producing a global maximum inside the histogram and appropriate measures have to be taken to guarantee sensitivity. This can be done by compensating pile-up in the recorded histogram, limiting incident backlight power on a system side or implementing ambient light suppression control algorithms, compare Section 3.2.

However, if there is a significant probability that the pulse return photons lead to a local maximum inside the histogram, this scheme can be used to calculate t_{ToF} from the index of the maximum bin i_{Max} and its temporal bin width t_{Bin}.

$$t_{\mathrm{ToF}} = \frac{c_0}{2} \cdot i_{\mathrm{Max}} t_{\mathrm{Bin}} \tag{5.1}$$

How probable it is to detect the maximum count value out of N_{M} measurements in a bin with number n_{i} can be calculated using the binomial distribution P_{B} from the probability of detecting more counts than a threshold i_{Thresh} inside the current bin and less than that in all other bins [58] to

$$P_i = \prod_{i_{\mathrm{L}}=0}^{N_{\mathrm{M}}} \left(P_{\mathrm{B}}(i_{\mathrm{L}}, N_{\mathrm{M}}, p_i) \prod_{\substack{i_{\mathrm{Bin}}=1 \\ i_{\mathrm{Bin}} \neq i}}^{N_{\mathrm{Bin}}} \sum_{k=0}^{i_{\mathrm{L}}-1} P_{\mathrm{B}}(k, N_{\mathrm{M}}, p_{i,\mathrm{Bin}}) \right). \tag{5.2}$$

The probability of finding the maximum for each histogram bin can be calculated with this expression. Note that this formulation neglects the case of maxima with the same counting value and is only valid for high triggering probability during the measurement, meaning that $F_{\mathrm{S}}(t_{\mathrm{Hist}}) \approx 1$.

When ambient light causes a count value higher than the laser pulse return, the system puts out a wrong distance value. If the same maximum count value occurs multiple times, the distance result becomes ambiguous.

The expected value for ToF of repeated measurements, in which maximum detection is used, can be given as

$$E(P_{\mathrm{max}}) = \frac{\sum_{i=0}^{N_{\mathrm{Bin}}} (\frac{1}{2} + i) t_{\mathrm{Bin}} P_i}{\sum_{i=0}^{N_{\mathrm{Bin}}} P_i}, \tag{5.3}$$

but a single measurement is limited to the TDC resolution. The advantage of maximum detection is that it is easily implementable and does not require calculation.

5.1.2 Interpolation

In general, measurement resolution is limited to the TDC resolution t_{Bin}. One approach which can overcome the TDC resolution on the basis of a single histogram is interpolation. The expected value of distance retrieved by interpolating depends on the triggering probability per bin p_i following equation (2.28) and is given by

$$E(P_{\text{mean}}) = \frac{\sum_{i=0}^{N_{\text{Bin}}} (\frac{1}{2} + i) t_{\text{Bin}} p_i}{\sum_{i=0}^{N_{\text{Bin}}} p_i}. \tag{5.4}$$

The performance under ambient light illumination can be improved by choosing only a small portion of the histogram to prevent interpolating between laser and found ambient counts. One way of restricting the interpolation to the generated laser counts is by first localizing the maximum of the distribution and cutting out a part of the histogram around it matching to the laser pulse width or shorter.

The result of comparing the two evaluation schemes, applied on simulated sech2 distributed laser pulse shapes, is shown in Figure 5.1. Mean estimated ToF is given as points and fluctuation of derived values as error bars. Both schemes successfully detect the ToF in the given scenario but also show a systematic, rate-dependent error that could theoretically be compensated for, e.g., using the technique introduced in Section 5.5. While the systematic error of maximum detection is smaller than that of interpolation, the standard deviation of distance results is higher. This is caused by the quantization noise in maximum detection. The value for ToF of a single measurement can only assume integer-valued bin numbers. Interpolation can partially circumvent this quantization noise and shows precision smaller than determined by a single bin size.

5.2 Finite Impulse Response Filter

By post-processing histograms with FIR filters, pulse signatures can be recovered which are too faint to be recognized by the algorithms introduced in the previous

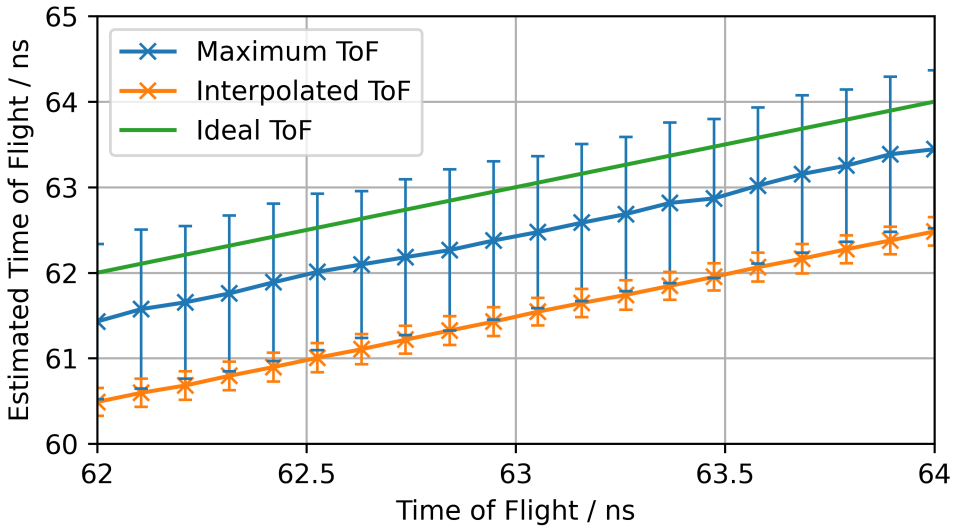


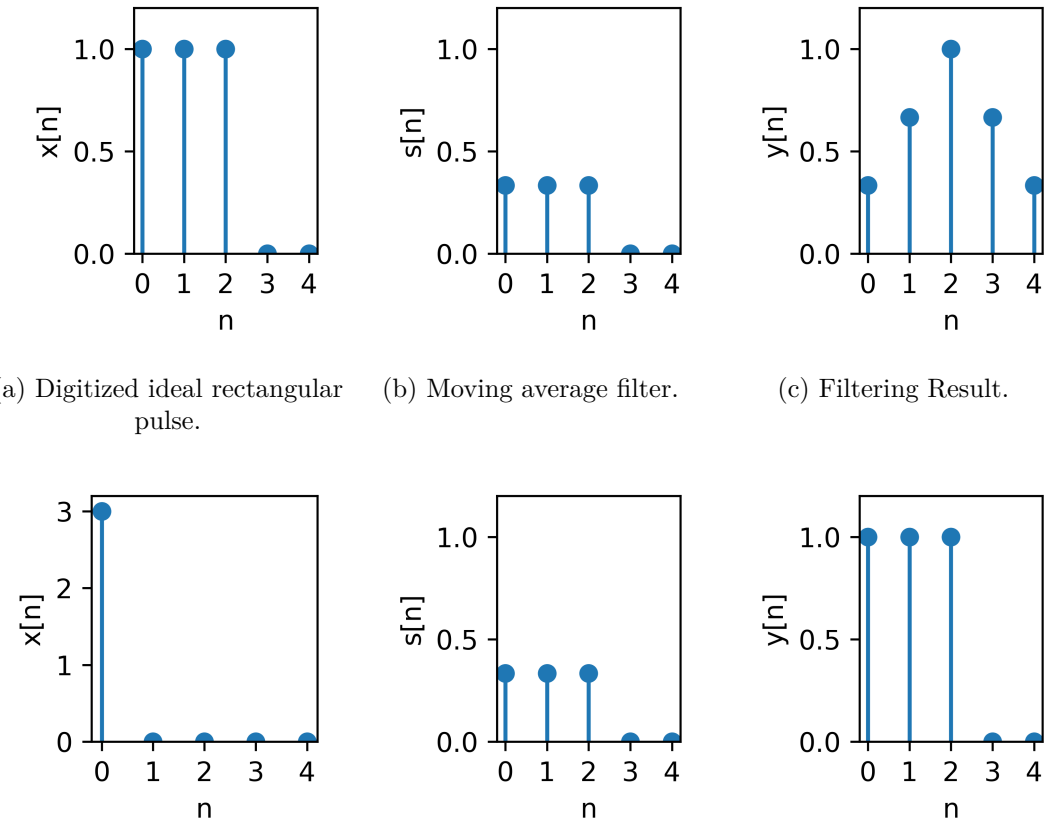
Figure 5.1: Mean and standard deviation of ToF estimation of a simulated sech2 pulse with $\tau_{\mathrm{P}} = 4.53\,\mathrm{ns}$ over different values of ToF via interpolation and maximum detection at a bin size of $t_{\mathrm{Bin}} = 1\,\mathrm{ns}$, laser event rate of $r_{\mathrm{L}} = 100\,\mathrm{MHz}$ and $N_{\mathrm{M}} = 1000$ accumulations. Standard deviation is calculated from the results of a total of 1000 simulation trials.

section. As signals from close targets return with high power, they are easily identifiable in the histogram. Those which stem from low reflectivity objects at high distances produce these faint histogram signals. For them, little pile-up distortion is present and a filter matched to the original pulse form promises good results. As previously laid out, pulse signature is not only intensity-dependent and thus easily predictable but suffers from dead time caused pile-up distortion. Application of a filter matched to the pulse signature is not possible for all event rate scenarios without estimating event rate.

The output of a time-discrete FIR filter $y[n]$ is determined by the convolution of the present signal $x[n]$ and the filters taps $s[n]$.

$$y[n] = s[n] * x[n] \tag{5.5}$$

Histogram processing by filtering is currently distributed widely. One approach is to use low-pass filters in either Infinite Impulse Response (IIR) [72] or FIR [73] implementation, but rank-order filters like median have also been reported. Now, two types of FIR filters are discussed.



(a) Digitized ideal rectangular pulse.

(b) Moving average filter.

(c) Filtering Result.

(d) Digitized dirac pulse.

(e) Moving average filter.

(f) Filtering Result.

Figure 5.2: Visualization of filtering a rectangular pulse with a moving average filter matched to temporal pulse width.

5.2.1 Moving Average

A moving average filter is a type of low-pass filter. For the rectangular pulse model and low rates, a moving average with a filter width equivalent to the temporal laser pulse width is shown in Figure 5.2 functions like a matched filter [57]. The result forms a maximum value when both filter and pulse form correlate temporally.

When high rates are present, pile-up leads to a narrowing of the temporal pulse form in the histogram. The edge case, assuming that IRF is smaller than a bin width Δt_{Bin}, is a temporal pulse form of a Dirac-impulse. The realized and rate-dependent laser distribution thus assumes different shape and the output $y[n]$ is shifted depending on event rate. Even after employing this filter, a pile-up-driven deterioration of precision is present.

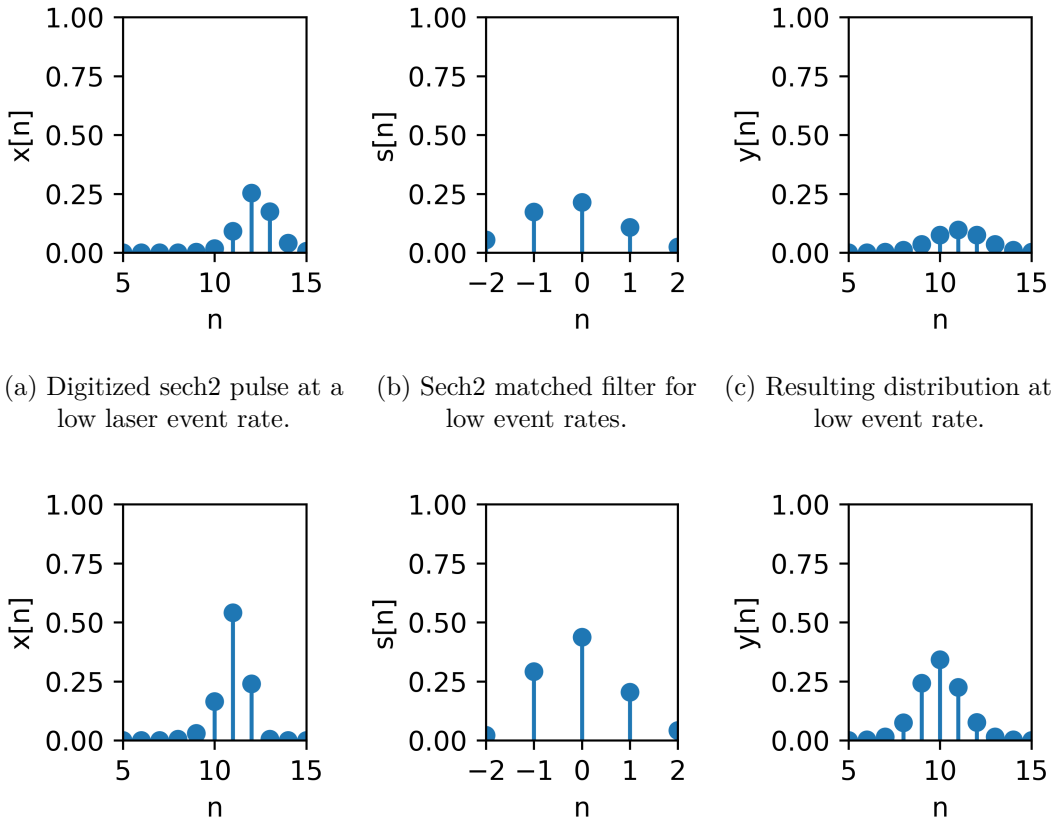
5.2.2 Matched Filter

Matched filters are used in signal processing to maximize SNR for a given realization of a signal buried in Additive White Gaussian Noise (AWGN). When evaluating a DToF histogram, finding the pulse position is the most important task. As the laser pulse distribution without pile-up distortion can be recorded in a characterization measurement, it can be used to construct a correlation filter by time reversal. The laser pulse distribution is recorded and inverted to retrieve the filter characteristic $s[n]$. A resulting filter kernel for $s[n]$ and expected values for the signal in the histogram with $N_{\mathrm{M}} = 1$ are shown in Figure 5.3 for two different laser return powers. From the respective pulse distributions, a matched filter can be derived. Without knowledge about the returned laser power, it is unclear how the filter kernel must be constructed to find the pulse information from the histogram ideally. Following the fact that the low return power is the more challenging case; the structure can be used for all cases with reduced performance in the high power scenario.

In a measurement, realizations will differ from the ideal expected values and worsen filter performance depending on the size of the accumulated statistical basis. Matched filter performance in SPAD DToF systems is limited by two factors. On the one hand, noise is not white because of the exponential background event shape in the histogram and because the counting values are not following a Normal but a Poisson distribution. In the first histogram bins, the highest noise is present with high variance. On the other hand, temporal laser pulse distribution depends on the intensity, and thus, direct correlation with a pulse form prediction is complicated and requires rate estimation. However, as finding the temporal pulse shape in the histogram is especially challenging for low return power, a matched filter fit to the original pulse form can be good enough.

5.3 Rank-Order Filters

Rank-order filters apply a sorting logic when determining filter output. They are computationally efficient in that they require no calculation to find an output value but simply choose a value that is already part of the filtered array. However, they require the use of hardware-compatible and area-efficient sorting structures.



(a) Digitized sech2 pulse at a low laser event rate.

(b) Sech2 matched filter for low event rates.

(c) Resulting distribution at a low event rate.

(d) Digitized sech2 pulse at a high laser event rate.

(e) Sech2 matched filter for high event rates.

(f) Resulting distribution at a high event rate.

Figure 5.3: Effect of pile-up and matched filter on histogram distribution. Original pulse form, matched filter, and convolution result are displayed. To generate the filter structure, $t_{\mathrm{Bin}} = 0.4\,\mathrm{ns}$, $\tau_{\mathrm{P}} = 4.53\,\mathrm{ns}$ and $r_{\mathrm{L},0} = 0.1\,\mathrm{GHz}$ or $r_{\mathrm{L},0} = 1\,\mathrm{GHz}$ respectively.

Figure 5.4: Sequence diagram of the minimum filter routine.

5.3.1 Sliding Minimum Filter

The output of the sliding minimum filter is determined by the lowest occurring value inside its filtering window. The basic structure of the sliding minimum filter routine is shown in Figure 5.4.

Again, the expected count value k_i depends on N_M and the respective triggering probability during a bin is determined by the CDF F_S.

$$k_i = N_\mathrm{M}[F_\mathrm{S}((i_\mathrm{Bin}+1)\cdot t_\mathrm{Bin}) - F_\mathrm{S}(i_\mathrm{Bin}\cdot t_\mathrm{Bin})] \tag{5.6}$$

When subtracting a pure ambient distribution from the measurement histogram, which is similar to correlated double sampling in analog detectors, the mean value of ambient counts is compensated for in these bins. The probability for a bin containing a filter output value greater than zero is determined by the filter width w_min and the count values of the bins neighbors in the filter width.

$$P_i = \prod_{k=i}^{i+w_\mathrm{Min}} \sum_{k=\lceil N_\mathrm{M} p_i \rceil}^{N_\mathrm{M}} P_\mathrm{B}(k, N_\mathrm{M}, p_i) \tag{5.7}$$

The probability of detecting a background light event gets smaller with increasing filter size. As the probability per bin is constant, the probability of false alarm does not depend on the absolute values of the ambient distribution. Over the histogram, a constant false alarm rate is achieved. The same condition for count values has to hold for the pulse to be detected. Contrarily to the pure background though, expected count values during the laser pulse return are increased. The detection probability for pulse return is thus higher than the false alarm rate.

5.3.2 Median Filter

The most common rank order filter is the median filter. There are many variants with wide application in image denoising [74], extending the pure rank-order approach of a simple median with mixed forms containing decision criteria and blending with FIR filters. The used variant in this work, however is the classical median filter in which values in the filter window around a center index are first sorted and then the middle value of the window is kept as the center indexes' value.

5.3.3 Quantile Filter

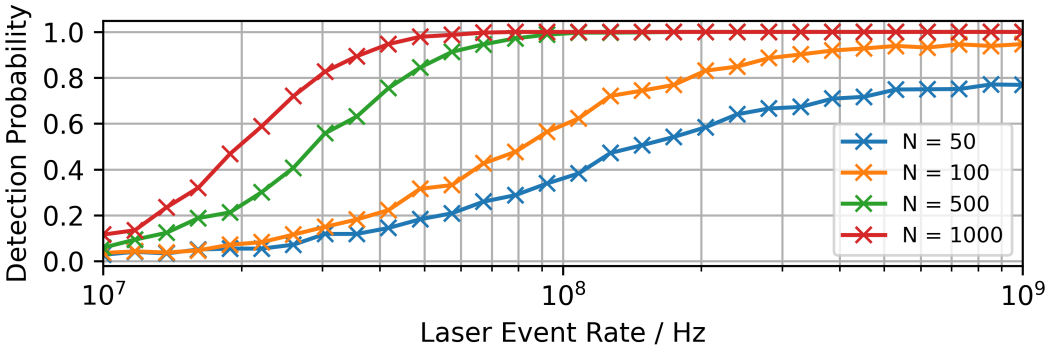
One more general modification of rank-order filters and focus of the following evaluation are the class of quantile filters [75]. They are similar to a sliding median or minimum filter, but the value position that they grab out of the window, called 'tap' after sorting is configurable.

Generally, the attenuation of ambient light is weaker than in the sliding minimum filter case. However, pulse attenuation is also less drastic and potentially ToF can be retrieved at weaker signal strength.

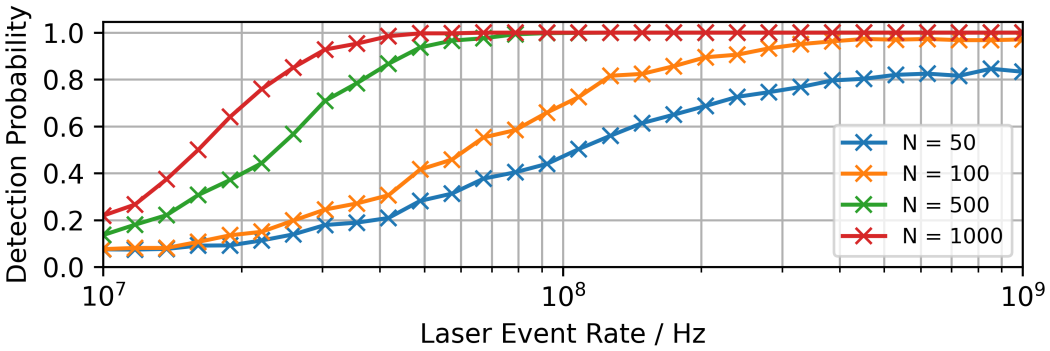
5.4 Filter Comparison

To estimate the performance of the different filter structures, photonic simulation as described in Chapter 4 is employed. The arrival distributions and histograms are generated using the sech2 pulse model and rejection sampling. The following considerations are based on a fixed background event rate of 10 MHz. Further parameters are $\tau_{\mathrm{P}} = 4.5\,\mathrm{ns}$, corresponding to a normal distribution FWHM of about 8 ns. The simulation results shown in Figure 5.5 are generated with $N_{\mathrm{M}} = 1000$ repetitions per plotted point.

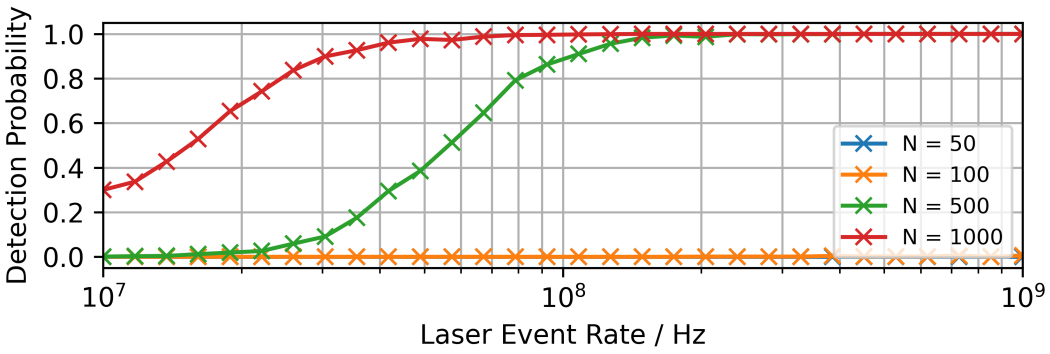
Moving average, matched, and quantile filters are compared. The considerations of the quantile filter include the sliding minimum and median filters, as they represent special cases where the lowest or middle tap of the filter is used as an output. The quantile and moving average filter use a filter width matched to the width of the arriving pulse. The results are generated for a different number of measurement repetitions N_{M} and shown in Figure 5.5. One can see that the



(a) Mean filter matched to pulse width.



(b) Matched filter.



(c) Quantil filter with a width of 20 bins and tap of 5.

Figure 5.5: Simulated detection probability for different filter structures and a number of measurements N_{M}. The best performance is achieved using matched filtering. A total of 1000 simulation trials was conducted.

best success rate overall is achieved using the matched filter structure. In this configuration, the rank-order filters suppress the occurring background and also reduce the number of instances where the pulse can be found. This results in lower detection probability across the board. In terms of detection probability for a given scenario, they are outperformed by the simple matched filter. The moving average filter is identified as an alternative to matched filtering with a simpler filter structure and easier hardware integration.

5.5 Computational Pile-Up Correction

Pile-up leads to an intensity-dependent change in histogram pulse signature. This introduces an error in ToF estimation, no matter which approach discussed in Section 5.1 is used. There is a change in occurrence time of the maximum value, an alteration of the laser rising edge, and the distance retrieved by interpolation.

This section lays out and investigates a different approach using the sech2-intensity function from equation (2.26). As derived in equation (3.8), the maximum position of the first-photon pulse PDF shifts depending on the functions rate parameter $r_{\mathrm{L},0}$. With this expression, the retrieved maximum position in the histogram t_{Max} can be corrected for, retrieving the actual time of maximum intensity t_{ToF}. However, knowledge of the rate parameter $r_{\mathrm{L},0}$ is required.

Measuring $r_{\mathrm{L},0}$ is complicated, so deriving an estimate $\hat{r}$ from the histogram is the next-best option. A straightforward option is looking at the total integral of the PDF. It gives a relationship between event rate and expected number of generated counts μ_{n} out of N_{M} measurements at a given pulse width of τ_{P}. The estimate is found by inserting the measured number of counts n.

$$\begin{aligned} r_{\mathrm{L},0} &= \frac{1}{2\tau_{\mathrm{P}}} \ln\left(1 + \frac{\mu_{\mathrm{n}}}{N_{\mathrm{M}}}\right) \\ \Rightarrow \hat{r} &= \frac{1}{2\tau_{\mathrm{P}}} \ln\left(1 + \frac{n}{N_{\mathrm{M}}}\right) \end{aligned} \tag{5.8}$$

This expression is feasible for estimating event rate but saturates as soon as each cycle detects a photon. The maximum output event rate for this estimator is yielded by setting $n = N_{\mathrm{M}}$, which means that a laser photon was detected in each measurement.

$$\hat{r}_{\mathrm{Max}} = \frac{1}{2\tau_{\mathrm{P}}} \ln(2) \approx 0.347 \tau_{\mathrm{P}}^{-1} \tag{5.9}$$

Again, this can be inserted into equation (3.8) to retrieve the maximum correctable pile-up error for integral estimation.

$$t_{\mathrm{Max}} = t_{\mathrm{ToF}} + 0.17 \cdot \tau_{\mathrm{P}}. \tag{5.10}$$

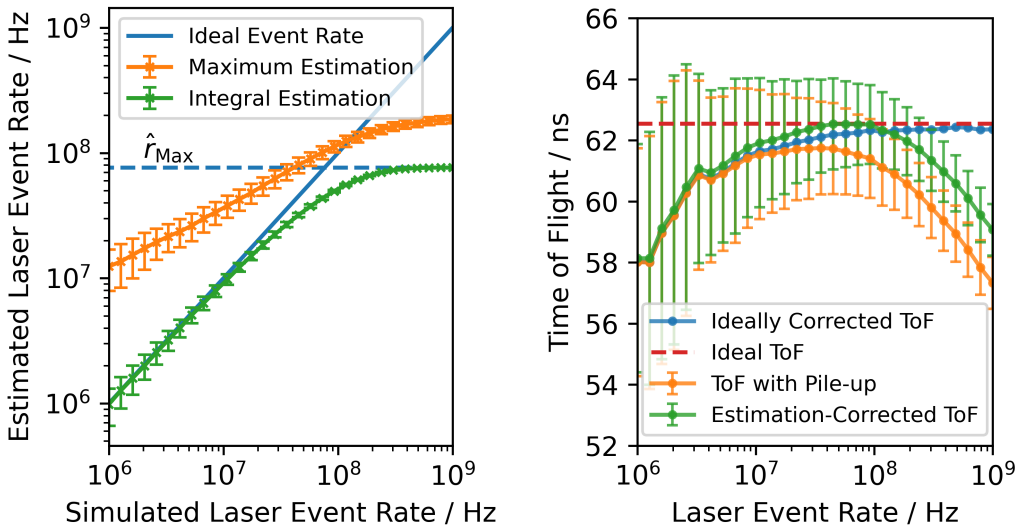
Depending on the laser pulse width, this fraction of τ_{P} can be marginal for ToF accuracy. If available, this insight further supports the attractiveness of high-power laser sources with short temporal output width. To summarize, the integral estimator helps to track laser event rate changes in low event rate or photon count regimes. This result is also logically sound because the pulse integral will saturate when a single event is detected per measurement cycle, and pile-up is an effect occurring only when multiple events are present during a pulse, but only one can be detected. Especially in high event rate regimes, estimation error becomes significant.

The previous approach used that the number of expected counts during the histogram depends on the event rate parameter $r_{\mathrm{L},0}$. However, saturation limits its usefulness. The idea now is to extend the maximum value of $r_{\mathrm{L},0}$ by estimating it from a single bin instead of the whole histogram. The condition to have all measurements fall into a single bin instead of occurring sometime during the histogram will be met at higher event rates.

The expected value of counts during the pulse maximum can be retrieved from triggering probability $p_{\mathrm{L,s}}$ from equation (3.9) and number of measurements N_{M}.

$$\begin{aligned} \mu_{\mathrm{Max}} &= N_{\mathrm{M}}\, p_{\mathrm{L,s}} \\ &= N_{\mathrm{M}} \quad \Big[\exp\Big(- r_{\mathrm{L},0}\tau_{\mathrm{P}} \tanh\Big(- \frac{t_{\mathrm{Bin}}}{2\tau_{\mathrm{P}}} - \frac{1}{2} \operatorname{arcsinh}(r_{\mathrm{L},0}\tau_{\mathrm{P}})\Big)\Big) \\ &\qquad - \exp\Big(- r_{\mathrm{L},0}\tau_{\mathrm{P}} \tanh\Big(\frac{t_{\mathrm{Bin}}}{2\tau_{\mathrm{P}}} - \frac{1}{2} \operatorname{arcsinh}(r_{\mathrm{L},0}\tau_{\mathrm{P}})\Big)\Big)\Big] \end{aligned} \tag{5.11}$$

Solving this equation for $r_{\mathrm{L},0}$ yields an estimate of the laser event rate for a given expected count value. The resulting estimated event rates of a Monte-



(a) Estimated laser event rates from simulation using integral and maximum estimation.

(b) ToF result using ideal pile-up correction with prior laser event rate knowledge, uncorrected results, and correction via maximum estimation.

Figure 5.6: Event rate estimation and ToF pile-up correction following equation (5.11). All results were generated with $N_{\mathrm{M}} = 1000$ measurements, $\tau_{\mathrm{P}} = 4.53\,\mathrm{ns}$, $t_{\mathrm{Bin}} = 100\,\mathrm{ps}$ and $t_{\mathrm{ToF}} = 62.55\,\mathrm{ns}$.

Carlo trial conducted with the sech2 pulse model is shown in Figure 5.6 a). They are given for $\tau_{\mathrm{P}} = 4.53\,\mathrm{ps}$ and $N_{\mathrm{M}} = 1000$. A total of 1000 simulation trials was conducted, and the resulting estimated event rate using the integral estimator (green) and maximum estimator (orange) are shown. Mean values are displayed with markers and standard deviation with error bars. In the low event rate regime, one can see that the integral estimator delivers values that are close to the underlying event rate. Increasing it further drives the estimator into saturation. The maximum estimator delivers an overestimation in low event rate regimes. However, at these rates, occurring pile-up is also limited. At medium event rates, the estimator follows the ideal event rate before it also saturates at slightly higher rates than the integral estimator. One can see that the maximum estimator approach is also limited. At low event rates, count values are low and count value variance is comparatively high. Here, event rates are overestimated because the highest realization from the histogram is evaluated. At high event rates, the evaluation scheme ceases to function similarly to the integral estimator.

If the temporal IRF is narrow enough, all counted values occur during the same bin. Also, the approach is sensitive to discretization effects. If t_{Max} lies at the boundary of two bins, counted values are split up, resulting in an estimation error. A narrow sliding sum filter could compensate for this.

Figure 5.6 b) shows the resulting ToF for different compensation techniques. In the low rate regime, deviations to smaller values of ToF are visible. This is caused by a low probability of detecting a distinct maximum in the histogram. Instead, a sequence of bins with either counting value '1' or '0' is formed, and as the maximum detection algorithm returns the first occurrence of multiple maxima with the same counting value, earlier ToF is retrieved. In high rate regimes, another deviation is visible. The orange graph shows pile-up behavior leading to smaller ToF. This deviation could be entirely compensated if the rate parameter was perfectly known, resulting in the blue ideally corrected ToF graph. However, only partial pile-up compensation can be achieved using the results from maximum estimation. While for event rates up to 0.2 GHz a close match to ideal ToF is shown, higher rates lead to similar deviations as no correction.

The pile-up compensation is investigated on measurement data in the following chapter.

Chapter 6

Experimental Investigation

In the following chapter, the previous theoretical predictions are studied using systems available at Fraunhofer IMS and a dedicated sensor testing setup. The working principle of the hierarchy model is shown on real measurement data. A test scene on targets with two different reflectivities is recorded, and the performance of the computational pile-up correction evaluated.

6.1 Measurement Setup

A measurement setup was built to supply reproducible signal and noise values to the Device Under Test (DUT). A schematic overview of the different components used is seen in Figure 6.1. A photograph of the realized setup is seen in Figure 6.2.

One axis holds a pulsed laser source which generates the signal in the ToF measurement. Here, an OSRAM PL90-3 laser diode [76] is driven via a PICOLAS LDP-AV 40-70 [77]. To change the signal intensity, the input current of the laser can be electrically controlled. Simulation of ambient signals is done via a continuous wave laser source in the spectral regime of the DToF system. The laser source is a Laser components QL90O7SA diode [78] with 100 mW optical output power at an emission wavelength of 905 nm. The Neutral Density Filter (NDF) bank can house a combination of filters and enables to set a wide range of possible input powers by only partially transmitting the laser power. The noise intensity can also be manipulated electrically and with NDF. A dichroic mirror is used as a beam combiner, forming the superposition of signal and ambient laser at one of its outputs. To delay the outgoing laser pulse, a Kontron PG8500 pulse

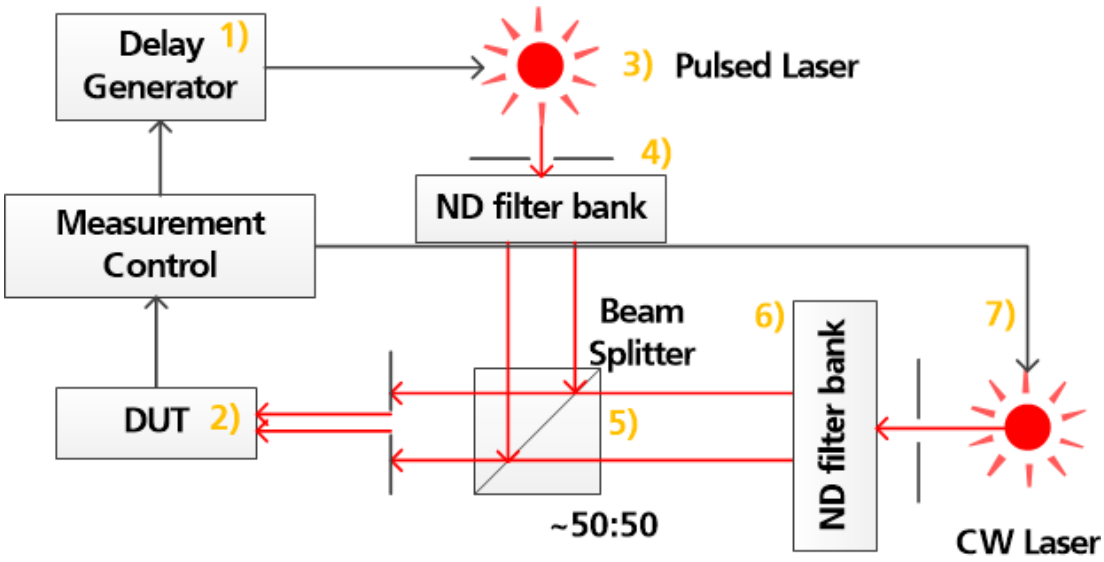


Figure 6.1: Experimental setup for testing LiDAR DUT. The DUT can be illuminated with different signal and ambient optical power set via neutral density filters and combined in a dichroic mirror.

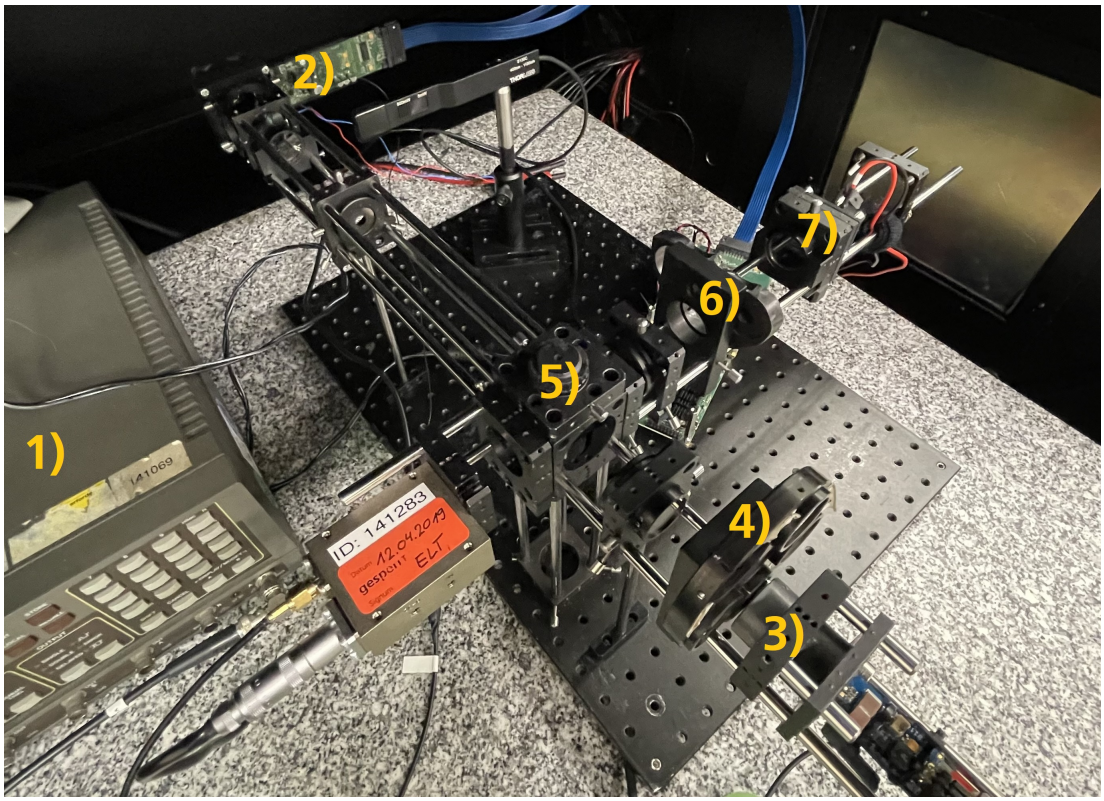


Figure 6.2: Photograph of the experimental setup.

generator is used. It is connected in the path between the control board and the laser driver.

6.2 Verification Systems

Two multi SPAD pixel systems are used for verifying the previously established performance predictions and evaluation schemes. They are used in place of the DUT in Figure 6.1.

6.2.1 Multi-Event Test Structure

This DUT is a 2×2 structure of four SPADs with a diameter of 14 µm. A block diagram of the pixel structure is given in Figure 6.3.

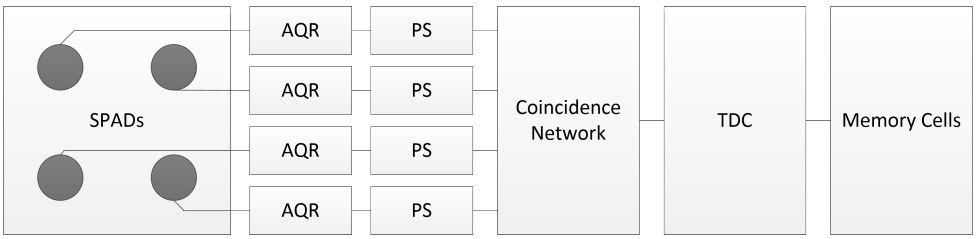


Figure 6.3: ME test pixel structure. The memory cells save multiple timestamps from the TDC.

As in most SPAD array integrated circuits, each diode is connected to an Active Quenching and Reset. The AQR exhibits a fixed dead time of about 20 ns per SPAD device. Coincidence is resolved using a configurable logic network built from OR gates between the AQR and the TDC input. The duration of pulses generated by the Pulse Shaper (PS) at the output of the AQR determines coincidence time. Connected to the coincidence circuit is a ring oscillator-based multi-hit TDC that can resolve four events per measurement cycle with a resolution of $t_{\mathrm{Bin}} = 312.5\,\mathrm{ps}$. After each event detection, it requires a hold-off time of about 18 ns.

In Figure 6.4, measured histograms recorded under constant ambient illumination are shown. The number of conducted accumulations is $N_{\mathrm{M}} = 861300$ so the results constitute long-term averages of incoming photon arrivals. Single-photon arrival levels (first arrival, second arrival and higher arrivals) are shown in different colors. The black scatter series displays the sum histogram containing all

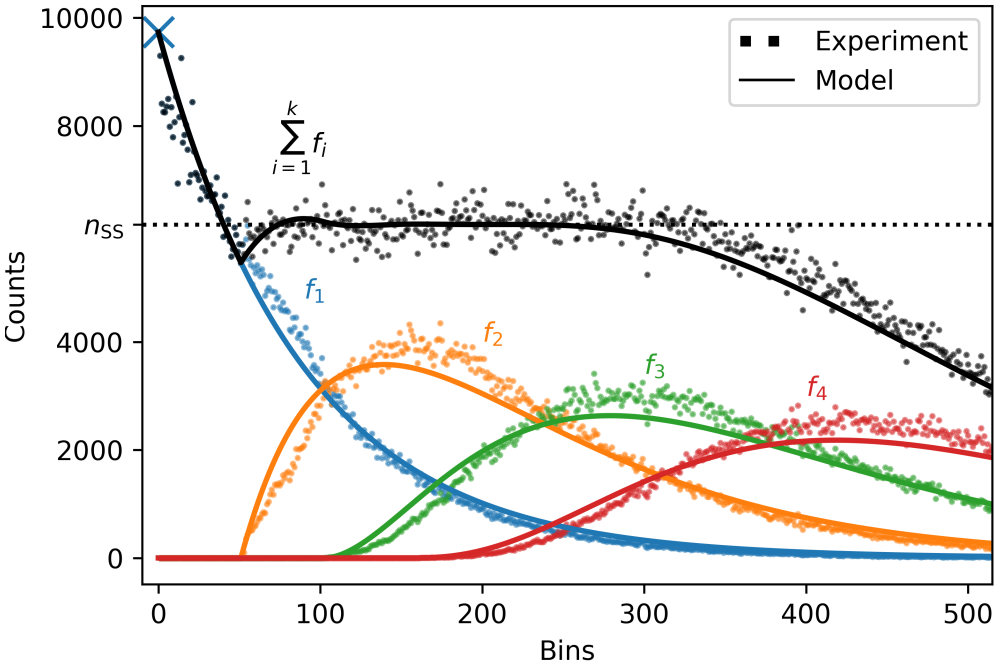


Figure 6.4: Measurement result of the ME DUT under constant ambient illumination. Colors show different ME levels. The black series shows the histogram of the sum of the first $k = 4$ events.

single event levels. A fit with a constant function was conducted between bins 55 and 300 to find the count value n_{SS} corresponding to the steady state rate r_{SS}. Using the dead time compensation from equation (2.40), a value for r_B is calculated. In the plot, it is marked with a blue cross. It fits well to the distribution of the first photon. That confirms the assumption that all SPADs are successfully readied for detection at time $t = 0$ because dead time influence did not reduce the counting values of the first triggered avalanche [79]. The value of background event rate r_B is used to calculate a simplified histogram mean distribution at constant illumination. It can be retrieved by integrating the ME PDF from equation (2.39) according to equation (2.28) and scaling it by the number of measurements. It is simplified in that it models a mutual dead time between events instead of the more realistic behavior of four SPADs with individual dead time and a TDC hold-off time. The distributions per level and a sum of all levels are given as solid lines. One can see that the measured counts rise more slowly than in the theoretical distribution. This could be caused by some SPADs already

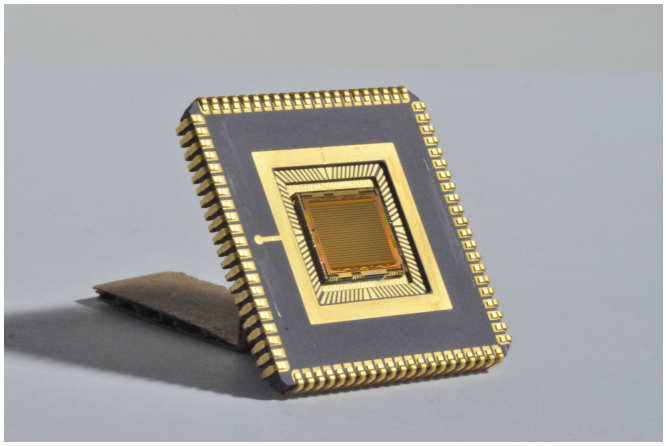


Figure 6.5: CSPADαlpha chip photograph.

having left their dead time while others take a little longer to be sensitive again. However, overall the simplification matches the distribution well.

6.2.2 CSPADαlpha

This DUT is a 2D SPAD array built at Fraunhofer IMS. It features a 64×48 grid of smart [80] pixels, each containing 2×2 SPADs. They can be used to record both intensity information through photon counting and timing information with TDC structures capable of generating time stamps. The pixels are implemented using the BSI technique. The SPAD chip is bonded to its auxiliary electronics on a dedicated readout wafer via wafer bonding, and the diodes are illuminated through their thinned substrate. It is the first implementation as a wafer-bonded BSI SPAD sensor at Fraunhofer IMS, which is why relatively area-intensive bonding vias were implemented, resulting in a FF of about $\eta_{\mathrm{FF}} = 3.6\,\%$. For the targeted operating wavelength at 905 nm, a PDP of about $\eta_{\mathrm{PDP}} = 2.4\,\%$ was determined.

The TDC uses a ring oscillator structure with a Voltage-Controlled Oscillator (VCO). It can resolve the arrival of events with a bin size of nominally 312.5 ps. The total output range is 1.28 µs which covers a range of about 200 m. Load- and geometry-dependent voltage drop-offs over the array means that pixel-individual VCO frequency can also be lower. Each counted bin will then cover an extended time range, resulting in a bigger bin size t_{Bin}. Median values for the realized bin size are around 400 ps but differ from chip to chip. To account for this difference between theory and implementation, the frequency drop of the VCOs can be

Figure 6.6: Pixel structure of CSPADαlpha pixels.

compensated for in a separate operational mode by counting the pixel clock cycles during a fixed number of master clocks (in-situ calibration). In intensity mode, all pixels are individually addressable and capable of being read out. Only half of the full resolution can be achieved in timing mode, resulting in an effective pixel number of 32×24. A chip photograph can be seen in Figure 6.5. The pixel structure is shown in Figure 6.6. It differs from the ME test structure in that it is only capable of first-photon detection and does not have associated memory cells.

6.3 Computational Pile-up Correction

Now, using the previously introduced CSPADαlpha, the pile-up correction algorithm is investigated. At first, the demonstrator is introduced, which combines the SPAD chip with optics and a laser source, and then an example scene with a planar object is used to show the advantages and limits of the approach.

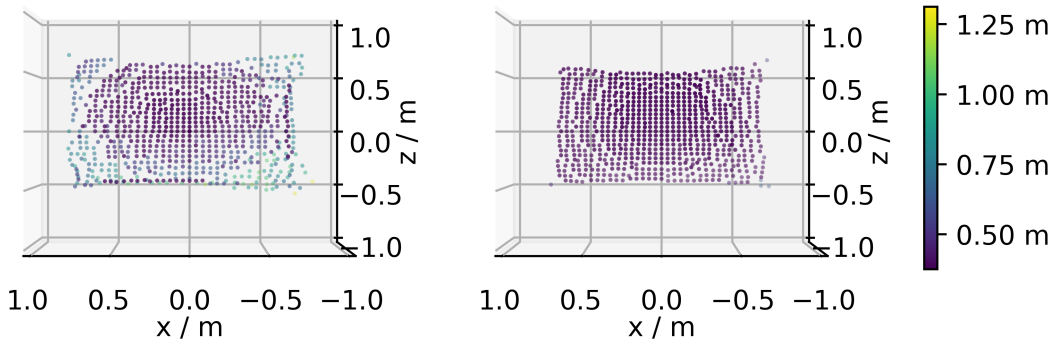
6.3.1 OWL Camera System

The OWL system is a demonstrator which complements the CSPADαlpha sensor board with two 220 W pulsed laser diodes of type LS9-220-8-S10 [81] from Laser Components and system optics for imaging the illuminated FoV. The sources exhibit a measured FWHM pulse width of about 8.4 ns. A lens system is used to illuminate a rectangular area approximately homogeneously, with its FWHM covering about 34.5° by 24.3°. The detector side optics have a focal length of $f = 12\,\text{mm}$. Incorporating the geometrical dimensions of the detector array, this equates to an observed FoV of roughly 38.2° by 29.1°. From those values follows that the outmost pixels see a slightly reduced power budget. An optical bandpass filter with transmission bandwidth $\Delta_{\text{BP}} = 30\,\text{nm}$ is integrated between sensor and

objective lens to attenuate broadband ambient influence.

6.3.2 Pile-up Correction in Example Scene

The Owl system is placed orthogonally to the target at a distance of 1 m. Two measurements were conducted at this distance, one using a target with black varnish and reflectivity $\rho = 10\,\%$ and the other using a white paper target with $\rho = 80\,\%$. The distance is evaluated from the histogram using maximum detection. The scenario is challenging for achieving high accuracy because of the low distances and the high dynamic range of returned laser power. Consequently, only partial correction of the pile-up phenomenon can be expected.



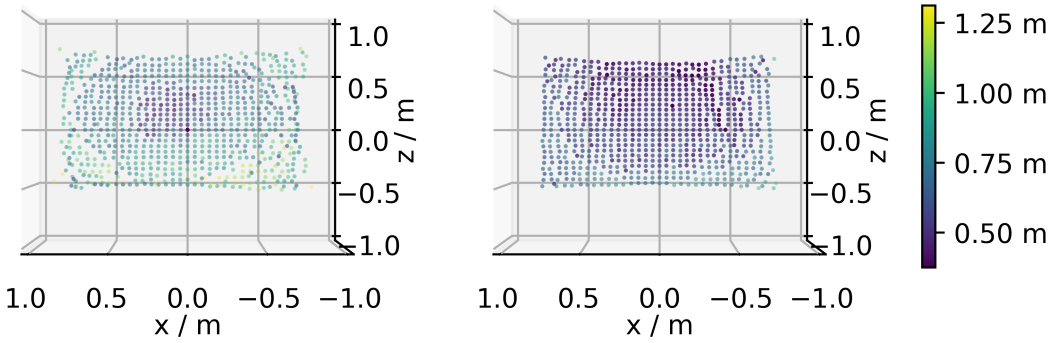
(a) Black target.

(b) White target.

Figure 6.7: Distance pointcloud of white paper (left) and black paper (right) target respectively. Each point was recorded with $N_\mathrm{M} = 1000$ pulses.

Figure 6.7 shows a point cloud recorded using $N_\mathrm{M} = 1000$ laser pulses. The resulting distances are systematically smaller because the pile-up effect emphasizes early detections. Measured distances for the black and white targets respectively are $d_\mathrm{Bl} = (0.73 \pm 0.20)\,\mathrm{m}$ and $d_\mathrm{Wh} = (0.41 \pm 0.11)\,\mathrm{m}$. The mean and variance of derived distance for the white target are both smaller because of the increase in returned laser power. However, even the black target is measured with a distance smaller than the actual distance to the target. Standard deviations of the results show a fluctuation of values in the range of tens of centimeters.

Applying the correction algorithm for pile-up sketched in Section 5.5 leads to the results in Figure 6.8. Accuracy and precision are improved over the previous



(a) Black target.

(b) White target.

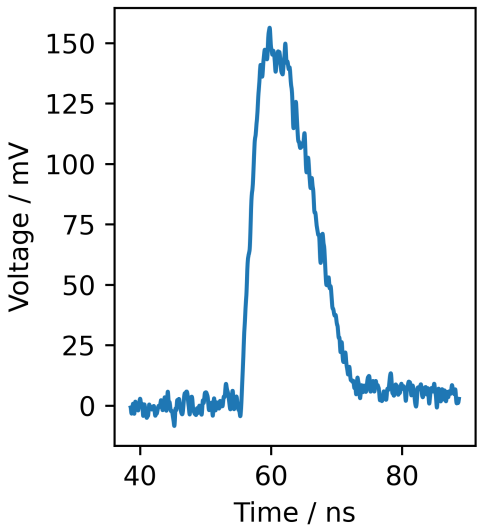
Figure 6.8: Corrected distance pointcloud of a planar target from Figure 6.7 using the correction algorithm laid out in Section 5.5.

results to $d_{\mathrm{Bl}} = (1.00 \pm 0.14)\,\mathrm{m}$ and $d_{\mathrm{Wh}} = (0.82 \pm 0.09)\,\mathrm{m}$. There is a noticeable improvement in distance deviation over the array for a given target. This hints at a reduction of reflectivity-dependent effects. At the same time, correction is only partially successful as the accuracy of the white target is still limited to about 18 cm.

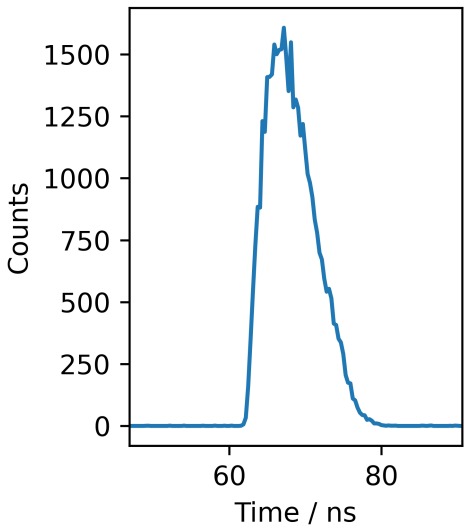
6.4 Calibration

Now the measurement setup from Section 6.1 is characterized in terms of the available signal and ambient light power using the ME test structure. Figure 6.10 shows the pulse event rates of the calibration measurement. To attenuate optical laser power far enough to imitate a diffuse reflection, multiple NDFs are employed. For all laser settings, a ND3.0 filter was used in conjunction with one additional filter level. The resulting laser event rates range from 40 to 400 MHz. The pile-up effect is visible, resulting in an intensity-dependent shift of the maximum position from its original value to the rising edge.

Figure 6.11 shows the estimated ambient light levels from cw diode illumination through three different neutral density filters. They were measured by using the ME test structure in timing mode and then conducting a non-linear least-squares fit on the resulting exponentially distributed histograms. The lowest transmission value shows a lower light level than expected. That could be the case because the

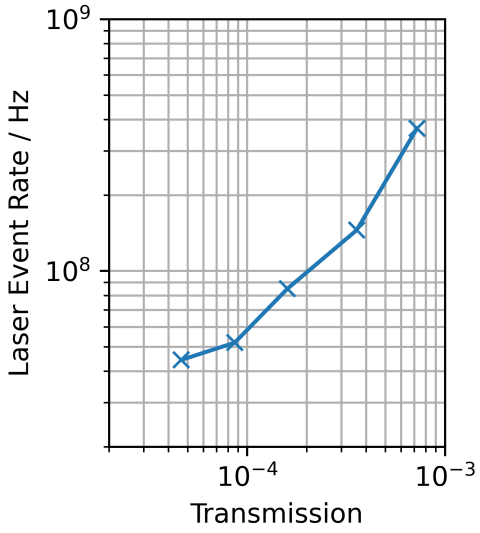


(a) Oscilloscope trace recorded using HSA-X-S-1G4-SI fast silicon photodiode [82] with 3dB-Bandwidth of 1.4 GHz.

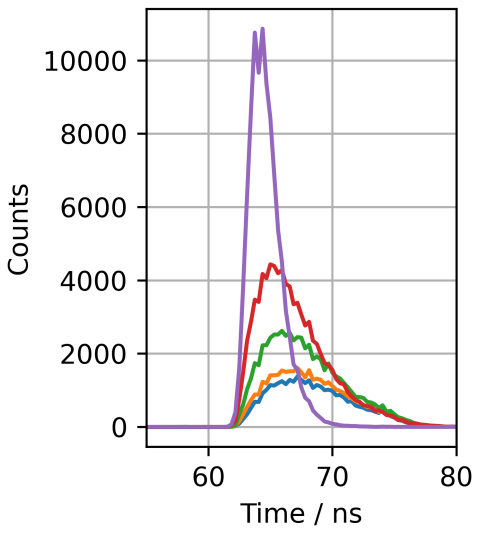


(b) Histogram recorded using ME test structure and $N_{\mathrm{M}} = 100000$ first-photon measurements at four active SPADs in the pixel.

Figure 6.9: Recorded laser trace of a SPL PL90-3 diode.



(a) Estimated laser event rates impinging on a ME DUT pixel for different neutral density filter transmissions.



(b) Photon pile-up at different laser event rates corresponding to the power values from the graph on the left.

Figure 6.10: Results from the calibration measurement of the laser source. Higher laser return power leads to narrower histogram distributions.

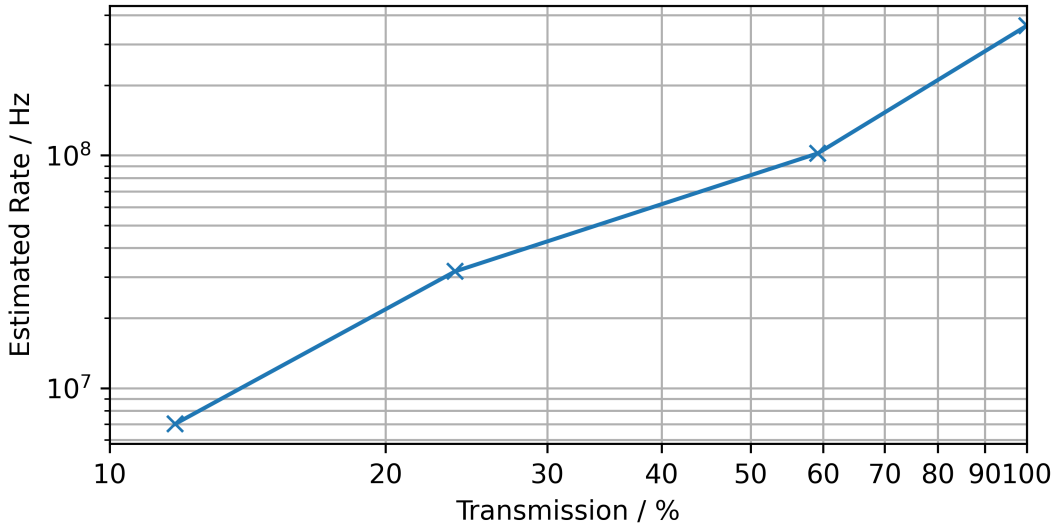


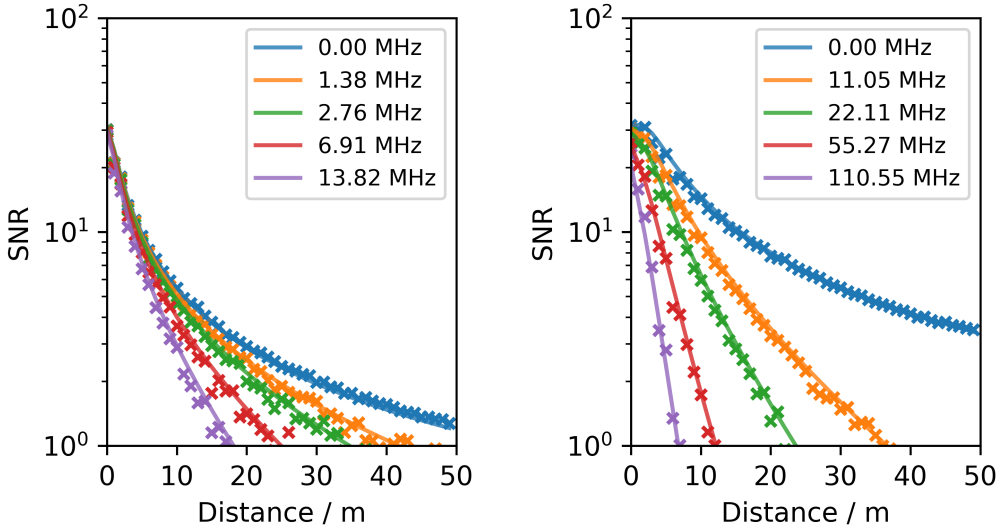
Figure 6.11: Estimated ambient event rates impinging on a ME DUT pixel for neutral density filter transmissions between 11.7 % and 100 %.

filters are reflective and thus illuminate the control photodiode inside the cw laser diode casing, so the laser output power might be lowered. Another possibility is that interference phenomena are present because of the coherent length of the used diodes and the relatively small scale of the setup.

6.5 Hierarchy Model

The theoretical considerations in Sections 3.2 and 4.3 have indicated an achievable gain by combining ME and coincident detection. This section aims to present and evaluate a possible control algorithm that has the following goal: Ensuring detection at high ambient illumination with the highest possible frame rate and accuracy. Improving the ratio of signal to background photons through coincidence detection helps have sufficient samples of the laser pulse. At the same time, higher coincidence levels in the available implementations lead to lower accuracy because later photons are emphasized, and the arrival distribution is widened. For ME detection, influence on accuracy is also existent, but the impact on data emergence is critical. Each level linearly increases the timestamp data generated at each pixel.

Both ME and CD detection, however manipulate the histograms SNR and



(a) SNR for low reflectivity ($\rho = 10\,\%$)

(b) SNR for high reflectivity ($\rho = 80\,\%$)

Figure 6.12: Theoretical SNR (lines) and bin wise estimation (scatter) of 100 histograms at different ambient light levels over distance.

detection probability.

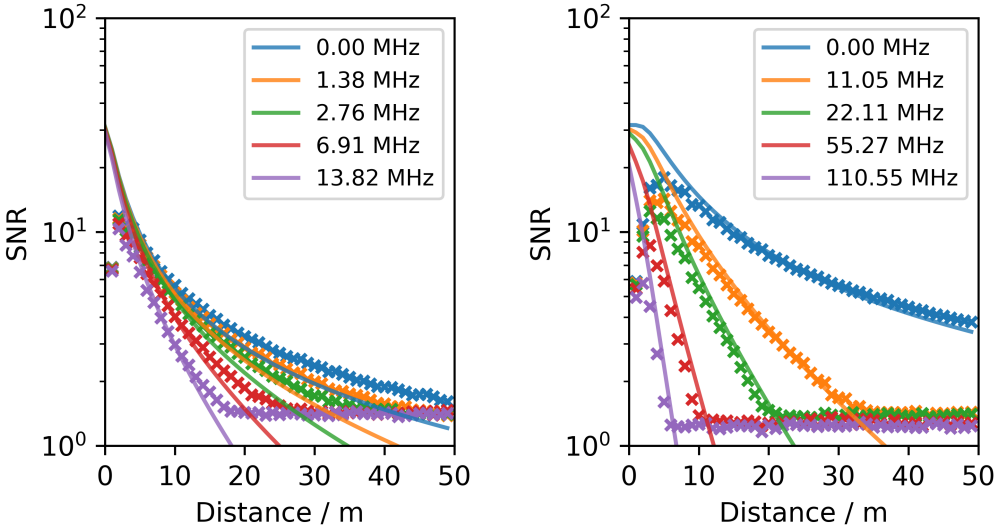
6.5.1 Estimation of SNR

Histograms from the automotive scenario are now used to evaluate two ways to derive SNR from measurement histograms. It will later be used as an input parameter for the hierarchy model. Both of these approaches use the bin wise SNR definition. The first approach is shown in Figure 6.12. This algorithm requires the analysis of multiple histograms and is mainly presented to motivate SNR retrieval from the histogram. To retrieve an estimate for the signals mean count value $\hat{n}_\mathrm{L}$ and standard deviation $\hat{\sigma}_\mathrm{L}$, the highest mean count value and its respective variance of the background compensated histograms are calculated. This value is compared to the standard deviation $\hat{\sigma}_\mathrm{B}$ of count values in a fixed bin before pulse return.

$$\hat{k}_\mathrm{SN} = \frac{\hat{n}_\mathrm{L}}{\sqrt{\hat{\sigma}_\mathrm{L} + \hat{\sigma}_\mathrm{B}}} \tag{6.1}$$

One can see that the retrieved values of $\hat{k}_\mathrm{SN}$ match closely to the theoretical SNR known from the simulation parameters.

Now, a different approach is evaluated aiming to estimate the SNR from a



(a) SNR for low reflectivity ($\rho = 10\,\%$). (b) SNR for high reflectivity ($\rho = 80\,\%$).

Figure 6.13: Theoretical SNR (lines) and mean sliding estimation (scatter) of histograms at different ambient light levels over distance.

single histogram without further knowledge about pulse position as an input parameter for sensor settings. Here, the mean can be estimated from a single bin. However, selecting the highest bin from multiple bins with comparable input signal biases to high values. Also, a single high ambient realization will lead to a wrong SNR estimation. So a small averaging window of two bins is employed for estimating mean signal count value. The ambient influence is estimated by using a sliding variance window with a comparatively broad filter width of 20 bins. The filter width must be broad enough to have a sufficient number of samples to estimate variance but not so broad that the non-constant distribution of histogram noise is ignored. As one can see, there are deviations from the theoretical values in both low and high signal cases, but the approach is well fit to estimate SNR between two and seven.

6.5.2 Combination Algorithm

A control criterion can be defined to choose between multi-event and coincident detection levels. In Figure 6.14, a sequence diagram illustrating a possible combination algorithm is shown. At first, a TCSPC measurement with all available ME levels is conducted, and the different levels are stored in memory. Then, the sys-

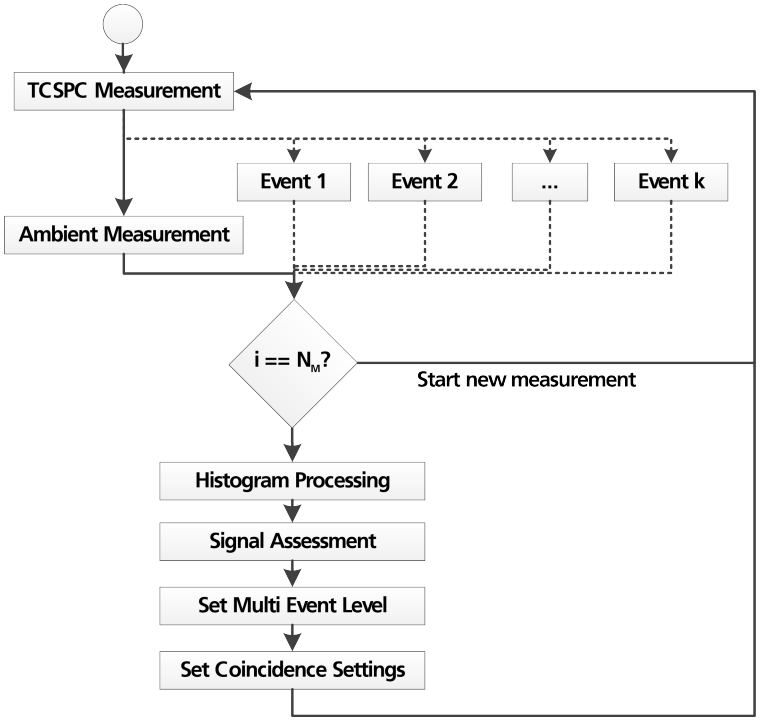


Figure 6.14: Sequence diagram for hierarchically combining ME and coincidence detection.

tem uses an ambient measurement to correct for the background influence in the respective accumulated histogram. This measurement can either be a TCSPC measurement with the laser source turned off or a SPC measurement. When laser PRF is limited to the kHz regime; the time between two pulses is in the magnitude of microseconds. This time can be used to conduct ambient light measurements with SPC or TCSPC without compromising FPS. The distribution of ambient light photons in the histogram is thus either yielded directly or has to be calculated from SPC data, and the background PDF predicted and subtracted ('Histogram Processing'). Next up is the estimation of signal strength ('Signal Assessment'). The straightforward solution is measuring the pulse return and background power on the sensor directly. For example, an intensity measurement without emission of a laser pulse can be conducted in advance of the TCSPC measurement for optimal level choice. Measuring the laser intensity with SPC is not directly possible using SPADs because of their characteristic dead time and short laser pulse times and requires a different measurement scheme with e.g. a photodiode as a second detector or estimation from the TCSPC histogram.

One approach to mention that does not use histogram analysis is to control

the sensor settings depending on the present, scene-dependent ambient light measured in a separate intensity measurement. An example of a coincidence control algorithm depending on intensity measurements is [58]. The advantage is that the sensor settings can be set before the actual measurement is conducted, and thus, no time is lost measuring and analyzing erroneous histograms. However, ambient light levels depend on both object reflectivity and scene parameters like the altitude of the sun, the incidence angle of ambient light, and observance angle between object and sensor. Thereby, little information about signal strength emitted by these objects can be retrieved because it additionally depends on object distance. It is thus unclear whether the ideal setting is no coincidence detection at all because the laser trace is faint in the histogram or a restrictive coincidence setting because there is plenty of power available.

Histogram analysis is the more flexible approach because it does not require any system side modifications. When ambient light is present, it can be compensated for in the histogram either by calculating an ambient light distribution in the histogram from an intensity measurement or by directly recording a correlation histogram without sending out a laser pulse.

The algorithm evaluated here is based on histogram analysis. The signal strength is evaluated from the ME histograms using sliding variance window estimation because of its low complexity. If the necessary processing power is available, a more elaborate algorithm like e.g., an optimization or fitting-based approach can be used to estimate the experimental SNR. Also, instead of SNR, a different measure for detectability like detection probability can be used. In the now demonstrated example implementation, a threshold SNR of 6 is defined over which a signal counts as sufficient.

As soon as the estimated SNR for all event levels is available, sensor settings can be derived. The algorithm chooses the lowest available ME level, which still meets the SNR requirement. If no ME level is sufficient, the set coincidence depth is instead increased by one. Should no setting meet the threshold, the measurement is unsuccessful and should be repeated, or more measurements have to be accumulated to retrieve target distance.

If a working setting for the pixel is found, it is not changed anymore. While this is sufficient for tracking an object through space, dynamic changes in the scene can require the pixel to react again. So when the SNR value goes below the

threshold or when the intensity measurement detects a drastic change in intensity, a new signal assessment is initiated. The disadvantage is that this reduces the available FPS.

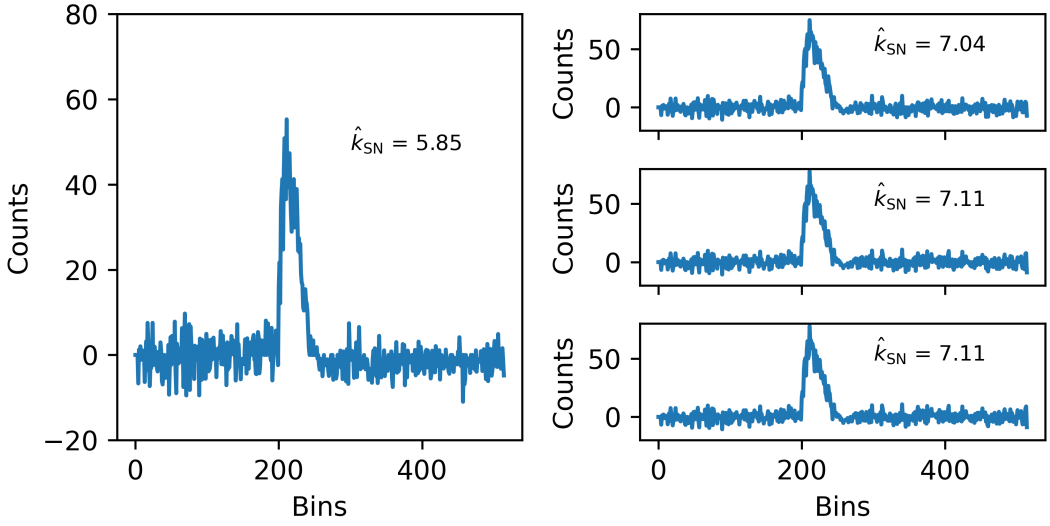
If knowledge about the target objects is available, the reaction time of the approach can further be optimized by discarding settings levels. For example, with the sensor architecture employed here and solely sensing diffuse targets, the user will rarely find any events in histograms with coincidence depth of four. This also depends on the number of implemented SPADs per pixel N_{S} as four events out of twenty is a much more relaxed requirement than four out of four diodes firing. Also, if maximum distance is limited and there is finite conversion time between events, it can make sense to implement only a small number of ME levels.

6.5.3 Example Scenario

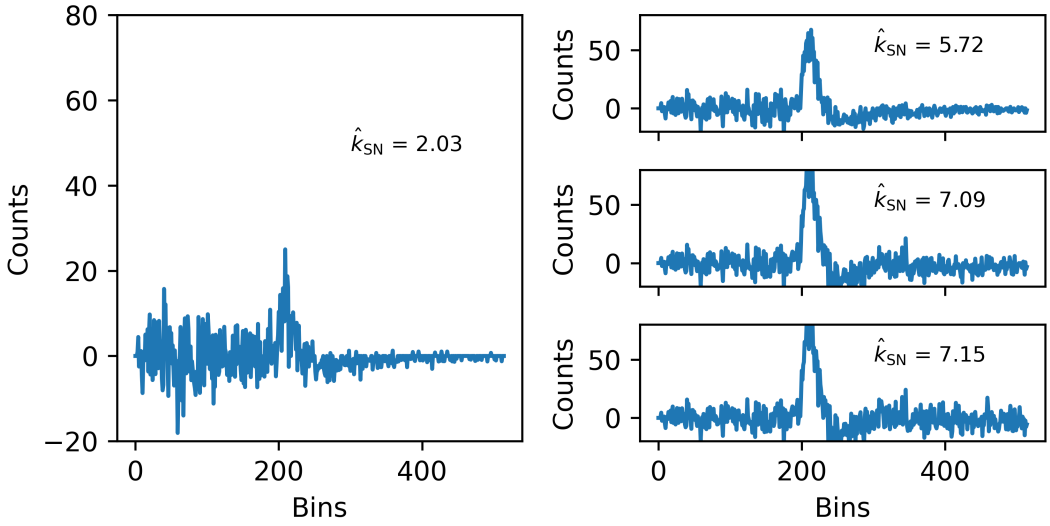
Now, the algorithm's reaction to input signals generated with the measurement setup sketched in Figure 6.1 is demonstrated. For the ambient signals, a single neutral density filter was used and the event rates from Figure 6.11 are reached. All histograms were generated using $N_{\mathrm{M}} = 4000$ accumulations and the SNR threshold is set at $k_{\mathrm{SN}} = 6$, because it is a convenient threshold assumed by the coarse neutral density filter levels and can be estimated well with the sliding window approach previously established. The figures display the first detected photon in the big panel on the left side. All higher photon arrivals are shown in the smaller panels on the right side, ranging from ME level two on the top to four at the bottom.

Results

The histograms in Figure 6.15 a) were measured with a low laser event rate of $r_{\mathrm{L}} = 44.3\,\mathrm{MHz}$ and an ambient event rate of $r_{\mathrm{B}} = 7.0\,\mathrm{MHz}$. The signal-related counts in the first-photon histogram are visible. However, evaluated SNR is below the necessary threshold for this setting. The higher photon arrivals increase derived SNR as expected. However, $k = 2$ with an SNR of $k_{\mathrm{SN}} = 7.04$ is already sufficient, and further ME levels solely increase data without improving signal detectability.



(a) Resulting histograms for $r_L = 44.3\,\text{MHz}$, $r_B = 7.0\,\text{MHz}$ and $n_C = 1$.

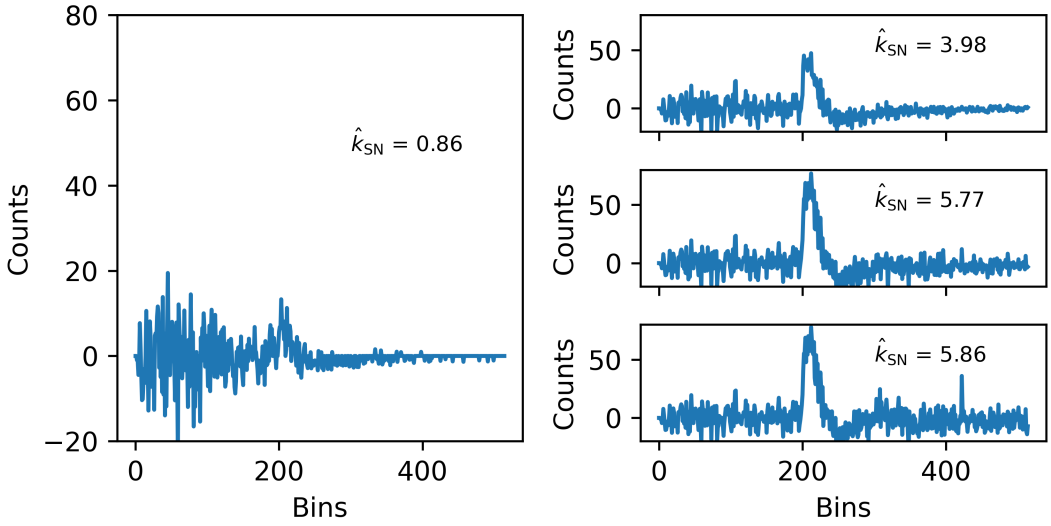


(b) Resulting histograms for $r_L = 85.0\,\text{MHz}$, $r_B = 101.9\,\text{MHz}$ and $n_C = 1$.

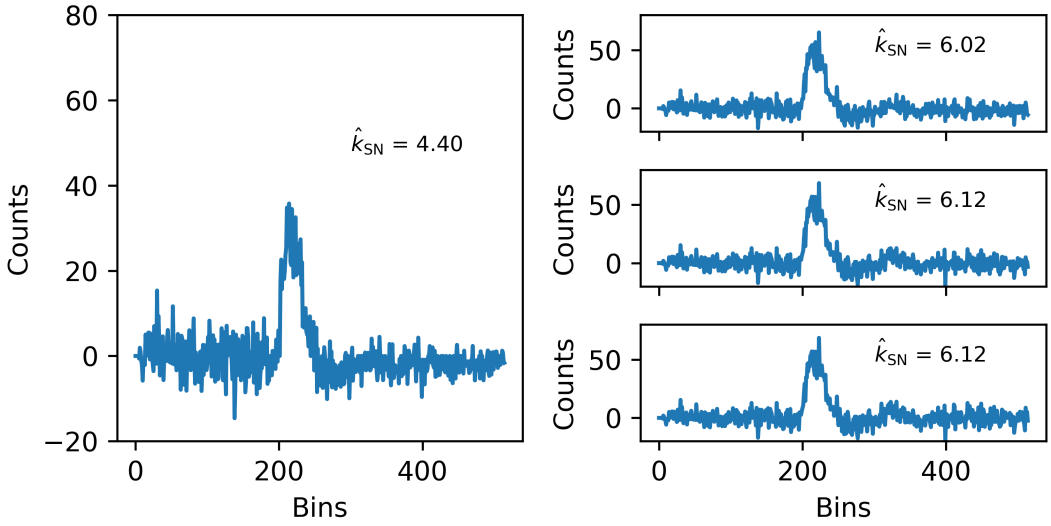
Figure 6.15: Sum histograms of $k = 4$ photon detections in low flux cases. Big panel on the left shows the first event while the panels on the right show two to four from top to bottom.

To emulate the presence of an object with higher reflectivity and higher ambient illumination, both laser and background event rate are increased by two filter levels. This leads to values of $r_L = 85.0\,\text{MHz}$ and $r_B = 101.9\,\text{MHz}$, and results in a higher background event rate than laser event rate. The results are shown in Figure 6.15 b). Because of the increased ambient light level, ME level $k = 2$

does not meet the required SNR anymore. This causes the system to conduct another evaluation measurement. All ME levels are recorded and $k = 3$ identified as sufficient for this scenario.



(a) Resulting histograms for $r_L = 85.0\,\text{MHz}$, $r_B = 360.2\,\text{MHz}$ and $n_C = 1$.



(b) Resulting histograms for $r_L = 85.0\,\text{MHz}$, $r_B = 360.2\,\text{MHz}$ and $n_C = 2$.

Figure 6.16: Sum histograms of $k = 4$ photon detections in high flux cases.

Increasing the ambient event rate to $r_B = 360.2\,\text{MHz}$ via the next available neutral density filter level results in the histograms in Figure 6.16 a). Here, no ME level with sufficient SNR can be found. Subsequently, coincidence depth is increased to two detector $n_C = 2$, and the results are displayed in Figure 6.16 b).

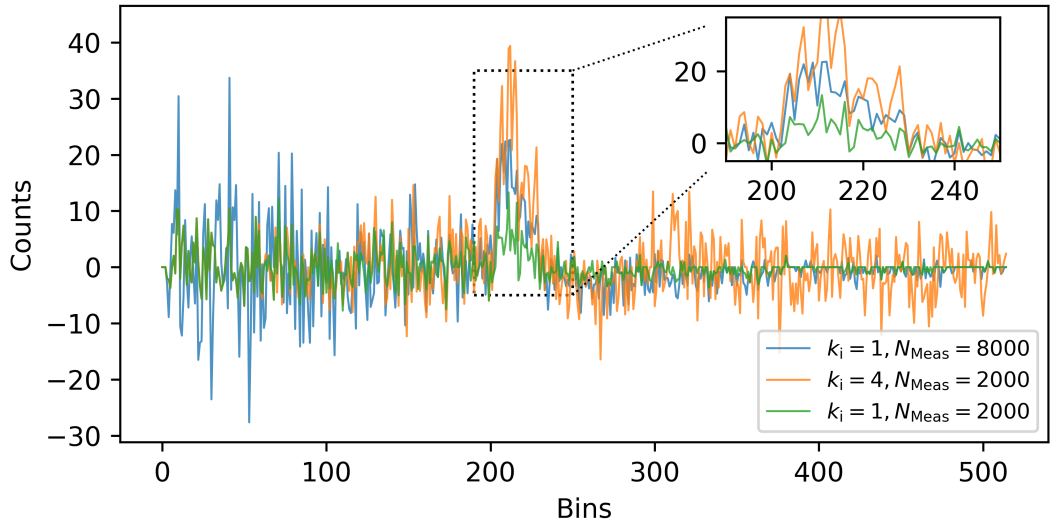


Figure 6.17: Comparison of an increased number of accumulated measurements and ME detection.

The lowest ME level capable of detection is $k = 2$, and the system will remain in it until the next signal assessment phase is conducted.

Accumulation versus Multi-Event Acquisition

As shown, the algorithm improves detectability and provides equal signal levels over different scenarios. In terms of data throughput, a sensor recording N_M timestamps with a ME level of k and a sensor conducting $k \cdot N_M$ measurements in first-photon mode are identical. In low light applications, the first-photon system is advantageous and will provide higher precision than the ME system because of the increased number of samples. In high ambient scenarios however, the exponential behavior in first-photon detection deteriorates the signal so far that further accumulation of measurements reaps a small benefit only. One such scenario is shown in Figure 6.17. As a baseline comparison, first-photon detection with $N_M = 2000$ is shown. The inlet shows that signal level is only marginally higher than the surrounding noise and successful evaluation is challenging. Employing four ME levels, this signal level is drastically increased. The noise level also increases but produces similar count values to the noise in the first-photon histogram. The difference is that in the first-photon histogram, noise is especially high in the first few bins because of the exponential behavior at constant light

levels. In contrast, in the ME histogram, noise is distributed more uniformly over the histogram. Increasing N_{M} to 8000, an increase in signal level by factor four can be seen. At the same time, noise behavior is emphasized and still selectively produces counts at the start of the histogram, impeding successful measurement evaluation. While this increase in accumulation provides a higher signal, the gain through ME detection is higher.

Summary

To summarize, the hierarchy control algorithm enables the system to choose its minimum required settings for detecting in a given scenario. It has an advantage over systems with fixed readout scheme if their readout speed limits them. There, higher FPS is achieved. Similarly, if a histogram is built on-chip, the number of memory accesses can be reduced significantly. If ToF is derived without building a histogram, the number of processed timestamps reduces, but this is only a marginal increase compared to the former ones. Disadvantages lie in the increased complexity of the system. For systems in controlled environments, not suffering from background light influence, like the robotics scenarios from Section 4.3, no advantage can be gained. Also, the signal assessment measurement will most of the time measure in settings that are not necessary for successfully evaluating the measurement. This also reduces FPS slightly depending on the scene and frequency of signal assessment measurements. Systems that are not limited by their readout speed at maximum ME level do not need the control algorithm. There, all ME levels can be evaluated without cost and this systematically yields the highest SNR. Systems containing only a single or few SPAD pixels are an example of this condition being easily met. However, reduction of data throughput in array implementations stays an attractive objective and should be pursued, with the hierarchy control algorithm being one measure to achieve it.

Chapter 7

Conclusions and Outlook

The aim of this work was to assess and extend SPAD DToF system capabilities under varying operation conditions. As performance is mainly limited by signal and ambient light levels, solutions for dealing with these influences were developed.

For this purpose, SPAD smart pixel structures were mathematically described and their performance evaluated. A detector comparison identified favorable operating regimes for the SPAD. Comparing APD and SPAD detectors has shown that the photon-efficient requirements of SPADs are better suited for low signal levels, but cannot tolerate as much ambient light as APDs used as ToF system detectors.

SPAD system behavior at low return power was considered. Two types of correlation-based digital filters were compared to a quantile filter signal chain. The quantile filter signal chain shows a stronger attenuation than the correlation-based filters, limiting its usefulness for achieving long range. Also, while no calculation on the individual histogram count values is required, the sorting necessary for the quantile signal chain is demanding for pixel-individual hardware integration. Further performance increase can be expected when using spatial correlations between neighboring pixels and image processing-based evaluation schemes, but these were not the focus of the investigation.

The high fluctuation in signal return power, present in optical systems built for long range and uncooperative environments, causes distortions in the measured raw data. A hardware-friendly pile-up correction algorithm was investigated, which comprises event rate estimation on the maximum count value of the arrival

time histogram, and a look-up of the event rate-dependent maximum shift. It differs from established algorithms in that it does not manipulate the histogram distribution but calculates a single correction factor for it. The results show an increase in achievable accuracy, and also the limitations of the approach caused by limited rate estimation capabilities. A future development could be combining the identified shift of maximum position with a more accurate rate estimation algorithm. Alternatively, returned optical power could be monitored using a second detector to correct the pile-up error more accurately, but this increases complexity on the system side.

The increase in ambient light tolerance by both coincidence and multi-event detection was made plausible by calculation, simulation and measurement. It was shown that the highest range is achievable by combining the two. For this purpose, the hierarchy model algorithm was developed. It works by analyzing histogram signal levels and can find the sufficient setting with the lowest amount of data output. APDs still perform better under high ambient illuminance than SPADs, but efficient combination of the ambient suppression algorithms can partially close the gap.

The hierarchy model works with discrete coincidence levels. Future implementations could use coincidence circuits with continuously adjustable coincidence time. This lets attenuation be chosen continuously and would require a modified control algorithm.

Both multi-event and coincidence detection have drawbacks in terms of silicon-area requirement, which increases cost and reduces optically active area. As this work does not specify a technology node for manufacturing, the question of how big their drawback is remains hard to answer. It is clear that the influence is less severe in BSI systems and in smaller technology nodes. However, additional process steps and smaller nodes increase the cost of the devices and there could still be a place for cheaper, less performant systems.

While this work makes a contribution to the ambient light rejection capabilities of SPAD-based smart pixels, requirements for autonomous driving are high and will remain a field of research. One frequent topic is the discussion of advantages and disadvantages for flash or scanning systems. With flash and scanning solutions competing in autonomous driving applications, and the current commercial pressure driving development, it is inevitable that both approaches will raise their

profile and that their ideal areas of application emerge. Anyhow, the methods evaluated in this thesis are applicable to both future development directions.

Bibliography

[1] Eric de Borniol, Pierre Castelein, Fabrice Guellec, Johan Rothman, Gautier Vojetta, Gérard Destéfanis, and Michel Vuillermet. A 320x256 hgcdte avalanche photodiode focal plane array for passive and active 2d and 3d imaging. In Bjørn F. Andresen, Gabor F. Fulop, and Paul R. Norton, editors, *Infrared Technology and Applications XXXVII*, SPIE Proceedings, page 801232. SPIE, 2011.

[2] Federica Villa, Fabio Severini, Francesca Madonini, and Franco Zappa. Spads and sipms arrays for long-range high-speed light detection and ranging (lidar). *Sensors*, 21(11):3839, 2021.

[3] Chankyu Kim, Yunho Jung, and Seongjoo Lee. Fmcw lidar system to reduce hardware complexity and post-processing techniques to improve distance resolution. *Sensors*, 20(22):6676, 2020.

[4] Sung-Hyun Lee, Wook-Hyeon Kwon, and Yong-Hwa Park. Amplitude-modulated continuous wave scanning lidar based on parallel phase-demodulation. pages 127–133. SPIE, 2021.

[5] S. Bellisai, D. Bronzi, F. A. Villa, S. Tisa, A. Tosi, and F. Zappa. Single-photon pulsed-light indirect time-of-flight 3d ranging. *Optics express*, 21(4):5086–5098, 2013.

[6] Paul McManamon, editor. *Field guide to lidar*, volume 36 of *SPIE field guides*. SPIE PRESS, Bellingham, Wash., 2015.

[7] Anna Vil, Anna Arbat, Eva Vilella, and Angel Dieguez. Geiger-mode avalanche photodiodes in standard cmos technologies. In Sanka Gateva, editor, *Photodetectors*. InTech, 2012.

[8] A. Rochas, M. Gosch, A. Serov, P. A. Besse, R. S. Popovic, T. Lasser, and R. Rigler. First fully integrated 2-d array of single-photon detectors in standard cmos technology. *IEEE Photonics Technology Letters*, 15(7):963–965, 2003.

[9] Danilo Bronzi, Federica Villa, Simone Bellisai, Bojan Markovic, Simone Tisa, Alberto Tosi, Franco Zappa, Sascha Weyers, Daniel Durini, Werner Brockherde, and Uwe Paschen. Low-noise and large-area cmos spads with timing response free from slow tails. In *2012 Proceedings of the European Solid-State Device Research Conference (ESSDERC)*, pages 230–233. IEEE, 17.09.2012 - 21.09.2012.

[10] J. M. Pavia, M. Scandini, S. Lindner, M. Wolf, and E. Charbon. A 1 × 400 backside-illuminated spad sensor with 49.7 ps resolution, 30 pj/sample tdcs fabricated in 3d cmos technology for near-infrared optical tomography. *undefined*, 2015.

[11] T. Al Abbas, N. A. W. Dutton, O. Almer, S. Pellegrini, Y. Henrion, and R. K. Henderson. Backside illuminated spad image sensor with 7.83 \textmu m pitch in 3d-stacked cmos technology. In *2016 International Electron Devices Meeting*, pages 8.1.1–8.1.4, Piscataway, NJ, 2016. IEEE.

[12] S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa. Avalanche photodiodes and quenching circuits for single-photon detection. *Applied optics*, 35(12):1956–1976, 1996.

[13] Frédéric Nolet, Samuel Parent, Nicolas Roy, Marc-Olivier Mercier, Serge Charlebois, Réjean Fontaine, and Jean-Francois Pratte. Quenching circuit and spad integrated in cmos 65 nm with 7.8 ps fwhm single photon timing resolution. *Instruments*, 2(4):19, 2018.

[14] Massimo Ghioni, Angelo Gulinatti, Ivan Rech, Franco Zappa, and Sergio Cova. Progress in silicon single-photon avalanche diodes. *IEEE Journal of Selected Topics in Quantum Electronics*, 13(4):852–862, 2007.

[15] A. E. Stevens, R. P. van Berg, J. van der Spiegel, and H. H. Williams. A time-to-voltage converter and analog memory for colliding beam detectors. *IEEE Journal of Solid-State Circuits*, 24(6):1748–1752, 1989.

[16] Bojan Markovic, Simone Tisa, Federica A. Villa, Alberto Tosi, and Franco Zappa. A high-linearity, 17 ps precision time-to-digital converter based on a single-stage vernier delay loop fine interpolation. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 60(3):557–569, 2013.

[17] Claudio Bruschini, Harald Homulle, Ivan Michel Antolovic, Samuel Burri, and Edoardo Charbon. Single-photon avalanche diode imagers in biophotonics: review and outlook. *Light: Science & Applications*, 8(1):87, 2019.

[18] Neale A. W. Dutton, Istvan Gyongy, Luca Parmesan, and Robert K. Henderson. Single photon counting performance and noise analysis of cmos spad-based image sensors. *Sensors (Basel, Switzerland)*, 16(7), 2016.

[19] P. Dudek, S. Szczepanski, and J. V. Hatfield. A high-resolution cmos time-to-digital converter utilizing a vernier delay line. *IEEE Journal of Solid-State Circuits*, 35(2):240–247, 2000.

[20] Pekka Keranen and Juha Kostamovaara. 256 x tdc array with cyclic interpolators based on calibration-free 2 x time amplifier. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 66(2):524–533, 2019.

[21] Roland H. Haitz. Mechanisms contributing to the noise pulse rate of avalanche diodes. *Journal of Applied Physics*, 36(10):3123–3131, 1965.

[22] Francesco Ceccarelli, Giulia Acconcia, Angelo Gulinatti, Massimo Ghioni, Ivan Rech, and Roberto Osellame. Recent advances and future perspectives of single–photon avalanche diodes for quantum photonics applications. *Advanced Quantum Technologies*, 4(2):2000102, 2021.

[23] Mohammadreza Dolatpoor Lakeh, Jean-Baptiste Kammerer, Enagnon Aguénounon, Dylan Issartel, Jean-Baptiste Schell, Sven Rink, Andreia Cathelin, Francis Calmon, and Wilfried Uhring. An ultrafast active quenching active reset circuit with 50% spad afterpulsing reduction in a 28 nm fd-soi cmos technology using body biasing technique. *Sensors*, 21(12):4014, 2021.

[24] Abdul Waris Ziarkash, Siddarth Koduru Joshi, Mario Stipčević, and Rupert Ursin. Comparative study of afterpulsing behavior and models in single photon counting avalanche photo diode detectors. *Scientific Reports*, 8(1):5076, 2018.

[25] Ivo Straka, Jan Grygar, Josef Hloušek, and Miroslav Ježek. Counting statistics of actively quenched spads under continuous illumination. *Journal of Lightwave Technology*, 38(17):4765–4771, 2020.

[26] Chen Wang, Jingyuan Wang, Zhiyong Xu, Jianhua Li, Rong Wang, Jiyong Zhao, and Yimei Wei. Afterpulsing effects in spad-based photon-counting communication system. *Optics Communications*, 443:202–210, 2019.

[27] Sahba Jahromi and Juha Kostamovaara. Timing and probability of crosstalk in a dense cmos spad array in pulsed tof applications. *Optics express*, 26(16):20622–20632, 2018.

[28] Igor Kompanets and Nikolay Zalyapin. Methods and devices of speckle-noise suppression (review). *Optics and Photonics Journal*, 10(10):219–250, 2020.

[29] J. W. Goodman. Some fundamental properties of speckle*. *Journal of the Optical Society of America*, 66(11):1145, 1976.

[30] Richard D. Richmond and Stephen C. Cain. *Direct-detection LADAR systems*. Tutorial texts in optical engineering v. TT85. SPIE, Bellingham, Wash. <1000 20th St. Bellingham WA 98225-6705 USA>, 2010.

[31] Markus-Christian Amann, Thierry M. Bosch, Marc Lescure, Risto A. Myllylae, Marc Rioux. Laser ranging: a critical review of usual techniques for distance measurement. *Optical Engineering*, 40(1):10, 2001.

[32] Sara Pellegrini, Gerald S. Buller, Jason M. Smith, Andrew M. Wallace, and Sergio Cova. Laser-based distance measurement using picosecond resolution time-correlated single-photon counting. *Measurement Science and Technology*, 11(6):712–716, 2000.

[33] E.P Baltsavias. Airborne laser scanning: basic relations and formulas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2-3):199–214, 1999.

[34] National Renewable Energy Laboratory. Am 1,5g spectrum.

[35] Din en 60825-1:2015-07: Safety of laser products - part 1: Equipment classification and requirements (iec 60825-1:2014); german version en 60825-1:2014, 2015-07.

[36] Peter R. Michael, Danvers E. Johnston, and Wilfrido Moreno. A conversion guide: solar irradiance and lux illuminance. *Journal of Measurements in Engineering*, 8(4):153–166, 2020.

[37] *Spectral luminous efficiency function for photopic vision*, volume 86 of *Publication CIE*. Commission Internationale de l'Eclairage, Vienna, 1. ed. edition, 1990.

[38] Elham Sarbazi, Majid Safari, and Harald Haas. Statistical modeling of single-photon avalanche diode receivers for optical wireless communications. *IEEE Transactions on Communications*, 66(9):4043–4058, 2018.

[39] Atul Ingle, Trevor Seets, Mauro Buttafava, Shantanu Gupta, Alberto Tosi, Mohit Gupta, and Andreas Velten. Passive inter-photon imaging.

[40] Dr. Rüdiger Paschotta. Sech^2-shaped pulses.

[41] Ruitong Zheng and Guanhao Wu. Constant fraction discriminator in pulsed time-of-flight laser rangefinding. *Frontiers of Optoelectronics*, 5(2):182–186, 2012.

[42] P. B. Coates. The correction for photon 'pile-up' in the measurement of radiative lifetimes. *Journal of Physics E: Scientific Instruments*, 1(8):878–879, 1968.

[43] Adithya K. Pediredla, Aswin C. Sankaranarayanan, Mauro Buttafava, Alberto Tosi, and Ashok Veeraraghavan. *Signal Processing Based Pile-up Compensation for Gated Single-Photon Avalanche Diodes*. 2018.

[44] Anant Gupta, Atul Ingle, Andreas Velten, and Mohit Gupta. Photon-flooded single-photon 3d cameras. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6763–6772, Piscataway, NJ, 2019. IEEE.

[45] Joshua Rapp, Yanting Ma, Robin M. A. Dawson, and Vivek K. Goyal. Dead time compensation for high-flux ranging. *IEEE Transactions on Signal Processing*, 67(13):3471–3486, 2019.

[46] Andre Buchner, Jan F. Haase, Jennifer Ruskowski, and Werner Brockherde. Laser event distribution and timing circuit design constraints in direct tof lidar applications. In Manijeh Razeghi, editor, *Quantum Sensing and Nano Electronics and Photonics XVII*, Proceedings of SPIE. 5200-, page 94, Bellingham, Washington, 2020. SPIE.

[47] Joshua Rapp, Robin M. A. Dawson, and Vivek K. Goyal. Dithered depth imaging. *Optics express*, 28(23):35143–35157, 2020.

[48] Tarek Al Abbas, Neale A.W. Dutton, Oscar Almer, Neil Finlayson, Francescopaolo Mattioli Della Rocca, and Robert Henderson. A cmos spad sensor with a multi-event folded flash time-to-digital converter for ultra-fast optical transient capture. *IEEE Sensors Journal*, 18(8):3163–3173, 2018.

[49] Sheldon M. Ross. *Stochastic processes*. Wiley series in probability and statistics. Wiley, New York, NY, 2. ed. edition, 1996.

[50] Philip Gatt, Steven Johnson, and Terry Nichols. Geiger-mode avalanche photodiode ladar receiver performance characteristics and detection statistics. *Applied optics*, 48(17):3261–3276, 2009.

[51] Juha Kostamovaara, Jussi Tenhunen, Martin Kögler, Ilkka Nissinen, Jan Nissinen, and Pekka Keränen. Fluorescence suppression in raman spectroscopy using a time-gated cmos spad. *Optics express*, 21(25):31632–31645, 2013.

[52] Jan F. Haase, Andre Buchner, Sara Grollius, Jennifer Ruskowski, and Holger Vogt. Measurement concept to reduce environmental impact in direct time-of-flight lidar sensors. In Manijeh Razeghi, editor, *Quantum Sensing and Nano Electronics and Photonics XVII*, Proceedings of SPIE. 5200-, page 9, Bellingham, Washington, 2020. SPIE.

[53] Cristiano Niclass, Mineki Soga, Hiroyuki Matsubara, Satoru Kato, and Manabu Kagami. A 100-m range 10-frame/s 340$\,\times\, $96-pixel time-of-flight depth sensor in 0.18-$\mum$ cmos. *IEEE Journal of Solid-State Circuits*, 48(2):559–572, 2013.

[54] Daniel G. Fouche. Detection and false-alarm probabilities for laser radars that use geiger-mode detectors. *Applied optics*, 42(27):5388–5398, 2003.

[55] Foad Arvani, Tony Chan Carusone, and Edward S. Rogers. Tdc sharing in spad-based direct time-of-flight 3d imaging applications. In *2019 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 1–5, Piscataway, NJ, 2019. IEEE.

[56] Joshua Rapp and Vivek K. Goyal. A few photons among many: Unmixing signal and noise for photon-efficient active imaging. *IEEE Transactions on Computational Imaging*, 3(3):445–459, 2017.

[57] Eberhard Hänsler. *Statistische Signale: Grundlagen und Anwendungen*. Springer Berlin Heidelberg, Berlin, Heidelberg and s.l., 3. auflage edition, 2001.

[58] Maik Beer. *SPAD-basierte Sensoren für die laufzeitbasierte Distanzmessung bei hoher Hintergrundlichtintensität*. 2018.

[59] Andre Buchner, Stefan Hadrath, Roman Burkard, Florian M. Kolb, Jennifer Ruskowski, Manuel Ligges, and Anton Grabmaier. Analytical evaluation of signal-to-noise ratios for avalanche- and single-photon avalanche diodes. *Sensors*, 21(8):2887, 2021.

[60] R. J. McIntyre. Multiplication noise in uniform avalanche diodes. *IEEE Transactions on Electron Devices*, ED-13(1):164–168, 1966.

[61] Robert Hildreth Kingston. *Optical sources, detectors, and systems: Fundamentals and applications*. Optics and photonics. Acad. Press, San Diego, 1995.

[62] Klaus Pasquinelli, Rudi Lussana, Simone Tisa, Federica Villa, and Franco Zappa. Single-photon detectors modeling and selection criteria for high-background lidar. *IEEE Sensors Journal*, 20(13):7021–7032, 2020.

[63] LC-APD. Silicon avalanche photodiode sar-/sarp-series.

[64] Karl-Heinz Waldmann and Werner Helm. *Simulation stochastischer Systeme: Eine anwendungsorientierte Einführung*. Springer Gabler, Berlin, Heidelberg, 2016.

[65] James E. Gentle. *Random number generation and Monte Carlo methods*. Statistics and computing. Springer, New York, 2. edition, corrected second printing edition, 2005.

[66] Justus Krüger. Will industry 4.0 create the dark factory?, 2019.

[67] M. Spies and H. Spies. Automobile lidar sensorik: Stand, trends und zukünftige herausforderungen. *Advances in Radio Science*, 4:99–104, 2006.

[68] Thinal Raj, Fazida Hanim Hashim, Aqilah Baseri Huddin, Mohd Faisal Ibrahim, and Aini Hussain. A survey on lidar scanning mechanisms. *Electronics*, 9(5):741, 2020.

[69] Giuseppe Intermite, Aongus McCarthy, Ryan E. Warburton, Ximing Ren, Federica Villa, Rudi Lussana, Andrew J. Waddie, Mohammad R. Taghizadeh, Alberto Tosi, Franco Zappa, and Gerald S. Buller. Fill-factor improvement of si cmos single-photon avalanche diode detector arrays by integration of diffractive microlens arrays. *Optics express*, 23(26):33777–33791, 2015.

[70] Kazuhiro Morimoto, Andrei Ardelean, Ming-Lo Wu, Arin Can Ulku, Ivan Michel Antolovic, Claudio Bruschini, and Edoardo Charbon. Megapixel time-gated spad image sensor for 2d and 3d imaging applications. *Optica*, 7(4):346, 2020.

[71] C. Niclass, A. Rochas, P.-A. Besse, and E. Charbon. Design and characterization of a cmos 3-d image sensor based on single photon avalanche diodes. *IEEE Journal of Solid-State Circuits*, 40(9):1847–1854, 2005.

[72] Augusto Ronchini Ximenes, Preethi Padmanabhan, Myung-Jae Lee, Yuichiro Yamashita, Dun-Nian Yaung, and Edoardo Charbon. A modular, direct time-of-flight depth sensor in 45/65-nm 3-d-stacked cmos technology. *IEEE Journal of Solid-State Circuits*, 54(11):3203–3214, 2019.

[73] O. Kumagai, J. Ohmachi, M. Matsumura, S. Yagi, K. Tayu, K. Amagawa, T. Matsukawa, O. Ozawa, D. Hirono, Y. Shinozuka, R. Homma, K. Mahara, T. Ohyama, Y. Morita, S. Shimada, T. Ueno, A. Matsumoto, Y. Otake, T. Wakano, and T. Izawa. 7.3 a 189×600 back-illuminated stacked spad

direct time-of-flight depth sensor for automotive lidar systems. In *2016 IEEE International Solid-State Circuits Conference (ISSCC)*, pages 110–112, 2016.

[74] Anwar Shah, Javed Iqbal Bangash, Abdul Waheed Khan, Imran Ahmed, Abdullah Khan, Asfandyar Khan, and Arshad Khan. Comparative analysis of median filter and its variants for removal of impulse noise from gray scale images. *Journal of King Saud University - Computer and Information Sciences*, 2020.

[75] R. M. Hodgson, D. G. Bailey, M. J. Naylor, A. L.M. Ng, and S. J. McNeill. Properties, implementations and applications of rank filters. *Image and Vision Computing*, 3(1):3–14, 1985.

[76] OSRAM. Pl 90 datasheet.

[77] Picolas ldp av 40-70 datasheet rev. 1905.

[78] Laser Components. Ql90o7sa datasheet.

[79] Andre Buchner, Bedrich Hosticka, Olaf Schrey, Jan F. Haase, Jennifer Ruskowski, and Anton Grabmaier. Acquisition of multiple events in direct time-of-flight lidar using single-photon avalanche diodes. In *2020 IEEE SENSORS*, pages 1–4. IEEE, 10/25/2020 - 10/28/2020.

[80] F. Villa, B. Markovic, S. Bellisai, D. Bronzi, A. Tosi, F. Zappa, S. Tisa, D. Durini, S. Weyers, U. Paschen, and W. Brockherde. Spad smart pixel for time-of-flight and time-correlated single-photon counting measurements. *IEEE Photonics Journal*, 4(3):795–804, 2012.

[81] Datasheet pulsed laser diode module ls9-220-8-s10-00, 02/2020.

[82] Femto. Datasheet hsa-x-s-1g4-si - ultra high-speed photoreceiver with si-pin photodiode.

DOI: 10.17185/duepublico/78563
URN: urn:nbn:de:hbz:465-20230705-153731-7