

Diploma Thesis

Design of a High Voltage Active Quenching Circuit in $0.15\,\mu m~CMOS$

submitted in satisfaction of the requirements for the degree of Diplom-Ingenieur at the TU Wien, Faculty of Electrical Engineering and Information Technology

Diplomarbeit

Entwicklung einer aktiven Hochvolt-Quencher-Schaltung in 0,15 µm CMOS

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines

Diplom-Ingenieurs

eingereicht an der Technischen Universität Wien, Fakultät für Elektrotechnik und Informationstechnik

von

Martin Jungwirth, BSc

Matr.Nr.: 01126062

unter der Anleitung von

Univ.Prof. Mag.rer.nat. Dr.techn. Horst Zimmermann

Dipl.Ing. Dr.techn. Michael Hofbauer, BSc

Institute of Electrodynamics, Microwave and Circuit Engineering Technische Universität Wien Karlsplatz 13, 1040 Wien, Österreich

Wien, am 13. Januar 2020



Abstract

This thesis is about the design of a high voltage active quenching circuit (AQC) for external single-photon avalanche diodes (SPADs) in a 0.15 µm complementary metaloxide-semiconductor (CMOS)-process. The designed AQC can work with an excess bias voltage up to 35 V and has a quenching time of 2.2 ns. With this circuit it is possible to investigate the performance improvement of SPAD-detectors due to increased excess bias voltage, such as the increased photon-detection probability (PDP). Additionally, the short quenching time should reduce the afterpulsing probability (APP) which is especially important when using high excess bias voltages.

In the introduction the operating principles of SPADs are reviewed along with the required so called quenching circuits for their proper operation. This quenching circuits come in two different kinds which are passive quenching circuits (PQCs) and active quenching circuits (AQCs), with their advantages and disadvantages. The main part of this thesis is dedicated to the design, development and selection of the different circuit blocks needed for this AQC. Special attention was given to the avalanche detection circuit and the level shifter due to their impact on the quenching time, followed by a description of the chosen and implemented design and considerations regarding the physical layout of the designed circuit. Lastly the performance of the implemented design is evaluated via post-layout simulation and compared to State of the Art high voltage AQCs.



Kurzfassung

Diese Diplomarbeit beschäftigt sich mit der Entwicklung einer aktiven Quencher-Schaltung (AQC) für hohe Excess-Spannungen zum Betrieb von Einzelphoton-Lawinen-Dioden (SPAD). Die Schaltung wird in einer 0.15 µm komplementären Metall-Oxid-Halbleiter (CMOS) Technologie entwickelt. Die entwickelte Schaltung kann mit einer Excess-Spannung von bis zu 35 V arbeiten und benötigt dabei 2.2 ns zum Löschen eines Lawinendurchbruchs. Dadurch ist es möglich, die Verbesserung der spannungsabhängigen Eigenschaften von Einzelphotonendetektoren zu untersuchen, im speziellen die Steigerung der Photonen Detektions-Wahrscheinlichkeit (PDP). Zusätzlich reduziert die kurze Zeitdauer, die zum Löschen eines Lawinendurchbruchs benötigt wird, die Wahrscheinlichkeit von sekundären Lawinendurchbrüchen. Das ist besonders wichtig bei hohen Excess-Spannungen.

Im ersten Teil der Arbeit werden die Grundlagen von SPADs und die zu deren Betrieb notwendigen sogenannten Quencher-Schaltungen besprochen. Diese Quencher-Schaltungen gibt es in zwei Ausführungsformen, einerseits als passive (PQC) und als aktive (AQC) Quencher-Schaltungen mit ihren jeweiligen Vor- und Nachteilen. Der Hauptteil dieser Arbeit beschäftigt sich mit dem Design, der Entwicklung und Auswahl der notwendigen Schaltungsblöcke für die zu entwickelnde aktive Quencher-Schaltung. Darauf hin folgt eine Beschreibung des ausgewählten und implementierten Designs sowie Überlegungen zum zugehörigen physikalischen Layout. Schließlich wird die Leistungsfähigkeit der implementierten Quencher-Schaltung an Hand von Post-Layout Simulationen untersucht und mit dem Stand der Technik verglichen.



Acknowledgment

I would first like to express my thanks to my supervisor Univ.Prof. Dr.techn. Horst Zimmermann for giving me the opportunity to write this master thesis in his research group at the Institute of Electrodynamics, Microwave and Circuit Engineering. I would also like to thank my colleagues from the Circuit Design group who helped me at the work on this thesis. Especially Univ.Ass. Dipl.-Ing. Dr.techn. Michael Hofbauer who had always an open ear for my inquires and who proofread this master thesis. Furthermore, I would like tho thank MSc. Alija Dervić for the fruitful discussions regarding the design of the circuit and Dipl.-Ing. Michael Hauser for his advice and help.

Finally, I must express my very profound gratitude to my brother for accompanying me throughout my study and especially to my parents who have given me the opportunity to study, and their support and encouragement throughout my study. This accomplishment would not have been possible without them. Thank you.



Contents

Glossary		iii	
Ac	rony	ms	v
Lis	List of Symbols vi		
Lis	st of	Figures	ix
1	Intro	oduction	1
	1.1	Problem - Motivation	1
	1.2	SPAD - Single-Photon Avalanche Diodes	1
	1.3	Quenching Circuits	5
		1.3.1 Passive Quenching Circuits (PQC)	6
		1.3.2 Active Quenching Circuits (AQC)	8
	1.4	State of the Art for High-Voltage AQCs	10
		1.4.1 High-Voltage AQC for Count Rates up to 80 M counts/s $\ .$	10
		1.4.2 High-Voltage Active Quenching and Resetting Circuit in $0.35\mu\mathrm{m}$	
		CMOS	11
		1.4.3 Monolithic AQC for Single-Photon Avalanche Detectors	11
	1.5	High-Voltage CMOS Process	13
2	Desi	ign Approaches	15
	2.1	Basic Structure of the High Voltage AQC	15
	2.2	SPAD Simulation Model	17
	2.3	Avalanche Detection	20
		2.3.1 Voltage-Mode Sensing	20
		2.3.2 Combined Sense Stage and Reset Switch	21
	2.4	Dimensioning of Quenching and Resetting Transistor	24
	2.5	Comparator	27
	2.6	Level Shifter	28
	2.7	Control Logic	32
	2.8	Output Driver	36

	2.9	Thermal Considerations	37
3	Imp	emented AQC Design	41
	3.1	Implemented Design	41
	3.2	Layout	44
		3.2.1 AQC-Core Layout	44
		3.2.2 Custom Pads and Padring	45
4	Res	Ilts and Comparison	49
	4.1	Post-Layout Simulation Results	49
		4.1.1 Exemplary Results for a Quenching Voltage of $35 V \dots \dots$	49
		4.1.2 Capability of Using different SPAD-Types	53
	4.2	Comparison to State of the Art	57
5	Con	clusion and Outlook	61
Bi	bliog	raphy	63

Glossary

Cadence	Cadence Design Systems, Inc. an American electronic
	design automation (EDA) software company.
LFoundry	LFoundry GmbH, a SMIC owned semiconductor manufac-
	turer.
DNWELL	deep n-well implantation used for isolation of circuits.
hold-off time	time the bias voltage is kept below the breakdown voltage
	after an avalanche.
LF15A	modular 0.15 $\mu\mathrm{m}$ RF CMOS process offered by $LFoundry.$
NISO	n-layer implant for isolating single devices.
NWELL	n-type well implantation.
PWELL	p-type well implantation.
ready state bias voltage	bias voltage of the SPAD when biased above breakdown
	and ready to detect photons.
reset time	the time it takes to recharge the SPAD to its ready state
	bias voltage.



Acronyms

APD APP	avalanche photo diode. afterpulsing probability.
AQC	active quenching circuit.
ASIC	application-specific integrated circuit.
CMOS	complementary metal-oxide-semiconductor.
\mathbf{CMS}	current-mode sensing.
DC	direct current.
DCR	dark count rate.
EDA	electronic design automation.
ESD	electro-static discharge.
HV	high voltage.
HVCS	high-voltage cascode switch.
HVMOS	high-voltage MOS.
LDMOS	lateral double-diffused MOS.
MIM	metal insulator metal.
MOS	metal-oxide semiconductor.
MOSFET	metal-oxide semiconductor field effect transistor.
PCB	printed circuit board.
PDK	Process Design Kit / Process Development Kit.
PDP	photon-detection probability.
PQC	passive quenching circuit.
PVT	process, voltage and temperature.
RE-SPAD	Red Enhanced SPAD [1].
SPAD	single-photon avalanche diode.
VMS	voltage-mode sensing.



List of Symbols

C_{SPAD}	SPAD junction and parasitic capacitance.
$M_{protect}$	p-channel LDMOS transistor which supports the quenching voltage.
M_Q	n-channel LDMOS transistor for quenching the avalanche.
M_{sense}	p-channel low voltage MOSFET for sensing the avalanche current.
P_T	avalanche triggering probability.
η	quantum efficiency.
Q_{EX}	theoretical avalanche charge, excess charge.
R_{SPAD}	intrinsic resistance of a SPAD.
t_{BD}	time to breakdown.
T_{reset}	resetting time.
V_{Bias}	SPAD reverse bias voltage.
V_{Break}	breakdown voltage of SPAD.
V_{EX}	excess bias voltage.
V_q	quenching voltage.



List of Figures

1.1	PDP for different excess bias voltages	3
1.2	Cross section of a thin SPAD-structure	4
1.3	Cross-section of a thick SPAD	5
1.4	Basic schematic diagram of a PQC	7
1.5	Cathode voltage and current waveform for a PQC	8
1.6	Basic schematic diagram of an AQC	8
1.7	Cathode voltage and current waveform for an AQC	9
1.8	Structure of Acconcia's AQC design.	11
1.9	Schematic of the high-voltage AQC with the HVCS	12
1.10	Schematic for the sense stage	12
1.11	Principle structure of the two different high voltage MOSFET-devices $% \mathcal{A}^{(n)}$.	13
1.12	MOSFET-symbols for the different voltage levels or gate oxides	14
91	Basic structure of the AOC design	16
$\frac{2.1}{2.2}$	Simple SPAD model	17
2.2	SPAD simulation model	18
2.0	Single photon response at the output of a POC for two discrete SPADs	10
2.5	Basic VMS stage	21
$\frac{2.0}{2.6}$	Test circuit for VMS sense stage	21
2.0 2.7	Current into SPAD-cathode during onset of an avalanche and its branch	
2.1	currents	23
2.8	Schematic of the principle design for the combined sense and reset stage.	<u>-</u> 3
2.9	Current into SPAD-cathode during onset of an avalanche.	25
2.10	Operational amplifier based free running comparator.	27
2.11	Declerg based Level shifter design.	29
2.12	Histogram for changes in the peak output voltage of the level shifter due	-
	to PVT-variations, evaluated with Monte Carlo simulation.	30
2.13	Moghe based Level shifter design.	31
2.14	Output signals for the Moghe based level shifter with the stair step like	
	edge clearly visible.	33

2.15	State diagram for the control logic	34
2.16	Schematic of a inverter with adjustable fall time	35
2.17	Schematic of the actual implemented control logic	35
2.18	One quenching sequence simulated with equal hold-off and reset time	36
2.19	Schematic for the 50 Ω output driver	37
3.1	Schematic of the implemented high-voltage AQC.	42
3.2	Layout image of the implemented high-voltage AQC-core	45
3.3	Layout of the overall high-voltage AQC-chip	47
3.4	Schematic of the testbench circuit for the high-voltage AQC to evaluate	
	the design	48
4.1	SPAD-cathode voltage waveform and corresponding <i>Out</i> -signal (top),	
	and SPAD-current into cathode (bottom).	50
4.2	Power consumption of the AQC-chip during one quenching and reset phase.	51
4.3	SPAD-cathode voltage at different process corners	52
4.4	SPAD-cathode voltage at different hold-off times by adjusting the resistor	
	$R_{holdoff}$ accordingly.	53
4.5	V_{sense} -node voltage during an avalanche and subsequent quenching phase.	54
4.6	Safe area for the usage of different SPADs	55
4.7	Quenching pulses for different quenching voltages.	57
4.8	SPAD-cathode voltage for evaluation of the maximum count rate	58
4.9	Waveform of the anode voltage at two different quenching voltages	59

Chapter 1

Introduction

1.1 Problem - Motivation

The detection of very faint optical signals down to the detection of single photons is nowadays a very important capability in many scientific and industrial fields as well as in telecommunication. Applications which are based on this capability are for example in biochemistry and biomedicine Single Molecule Fluorescence Spectroscopy (SMFS), Förster Resonance Energy Transfer (FRET), in industry and automotive Laser Imaging Detection and Ranging (LIDAR) and in telecommunications Quantum Key Distribution (QKD) [2], [3]. Furthermore, single-photon detectors are the major prerequisite for optical quantum information (QI) applications like the ambitious linear optical quantum computing (LOQC) [4].

Single-photon avalanche diodes are now the most used single-photon detectors which replaced the earlier used photo multiplier tubes (PMT) due to the advantages of solid state detectors [2]. The advantages over PMTs are a smaller size, lower supply voltage, ruggedness, high reliability and insensitiveness to magnetic fields [2], [3]. In recent studies it was shown that the already high detector sensitivity can be further increased by applying a higher excess voltage to the detector [5]–[7]. To exploit this effect Acconcia, Rech, Gulinatti, *et al.* [3] published a high voltage active quenching circuit in 2016 with the ability to apply excess voltages up to 50 V, only limited by the limitations of the used semiconductor technology. The goal of this master thesis is to modify and improve this active quenching circuit to reduce the quenching time and improve the overall circuit performance.

1.2 SPAD - Single-Photon Avalanche Diodes

Single-photon avalanche diodes are very sensitive semiconductor photodetectors with the ability to detect single photons. These photodetectors can either be fabricated with standard complementary metal-oxide-semiconductor (CMOS) technologies or custom technologies like GaAs. A SPAD as well as an avalanche photo diode (APD) is a special kind of photo diode which exploit the inner photoelectric effect for detection of photons. The inner photoelectric effect describes the photon triggered excitation of an electron from the valence band into the conduction band [8]. The excitation of an electron into the conduction band can only happen when the photon energy is equal or greater than the band gap energy of the absorbing material. This electron in the conduction band is only loosely bound to its nucleus and can therefore take part in electron current, also called photo current. A SPAD is operated in a so called "Geiger" mode, in this mode of operation the diode is reverse biased with a voltage V_{Bias} way above its breakdown voltage V_{Break} . This voltage V_{Bias} will be called from now on ready state bias voltage. The difference between the bias voltage V_{Bias} and the V_{Break} is called excess bias voltage V_{EX} . The high bias voltage causes a high electric field in the space-charge region of the SPAD. If a photon is absorbed in the space-charge region of such a SPAD and excites an electron into the conduction band, thereby generating a electron-hole pair, this electron becomes accelerated by the electric field. Owing to the high electric field strength this electron gets accelerated to such high energies that it can ionize silicon atoms by impact ionization and generate secondary electron-hole pairs. The secondary electrons are again accelerated in the high electric field and generate additional electron-hole pairs which in turn generate additional electron-hole pairs leading to a self-sustaining avalanche current [9]. Due to this so called avalanche multiplication process it is possible to detect single photons with SPADs. Once a avalanche is triggered by a photon the SPAD as photon detector is essentially blind for further photons. Because of the self-sustaining avalanche a further photon generated electron-hole pair can not be distinguished. For this reason a SPAD needs to be operated with a suitable circuit to quench the avalanche current and restore the bias condition and generate a voltage pulse to signalize a photon absorption to a subsequent processing unit. Such circuits are called quenching circuits because they interrupt the avalanche current by lowering the bias voltage V_{Bias} below the breakdown voltage V_{Break} like a fuse quenches an arc. The details of such quenching circuits will be discussed in Section 1.3.

The photon-detection probability (PDP) is an important characteristic of SPADs and describes how well and efficient photons are detected. The PDP is defined as ratio between the incoming photon count and the triggered avalanche current pulses [5]. The PDP depends on many factors, most of them are intrinsic parameters of the SPAD like the reflection coefficient R, absorption coefficient α for silicon, ionization coefficients α_n, α_p for electrons and holes and the geometry of the depletion region. Only the dependence of the PDP on the excess bias voltage V_{EX} as stated in [5]–[7] is of interest for the design of quenching circuits, since this gives the opportunity to tune the detector performance extrinsically by the quenching circuit. The PDP can be described as the product of quantum efficiency η and avalanche triggering probability P_T as shown in Equations 1.1 [5]–[7].

$$PDP(\lambda, V_{EX}) = \eta(\lambda) \cdot P_T(V_{EX}) \qquad P_T = 1 - e^{-(V_{EX}/V_C)}$$
(1.1)

The right formula in Equations 1.1 is an approximation for the excess voltage V_{EX} dependent triggering probability P_T where the characteristic voltage V_C incorporates different SPAD-parameter. This dependence of the PDP from the excess bias voltage V_{EX} is depicted in Fig. 1.1 for two different structures of SPADs. Beside this ideal



Fig. 1.1: PDP for different excess bias voltages V_{EX} for thin SPADs (left) and for thick SPADs (right) [5, Fig. 3].

behavior of a SPAD as a photon detector there are some non ideal and parasitic effects worth mentioning. One effect which can limit the photon detection ability towards low photon counts are thermally generated carriers. These thermally generated electronhole pairs can trigger an avalanche even if the SPAD is kept in dark conditions [5], [9]. Therefore, the number of thermally generated avalanche pulses per time interval is called dark count rate (DCR). This DCR is strongly temperature dependent due to its underlying physical process. Furthermore, high electric fields in the multiplication region can increase the DCR through a process called field-enhanced thermal generation [1]. Another contribution to the avalanche pulse counts is caused be defects and impurities in the depletion layer with deep energy levels between mid-gap and the band edge. These energy levels are called traps because they can capture electrons during an avalanche [5]. These captured electrons are released with considerable delay after their initiating avalanche and can trigger another avalanche current pulse. Due to their occurrence shortly after the initiating avalanche, this effect is called afterpulsing.

The physical structure of SPADs can be classified in two groups, which are thin and thick SPADs. When realized as planar diodes with a depletion region of about $1 \,\mu\text{m}-2 \,\mu\text{m}$ thickness, they are called thin SPADs. A cross-section for such a SPAD-structure can be seen in Fig. 1.2. This structure has a low breakdown voltage ($15 \,\text{V}-40 \,\text{V}$), a



Fig. 1.2: Cross section of a thin SPAD-structure in this case a double-epitaxial SPAD[1, Fig. 2].

peak PDP up to 40% and excellent time resolution [10]. Usually the anode contact, the red semiconductor layer in Fig. 1.2, has a much larger structure than its opposite contact and therefore a higher parasitic capacitance against the substrate. Thus the cathode contact is best suited to be connected to the quenching circuit. Thin SPADs are compatible with planar CMOS-circuits and can therefore be integrated with their quenching circuit on the same die. Due to the isolation sinker thin SPADs are suitable for detector arrays on a chip.

Thick SPADs have a depletion region thickness of about $30 \,\mu\text{m}-40 \,\mu\text{m}$, a high breakdown voltage in the range of $250 \,\text{V}-470 \,\text{V}$ and a peak PDP above 50% [5], [10] as can be seen in the diagram on the right-hand side of Fig. 1.1. These diodes use the whole thickness of a back etched wafer as absorption zone shown in Fig. 1.3, therefore this structure is less suited for detector arrays or direct integration with the quenching circuit. Such thick SPADs are often built in a special semiconductor process other than CMOS which also prohibits the direct integration with the quenching circuit. At first glance the anode and cathode contact have nearly the same size which results in similar parasitic against the surrounding. Therefore, there is no preference for which contact is connected to the AQC. Although, in case of the thick SPAD in Fig. 1.3, due to mounting the anode contact is preferred to be connected to the AQC.

There are also SPADs which do not clearly fit into one of the above mentioned two groups, like the SPAD used in the work by Zimmermann et al. [11] which is kind of



Fig. 1.3: Cross-section of a thick SPAD[5, Fig. 8].

a intermediate type. Their layer structure is more like a thick SPAD by utilizing the p-epitaxy layer as absorption zone with a thickness of about 12 µm which is lower than for usual thick SPADs. On the other hand these SPADs are fabricated as planar devices similar to thin SPADs, which allows them to be integrated with the detection circuit in one chip. Due to their thick absorption zone they are best suited for the red and near infrared spectral range.

1.3 Quenching Circuits

As introduced in Section 1.2, the task of a quenching circuit is to stop the self-sustaining avalanche current and to restore proper operating conditions. Therefore, the quenching circuit has to fulfill following functional tasks [5]:

- 1. detect the avalanche current
- 2. interrupt the avalanche current by lowering the bias voltage V_{Bias}
- 3. generate an output pulse synchronous to the detection
- 4. restore operating conditions by charging the SPAD to its ready state bias voltage

Apart from performing these tasks in the given sequence to operate the SPAD as photon detector, the quenching circuit can reduce the influence of non idealistic effects. As mentioned in Section 1.2 the afterpulsing probability (APP) is strongly dependent on the amount of avalanche charge flowing throw the depletion region. There are some methods to reduce the APP, which mostly depend on the quenching circuit. Obviously the avalanche charge depends on the size of the SPAD-capacitance and the parasitic capacitance [2]. The SPAD-capacitance is usually given by the used SPAD and therefore cannot be much reduced. The parasitic capacitance can partially be reduced by a thorough physical layout of the circuit. But usually the contribution from integrated quenching circuits to the overall capacitance is small and therefore have only minor influence on the APP. A further method is to detect and quench the avalanche as fast as possible so only a small part of the charge stored in the SPAD-capacitance can flow through the depletion region [2]. Therefore, the performance of the circuit, especially the speed in which the tasks one and two can be executed, influences the APP. Another way to reduce afterpulsing is by introducing a hold-off time after quenching the avalanche. During this hold-off time the bias voltage of the detector is kept below the breakdown voltage V_{Break} to give the trapped carriers time to be emitted, and to recombine without triggering an additional avalanche [2]. After the hold-off time the bias voltage is reset to its ready state bias voltage. The following list summarizes the important time parameters for the design and characterization of quenching circuits.

quenching time Elapsed time from onset until the avalanche is quenched.

hold-off time Time the bias voltage is kept below the breakdown voltage after an avalanche.

reset time The time it takes to recharge the SPAD to its ready state bias voltage.

dead time Sum of all three time parameters above and denote the time the detector is blind for photons.

The different approaches to quench an avalanche can be roughly grouped into two kinds of quenching circuits, namely active quenching circuits (AQCs) and passive quenching circuits (PQCs).

1.3.1 Passive Quenching Circuits (PQC)

Passive quenching circuits are the simplest way to accomplish the tasks associated with quenching an avalanche. Passive quenching is basically done through a high value ballast resistor as shown in Fig. 1.4. When waiting for a photon to trigger an avalanche the SPAD is biased at its ready state bias voltage through the ballast resistor R_B . At the onset of an avalanche the current swiftly rises to its peak value defined by the excess bias



Fig. 1.4: Basic schematic diagram of a passive quenching circuit (PQC) [2].

voltage V_{EX} and the intrinsic resistance R_{SPAD} of the SPAD. This intrinsic resistance R_{SPAD} discharges the parasitic capacitance C_{SPAD} until the avalanche current reaches its final value which is around $\frac{V_{EX}}{R_B}$. If this final avalanche current is below the so called "quenching threshold" of about 100.0 µA the avalanche quenches itself, according to [2]. Due to the statistical nature of this avalanche multiplication process the time duration after which self-quenching happens will randomly vary around a mean value. As a consequence of self-quenching nearly the whole excess charge $Q_{EX} = V_{EX} \cdot C_{SPAD}$ will flow through the SPAD's space charge region and can cause afterpulsing. After the avalanche current quenches itself, the parasitic capacitance C_{SPAD} is recharged through the ballast resistor R_B to its ready state bias voltage and the SPAD is ready to detect another photon. Due to the fact that R_B must be in the range of a few $100 \,\mathrm{k\Omega}$ to successfully quench an avalanche, the recharge of C_{SPAD} is rather slow and therefore the reset time is long. Exemplary cathode voltage and current waveforms for a PQC can be seen in Fig. 1.5. During the long reset transition the bias voltage is above the breakdown voltage all the time as can be seen in the top graph in Fig. 1.5. Therefore, an avalanche can be triggered during the reset phase. When this happens the bias conditions are not as desired, causing a lower current amplitude, which is harder to detect. Due to this non ideal bias conditions the performance of the diode is degraded in terms of detection efficiency and timing response [2]. This long reset time is one of the major drawbacks of the passive quenching method apart from a not well defined dead time. A further drawback are the limited possibilities to reduce the APP. On the other hand a PQC is simple and has a lower area occupation if integrated compared to AQC, which is advantageous for SPAD-arrays.



Fig. 1.5: Cathode voltage and current waveform for a passive quenching circuit [2].

1.3.2 Active Quenching Circuits (AQC)

Active quenching circuits (AQCs) are the attempt to overcome the drawbacks of PQCs, which is pursued with different types of AQCs. A common approach for an AQC design is shown in Fig: 1.6. To be precise Gallivanoni et al. [2] characterizes this principle diagram in Fig. 1.6 as mixed active-passive quenching circuit. However, in the literature it is a widely used structure for a quenching circuit and commonly referred to as AQC. A solely AQC is rarely used due to the drawback of non instant quenching at the onset of an avalanche. Whereas the active quenching structure in Fig. 1.6 uses one of the



Fig. 1.6: Basic schematic diagram of a mixed active-passive quenching circuit usually named AQC [2].

great advantages of passive quenching, which is instant quenching at the onset of an

avalanche. Since the passive quenching starts immediately and has no reaction time and delay. Passive quenching is accomplished through the ballast resistor R_B , which also acts as sense resistor for the active part to detect an avalanche. The passive quenching phase lasts from triggering of an avalanche until the switch S_Q closes and the active quenching phase begins. Usually the switch S_Q keeps the bias voltage of the SPAD below the breakdown voltage V_{Break} for the hold-off time, to allow trapped carriers to recombine [2].

After the hold-off time is over the switch S_Q opens and the switch S_R is closed to recharge the parasitic capacitance C_{SPAD} to the ready state bias voltage. The reset switch S_R is only closed for a short reset duration which should only be as long as needed to recharge C_{SPAD} to the required bias voltage. A short reset duration reduces the probability that avalanches are triggered during the reset phase. When this happens the avalanche cannot be detected immediately due to the low impedance path through S_R to the supply voltage. An avalanche during the reset phase can only be detected after the reset phase is over, therefore the avalanche charge is much larger and increases the APP. Exemplary cathode voltage and current curves for an AQC are shown in Fig. 1.7. The exponential decay of the cathode voltage beginning at 0.0 ns is due to



Fig. 1.7: Cathode voltage and current waveform for an AQC [2].

the passive quenching until the active part intervenes and swiftly lowers the voltage below breakdown, thereby quenching the avalanche. Contrary to a passive quenching circuit (PQC), the hold-off time is well-defined and adjustable. A mixed passive-active quenching circuit combines the main advantages of PQCs and AQCs. These are instant quenching from the PQC, fast quenching and reset and well defined hold-off time from the AQC.

1.4 State of the Art for High-Voltage AQCs

1.4.1 High-Voltage AQC for Count Rates up to 80 Mcounts/s

The High-voltage integrated active quenching circuit for single photon count rate up to 80 Mcounts/s [3] proposed by Acconcia et al. is the first integrated AQC able to apply quenching voltages up to 50 V, according to the authors. This high voltage AQC is designed for external SPADs, especially the Red Enhanced SPAD (RE-SPAD) from Gulinatti et al. [1]. The circuit is fabricated in a high-voltage 0.18 µm CMOS technology. The structure of Acconcia's AQC is shown in Fig. 1.8 and is similar to the basic structure of a mixed active-passive quenching circuit shown in Fig. 1.6. The main difference is how the SPAD is connected to the AQC. Contrary to the usual AQC-structure, the anode of the SPAD is connected to the quenching circuit and the cathode is supplied with a constant bias voltage. The five main blocks of this circuit are the SENSE STAGE, HIGH-SIDE LOGIC, LOW-SIDE LOGIC and the high voltage transistors M_Q and M_R . The separation of the control logic into a high-side and low-side portion allows the use of fast low voltage transistors for the control logic. All the logic blocks use a rail-to-rail supply of 1.8 V independent of the voltage swing for the SPAD. This guarantees that all the gate-source voltages for the high voltage transistors stay within the technological limit. By closer examining the structure of Acconcia's quenching circuit, it becomes obvious that this circuit has a few issues, which limit the quenching speed. First, the anode of the used RE-SPAD has a larger physical structure than the cathode, as can be seen in the cross-section in Fig. 1.2. The anode is built out of a p+ epitaxy layer above the substrate and is therefore forming a large parasitic capacitance against the substrate. During quenching, the transistor M_Q has to discharge this parasitic capacitance and the SPAD-capacitance very fast to quench the avalanche current. There comes the second issue into play. Due to the structure of this circuit the quenching transistor is a p-channel metal-oxide semiconductor field effect transistor (MOSFET). The majority carriers in p-channel MOSFETs are holes and their mobility in silicon is typical around one third of the electron mobility [9]. This reduced carrier mobility of p-channel MOSFETs decreases also the switching speed compared to n-channel transistors. This means the lower carrier mobility in combination with a large parasitic capacitance limits the quenching speed of this design.



Fig. 1.8: Structure of Acconcia's AQC design [3].

1.4.2 High-Voltage Active Quenching and Resetting Circuit in 0.35 μm CMOS

Dervić et al. propose a High-voltage active quenching and resetting circuit for SPADs in $0.35 \,\mu\text{m}$ CMOS for raising the photon detection probability in [6]. The proposed circuit in Fig. 1.9 uses a modified version of the comparator COMP from [12] in combination with a three stage high-voltage cascode switch (HVCS). The high-voltage cascode switch is constructed out of isolated 3.3 V logic-MOSFETs to increase the quenching voltage swing to 9.9 V. This approach allows to use fast logic-level MOSFETs which can quench the avalanche very fast but limit the voltage swing to three times the nominal voltage of the MOSFETs. To keep the gates of the transistor within their technological limit the high-voltage cascode switch (HVCS) needs adaptive gate biasing. If one wants to increase the quenching voltage swing for example up to 50 V like in [3] by using low voltage MOSFETs the complexity of the HVCS would rapidly increase. The required adaptive gate biasing circuit would introduce further delays and would slow down the quenching speed. This approach is only feasible if a quenching voltage of two to three times the nominal supply voltage is sufficient.

1.4.3 Monolithic AQC for Single-Photon Avalanche Detectors

Zappa et al. [10] proposed one of the first integrated AQC, according to the authors, designed in a high voltage 0.8 µm CMOS process. The used process allows the circuit



Fig. 1.9: Schematic of the high-voltage AQC in [6] with the HVCS.

to generate quenching voltage pulses up to 20 V. The goal of this work was to build a general purpose quencher, which can work with any available SPAD detector. The input sense stage as key part of the integrated AQC is shown in Fig. 1.10. The accompanying control logic and hold-off circuit is built as standard 5.0 V logic and not shown. In



Fig. 1.10: Schematic for the sense stage of Zappa et al.'s integrated AQC [10].

Fig. 1.10 it can be seen that the avalanche detection is realized via the amplifier P_1 and thereby fixing the detection threshold to the threshold voltage of the used high voltage MOSFET. Another disadvantage of this circuit is the usage of symmetrical high voltage MOSFETs for the detection which are rather large to support the voltage and therefore have higher parasitic capacitance (see Section 1.5). This parasitic capacitance in combination with the fixed threshold voltage increases the reaction time of the circuit. Therefore, the delay until the start of the active quenching phase is more than 20.0 ns. The only adjustable parameters in this circuit are the hold-off time and the quenching voltage. Additionally, the control logic has a GATE-input to deactivate the detector. The missing ability of adjusting the reset time is imposing an upper limit for the usable SPAD-detector in terms of its parasitic capacitance C_{SPAD} . With the given size of the switch S_{RESET} , only a certain amount of charge can be transferred into C_{SPAD} during the fixed reset time, SPADs which require more charge will not properly work with this AQC.

1.5 High-Voltage CMOS Process

For this thesis the LF15A process from the manufacturer *LFoundry* with the Process Design Kit / Process Development Kit (PDK) LF15Ai in revision 1.5.0 PRE3 is used. The LF15A process is a modular 150 nm RF CMOS process with 6 aluminum metal layers and three different gate oxide thicknesses for the voltage domains 1.8 V, 3.3 V and 5.0 V. Additionally, the LF15A PDK includes the HV LDMOS 40 V transistors, which are essential for the goal of this thesis, which is designing a high voltage AQC. There are two commonly used structures for high-voltage MOSFETs, as shown in Fig. 1.11. One is the asymmetrical lateral double-diffused MOS (LDMOS) and the other one is the symmetrical high-voltage MOS (HVMOS) [13]. *LFoundry* offers n- and p-channel asymmetrical LDMOS with a nominal gate voltage of 3.3 V in its LF15Ai PDK. Therefore, the LDMOS is compatible with the 3.3 V logic domain. A LDMOS



Fig. 1.11: Principle structure of the two different high voltage MOSFET-devices [13, Fig. 5].

has a lightly doped drift region between channel and drain contact, which is the ndoped region in the top diagram in Fig. 1.11. This drift region enlarges the depletion zone into the n- doped region and thereby consuming most of the potential difference applied between drain and source. Thus, the potential in the proximity of the gate oxide is low enough that no oxide breakdown can happen, because the gate oxide can only withstand voltages corresponding to the limits for the nominal gate voltage of 3.3 V. Due to this asymmetrical structure only the drain can withstand higher voltages against source and bulk [14]. Whereas the voltage between source and bulk, and between gate and bulk are limited to nominal 3.3 V. Compared to a symmetrical HVMOS-structure, a LDMOS is built with a shorter channel length and the same oxide as its compatible low voltage MOSFET [15]. The short channel causes less gate-source capacitance, therefore the LDMOS can be switched faster. Since these LDMOSs are optimized for switching applications their matching with regard to analog design is very poor. The MOSFET-symbols for the different gate oxides and corresponding voltage levels which will be used in Chapter 2 and 3 are depicted in Fig. 1.12.



Fig. 1.12: MOSFET-symbols for the different voltage levels or gate oxides.

Additionally, the LF15Ai PDK includes two different isolation schemes to isolate devices in a PWELL from the p-type substrate. This is especially important when p-type LDMOS are used to isolate their driver circuitry. One scheme uses a n-layer implant NISO to isolate single devices from the substrate. The other scheme uses a deep n-well implantation DNWELL as bottom isolation enclosed within a NWELL and NISO ring to isolate more devices as circuit block. The DNWELL isolation can be used up to voltage levels of 60 V.

Chapter 2

Design Approaches

2.1 Basic Structure of the High Voltage AQC

The basic design approach for this high voltage AQC design is inspired by the structure in Fig. 1.8 from Acconcia, Rech, Gulinatti, *et al.* [3] and this initial structure is shown in Fig. 2.1. Compared to the structure in Fig. 1.8 the SPAD-cathode is now connected to the AQC. Therefore, the quenching transistor is now a n-channel MOSFET, which has a higher carrier mobility and should therefore quench the avalanche faster, as mentioned in Section 1.4.1. Due to this configuration, most of the control logic is in the *HIGH-SIDE LOGIC*-block accompanying the *SENSE STAGE*. This control logic and the *SENSE STAGE* will be placed inside a DNWELL-isolation block marked as dashed orange rectangle in Fig. 2.1. This structure of the AQC raises one important requirement for the *SENSE STAGE*, which is to protect the low-voltage sensing circuitry and control logic from the high quenching voltage swing at the SPAD-cathode terminal.

This high-voltage AQC will be designed for external SPADs especially the type SAP500 from Laser Components [16] and C30902SH from Excelitas [17]. These commercial thick SPADs have a capacitance C_{SPAD} in the range of a few pF and their cross section is similar to the one in Fig. 1.3. To have some margin a capacitance of 5.0 pF is chosen for designing the AQC. As mentioned in Section 1.2, the avalanche charge is an important parameter for quenching circuits. For this design the high quenching voltage in combination with a SPAD capacitance in the pF-range causes a large charge stored in C_{SPAD} . At least the excess charge of $Q_{EX} = C_{SPAD} \cdot V_{EX}$ must be removed to successfully quench an avalanche and lower the bias voltage V_{Bias} down to the breakdown voltage V_{Break} of the SPAD. Therefore, the AQC to be developed should detect and quench an avalanche as fast as possible to prevent that all the charge Q_{EX} flows through the depletion region of the SPAD. This will be the main focus in the following quencher design with the intention to reduce the afterpulsing probability (APP) [12].



Fig. 2.1: Basic structure of the AQC-design inspired by [3]

In this thesis, the term quenching voltage is used to describe the voltage swing the quencher applies onto the SPAD-cathode. In the case of Fig. 2.1 the quenching voltage V_q is equal to the supply voltage $V_{dd,HS}$ which is the highest potential in the circuit. This term is used to distinguish the quenching voltage V_q from the excess bias voltage V_{EX} which is defined as difference between SPAD ready state bias voltage V_{Bias} and breakdown voltage U_{Break} . The distinction is made because an active quencher can lower the bias voltage during quenching below V_{Break} , than the quenching voltage is greater then V_{EX} . The potential of V_{Anode} must be negative with respect to GND to reverse bias the SPAD. The ready state bias voltage is then given as $V_{Bias} = V_q - V_{Anode} = V_{dd,HS} - V_{Anode}$, which is also equal to V_{EX} the potential V_{Anode} must be $-V_{Break}$. Thus, in operating conditions where V_q is set equal to V_{EX} the potential V_{Anode} must be $-V_{Break}$ for proper reverse biasing the SPAD. In a general case the potential V_{Anode} is greater or equal to $-V_{Break}$. The safe operating area of the n-channel LDMOS limits the usable quenching voltage range to 35 V. This means the maximum excess bias voltage V_{EX} will be 35 V when V_{EX} is equal to the quenching voltage V_q .

Initially, for the supply voltage of the *HIGH-SIDE LOGIC*-block a value of 1.8 V is chosen, which allows to use the faster 1.8 V-MOSFETs for the *SENSE STAGE* and the *CONTROL LOGIC*. The voltage level of the gate signals for the p-channel LDMOS will

also be 1.8 V despite the fact they could also be controlled via a 3.3 V signal. Thereby avoiding an additional delay of the necessary level shifter to drive the gate of M_R . To exploit the full potential of the n-channel LDMOS quenching transistor M_Q a voltage level of 3.3 V will be used for the *LOW-SIDE LOGIC*-block.

2.2 SPAD Simulation Model

One of the major prerequisite for designing a high voltage AQC is a proper electrical model of the SPAD. This model should capture the electrical behavior of a SPAD from the viewpoint of the quenching circuit, to use it for simulation of the design and evaluation of different quenching circuit structures. An accurate SPAD model is especially important for designing and optimizing the *SENSE STAGE*.

For the first design iterations the simple SPAD model in Fig. 2.2 was used. The



Fig. 2.2: Simple SPAD model.

controlled current source simulates a photon triggered avalanche current. Therefore, the current source generates a current pulse with 1.0 ns duration. The current can be derived from the avalanche charge $Q_{av} = M \cdot e$ with an intrinsic avalanche gain of $M = 10^8$ according to [17], the avalanche charge is $Q_{av} = 16.02 \,\mathrm{pC}$. In comparison, the charge of the SPAD-capacitance when the excess bias voltage V_{EX} is equal the maximum quenching voltage of 35 V is $Q_{EX} = C_{SPAD}V_{EX} = 175 \,\mathrm{pC}$. Only the quenching voltage is used to calculate the charge because this is the voltage swing during quenching of an avalanche and only this portion of the charge is removed from C_{SPAD} . In this case, the simulated avalanche only removes a tenth of the excess charge from the SPAD capacitance C_{SPAD} , thereby causing a very shallow voltage slope at the SPAD cathode. Due to the limited time of 1.0 ns the avalanche current flows, there is the possibility that the simulated AQC does not detect the avalanche in case the voltage drop at the SPAD cathode is too low. Thus, this model is only useful to simulate the onset of the photon triggered avalanche. First simulations of the principle quencher structure showed that the active quenching time is roughly 1.0 ns when quenched with a 1.0 mm wide LDMOS. Due to additional unavoidable delays in the real circuit it is obvious

that a more accurate SPAD-model is preferable in order to be able to optimize the quenching circuit. Furthermore, the model should have the ability for passive quenching to check if the designed quencher really actively quenches the SPAD.

One promising SPAD-model for circuit simulation was proposed by Mora et al. [18]. An adopted version, which is used for all further design iterations of this high voltage



Fig. 2.3: SPAD simulation model adopted from the model in [18]

AQC design is shown in Fig. 2.3. The bidirectional terminals Cathode and Anode represent the corresponding physical cathode and anode connection. The terminals *Photon, Avalanche* and *COM* have no electrical counterpart and are only used for simulation purpose to control the model. The original model in [18] simulates the SPAD I - V characteristic with a piecewise linear approximation through a nonlinear voltage generator. This I - V characteristic is approximated much simpler in Fig. 2.3 through the direct current (DC) voltage source V_{Break} for the break down voltage and the resistor R_{SPAD} for the voltage slope at breakdown. From a physical point of view R_{SPAD} captures the diode resistance of the space-charge region and the neutral region which are passed by the avalanche current [18]. C_{SPAD} captures the SPAD-capacitance and all parasitic capacitances against substrate, contrary to the original where the substrate node and corresponding parasitic capacitance are modeled separately. In the proposed AQC-design the anode of the SPAD is always connected to the constant negative V_{Anode} -potential. Therefore, all parasitic capacitances against substrate are modeled by one capacitor. For the value of R_{SPAD} Cova et al. [19] state a value of less than 500 Ω for thick depletion layer SPADs which closely matches with the values Mora
et al. [18] used. From the measurement results of a thorough characterization of the SAP500 in [20] depicted in Fig. 2.4, a value of approximately 800Ω can be extracted.



Fig. 2.4: Single photon response at the output of a PQC for two discrete SPADs [20].

To simulate a photon triggered avalanche in the model of Fig. 2.3 a short logic pulse is applied between the terminals *Photon* and *COM*. The duration of the logic pulse should be below 1 ns, usually a value of 500 ps is used. Due to the ideal voltage controlled switches SW_1 and SW_2 the common potential on terminal *COM* for the control signal *Photon* and *Avalanche* can be freely chosen. This allows convenient comparison between the simulated photon pulse and the response of the AQC. The switch SW_1 initiates the avalanche current $I_{Avalanche}$ when switched on. To simulate the self-sustaining behavior of the SPAD the current controlled voltage source G_1 enables switch SW_2 . The transimpedance h and the switching threshold of SW_2 are therefore chosen that switch SW_2 stays closed as long as the avalanche current is higher than the "quenching threshold" of $100 \,\mu\text{A}$ [2]. Thus, the feedback circuit out of G_1 and SW_2 simulates the self-sustaining and self quenching behavior of a SPAD. The voltage of G_1 can be examined via terminal *Avalanche* during simulation to see if the designed AQC actually quenches the SPAD. The final parameters used for the SPAD-model during development of the quenching circuit are given in following table.

Tab. 2.1: Simulation parameter used in the SPAD-model in Fig. 2.3 for the design of the AQC.

R_{SPAD}	800Ω
C_{SPAD}	5 pF
V_{Break}	120 V

2.3 Avalanche Detection

The SENSE STAGE in Fig. 2.1 has an important role in the overall quenching circuit. First, it needs to detect the avalanche as fast as possible and initiate the quenching. Second, the SENSE STAGE has to protect its own circuitry and the following control logic from the high quenching voltage, which would otherwise destroy the low voltage MOSFETs due to gate breakdown. A common approach in literature is to sense the avalanche by detecting a voltage change across a kind of shunt resistor. This approach is called voltage-mode sensing (VMS) [2]. Another less-used approach is the direct detection of the avalanche current via a current mirror or current comparator or similar current sensing circuitry. This technique is called current-mode sensing (CMS). In the following, different approaches for detection of the avalanche are examined and the final approach will be developed, which is used in Chapter 3.

2.3.1 Voltage-Mode Sensing

For the basic VMS approach, a adapted version of the detection circuitry in the design from Enne, Steindl, Hofbauer, et al. [12] is used. Contrary to [3] where the avalanche is sensed via a common source amplifier with a similar circuit as in Fig. 1.10. In [12] the avalanche is sensed via a fast free running comparator and a current mirror on whose output impedance the avalanche current is translated into a voltage. The detection threshold $V_{Reference}$ can be freely adjusted down to 100.0 mV in contrast to [3] where the detection threshold is fixed to the gate source voltage of the amplifier. The adaption of the detection circuit from [12] for high quenching voltages is shown in Fig. 2.5. The output impedance of the low voltage transistor M_{sense} is used as current shunt to translate the avalanche current into a voltage drop. This voltage drop is then detected by the fast comparator U_1 with an adjustable reference voltage $V_{Reference}$, reviewed in Section 2.5. The LDMOS $M_{protect}$ prevents the potential at the input of comparator U_1 from going below $V_{ss,HS}$ to keep the drain-source voltage of M_{sense} within its technological limits. The minimum width of the protection LDMOS is technologically limited to $20.0\,\mu\text{m}$. Thus, this transistor is introducing a relatively large parasitic capacitance between its drain and source compared to low voltage MOSFETs. During the fast transitions, when M_Q in the test circuit of Fig. 2.6 is switched on, this drain source capacitance causes voltage spikes at the drain to source voltage of M_{sense} . To keep this voltage spikes within the technological limits, the gate of $M_{protect}$ is connected to an adjustable DC voltage source instead of connecting it to $V_{ss,HS}$. Due to the lower gate voltage, $M_{protect}$ switches into OFF-state earlier when the drain source voltage of M_{sense} rises. Furthermore, it is necessary to increase the width of



Fig. 2.5: Basic VMS stage adapted from [12].

 M_{sense} to keep its drain-source voltage within the technological limits. This in turn reduces the voltage drop on the negative input of U_1 at the onset of the avalanche and therefore increases the detection time delay. The described method of protecting the low voltage MOSFET from the high quenching voltage can be viewed as a passive protection method, because none of the involved transistors are actively switched when the voltage rises during an avalanche. This fact will be changed in Section 2.3.2.

The basic AQC in Fig. 2.6 is used for evaluating the VMS stage of Fig. 2.5. For the evaluation of this VMS-stage, in addition to the sense stage only the control logic, the quenching-MOSFET M_Q and the reset-MOSFET M_R are realistically modeled. The level shifter and buffers are ideal parts with no internal delay. Of interest in this circuit is the current contribution of all components connected to the SPAD-cathode during onset of the avalanche, which is shown in Fig. 2.7. On the green trace it can be seen that the largest contribution to the avalanche current comes from the switched off reset-MOSFET M_R . This is due to the relatively large width and therefore large drain-source capacitance compared to the MOSFETs in the SENSE STAGE. This means that most of the avalanche current will be shunted to $V_{dd,HS}$ through M_R and is not available for the SENSE STAGE and will therefore reduce its sensitivity.

2.3.2 Combined Sense Stage and Reset Switch

The issues with the voltage-mode sensing (VMS)-stage led to the consideration to use the protection-MOSFET $M_{protect}$ and sensing-MOSFET M_{sense} also for recharging the SPAD during the reset phase by increasing their width. It is a similar sensing method as used in Fig. 1.9 with the HVCS [6]. The main reason for this approach is



Fig. 2.6: Test circuit for the VMS sense stage of Fig. 2.5

to guide most of the avalanche current through the sense stage. Another reason is the technologically enforced minimum width of a LDMOS which requires the use of a larger protection-MOSFET $M_{protect}$ than necessary. An obvious advantage of this approach is the reduction of transistors directly connected to the SPAD, at least by one, thereby reducing parasitic capacitance at the SPAD-cathode node, which otherwise would slow down the sensing circuit. The principle design of the combined sense and reset stage is shown in Fig. 2.8. Contrary to the voltage-mode sensing (VMS) stage in Fig. 2.5 the transistors M_{sense} and $M_{protect}$ are now actively controlled by the control logic, which is necessary to switch between sensing mode and reset mode. The control logic will be discussed in Section 2.7. The dashed connection in Fig. 2.8 is for demonstrating how the comparator U_1 and the control logic will control the transistors during the detection of an avalanche to protect itself and M_{sense} from the high voltage. In this configuration (dashed line) the circuit could be used as SENSE STAGE but then resetting has to be carried out via a separate transistor like M_R in Fig. 2.6.

To the operating principle of this stage: In the ready state, when the SPAD is biased above breakdown, the sense transistor M_{sense} is in the *OFF* state and ready for detecting an avalanche. Whereas the protection transistor $M_{protect}$ is in the *ON* state when waiting for an avalanche. The M_{sense} *OFF*-resistance is the shunt resistor



Fig. 2.7: Current into SPAD-cathode during onset of an avalanche and its branch currents for VMS-stage in Fig. 2.5.

at which the voltage drop caused by the avalanche current will be detected by the comparator U_1 . Immediately after the detection, the control logic switches M_{sense} on and $M_{protect}$ into OFF state. Now the switched off protection transistor consume the high quenching voltage and keeps the low voltage sense transistor and the comparator within the technological limits. Switching on M_{sense} reduces its drain source resistance and thus helps to keep its drain source voltage below the limits during active quenching. Especially a capacitive coupled voltage spike from the fast current transition during quenching is much lower in this approach compared to the VMS-stage, due to the low ON-resistance of M_{sense} . When the hold-off time is over, the control logic switches the protection transistor $M_{protect}$ into the ON-state after switching off the quenching transistor. At this point, both transistors $M_{protect}$ and M_{sense} of the sense stage are in ON-state and the quenching transistor M_Q is OFF. The SPAD can now recharge through $M_{protect}$ and M_{sense} to its ready state bias voltage. The recharge duration T_{reset} is controlled by the control logic, thereby allowing the usage of different SPADs with different capacitance C_{SPAD} . After T_{reset} is over M_{sense} is switched off and the AQC is now ready to detect avalanches again. Fig. 2.9 shows the current contribution from the AQC during onset of an avalanche to the SPAD-cathode current. Compared to Fig. 2.7 most of the current now flows through the sense stage and contributes to the



Fig. 2.8: Schematic of the principle design for the combined sense and reset stage.

detection of the avalanche. During the first 300.0 ps, until the quenching process starts, the cathode current is around 5.0 mA lower compared to Fig. 2.7. This indicates a slightly lower avalanche charge during the passive quenching phase. This time the sense and reset stage in Fig. 2.9 was tested with the real level shifter, which will be discussed in Section 2.6. Therefore, the passive quenching part, while current is flowing into the SPAD-cathode, lasts longer due to the delay of the realistic level shifter. That is the reason why only the first 300.0 ps are compared with Fig. 2.7. The peak at t = 0.4 ns shows the transition of the *OFF*-state from M_{sense} to $M_{protect}$ after detection. The current peak arises from not perfectly synchronized gate signals for both transistors and different switching delays. Obviously M_{sense} switches faster into the *ON*-state than $M_{protect}$ into the *OFF*-state which causes the current peak.

2.4 Dimensioning of Quenching and Resetting Transistor

In this section the dimensioning of the main transistors M_Q , $M_{protect}$ and M_{sense} will be discussed. For a first order approximation of the required width W_Q for M_Q the SPAD in Fig. 2.1 can be replaced with its parasitic capacitance C_{SPAD} . This can be done because during switching operation the load, which is seen from M_Q , is the capacitance of the SPAD. With this equivalent circuit it can be seen that the quenching transistor basically has to discharge a capacitance during a transition. This means at onset of



Fig. 2.9: Current into SPAD-cathode during onset of an avalanche and split into its partial currents. This time with a realistic level shifter.

the switching transition C_{SPAD} is fully charged to its ready state bias voltage and the gate of M_Q is at 3.3 V. Thus M_Q is in the saturation region until its drain-source voltage is below the overdrive voltage, which is true for most of the switching transition. Consequently, the quenching transistor can be approximated as constant current source to estimate the width W_Q for a targeted active quenching time of 0.5 ns. By assuming a constant current source the time derivative of the cathode voltage can be replaced by its finite difference.

$$I_Q = C_{SPAD} \frac{dU_{cathode}}{dt} \to I_Q = C_{SPAD} \frac{\Delta U_{cathode}}{\Delta t}$$
(2.1)

With $-\Delta U_{cathode}$ equal to the maximum quenching voltage of 35 V and $-I_Q$ is given by the drain current of M_Q in saturation according to following Equation:

$$-I_Q = \frac{1}{2}\mu C_{ox} \frac{W_Q}{L_Q} (U_{GS} - U_{th})^2$$
(2.2)

With $U_{GS} = 3.3 \text{ V}$, the carrier mobility times the gate oxide capacitance $\mu C_{ox} = 1.06 \text{ mA/V}^2$, $L_Q = 1.95 \text{ µm}$ and $U_{th} = 0.5 \text{ V}$. Combining Equ. 2.1 with Equ. 2.2 gives a minimum width for the quenching transistor of $W_Q = 163.5 \text{ µm}$. The actually used width in Chapter 3 in Fig. 3.1 is ten times larger than this rough approximation. This

is due to the fact that Equ. 2.2 only roughly describes a modern deep-sub μm CMOS transistor. Especially a high voltage LDMOS with its low doped drain region to support the high voltage behaves differently. One important point which is missed in this approximation is that the parasitic capacitance of M_Q itself imposes a lower limit for the achievable active quenching time. This simple approximation is misleading because it suggests that the active quenching time can be drastically reduced by increasing the width W_Q of M_Q . This suggestion has a major drawback since increasing W_Q enlarges also the gate-source capacitance according to $C_{gs} = W_Q L_Q C_{OX}$. A larger gate-source capacitance requires a stronger gate driver, which is usually built with a chain of inverters with increasing strength as will be discussed in Section 3.1. A larger MOSFET requires a longer inverter chain to achieve the required drive strength, which in turn has a longer delay and therefore increases the overall quenching time. Furthermore, the effect of the parasitic capacitance from $M_{protect}$ is neglected in this approximation since its capacitance will be much lower than C_{SPAD} and at this state of the design it is not known. Therefore, the final dimensions for the three main transistors in Fig. 3.1 are found with the optimizer tool in the *Cadence* EDA design suite. The optimizer tool was used to derive an optimal M_Q -width for an as short as possible overall quenching time. This has to be a compromise between short active quenching time and short driver delay.

The size of the protection transistor $M_{protect}$ can similarly be approximated as it was done for M_Q , but this transistor has to fulfill additional requirements which are contradictory. Therefore, the sizing was done solely via the optimizer tool. First and foremost $M_{protect}$ has to keep the V_{sense} potential within the technological limits of the 1.8 V transistors. The sizing therefore has to be done in consensus with the sizing of the sense transistor. This means the drain-source capacitances need to have a certain ratio to prevent a violation of the technological limits due to capacitive coupling during fast transitions. Furthermore, the transistors $M_{protect}$ and M_{sense} need to be large enough to carry the recharge current during resetting of the SPAD to its ready state bias voltage. The size of the protection transistor is a compromise between a fast resetting phase and a fast quenching time. A fast resetting time requires a large transistor for charging the parasitic capacitance C_{SPAD} , which is preferable to prevent avalanches during the reset phase. This in turn increases the overall capacitance, which M_Q needs to discharge during quenching and thereby slows down the quenching phase, even when increasing the size of M_Q due to the longer driver chain. In favor of a fast quenching transition, $M_{protect}$ is sized in such a way that a reset time of about 10 ns is achieved with the given SPAD-parameters of Section 2.2. To fulfill all these contradicting requirements the optimizer tool in *Cadence* was used to find a suitable solution. For the found solution, it was confirmed with simulations at different process corners, that the voltage drops across the low-voltage transistors, especially $M_{protect}$, stay below their technological limits.

2.5 Comparator

The VMS detection circuits in Section 2.3.2 need a comparator to decide if an avalanche takes place and to translate this information into a logic signal. The used comparator needs to be a free running one and should switch very fast. In general, there are two operation modes of comparators. A continuously operating comparator compares the input voltages all the time and reacts immediately when the signal crosses the reference voltage. This type is usually called free running comparator. Such comparator designs are usually based on an open loop operational amplifier. The other mode is the sampling operation where the signal is compared to the reference only at transitions of a clock signal. The highest feasible frequency for an application-specific integrated circuit (ASIC)-design is around 1.0 GHz. A clocked comparator with a clock frequency of 1.0 GHz would introduce a random quenching delay due to the fact that the photon triggered avalanche are asynchronous to the clock signal. In the worst case, this delay is 1.0 ns. Therefore an existing free running comparator design from Alija Dervić as shown in Fig. 2.10 is used for the final design in Chapter 3. The central part of this



Fig. 2.10: Operational amplifier based free running comparator.

comparator is the differential pair M_7 , M_8 , which amplifies the input signal. The load of the differential pair is the cross-coupled pair of the MOSFETs M_3 , M_4 in parallel with diode-connected MOSFETs M_2 , M_5 . Due to the cross coupling the differential resistance between the drains of M_3 and M_4 becomes negative equal to $-\frac{2}{g_{m3}}$ [21]. The differential resistance between the two diode-connected MOSFETs M_2 , M_5 is $\frac{2}{g_{m2}}$. The parallel connection of the diode-connected MOSFETs M_2 , M_5 and the cross-coupled pair M_3 , M_4 gives a differential resistance of

$$\frac{2}{g_{m2} - g_{m3}}$$

Due to the sizing of the MOSFETs the transconductance g_{m3} of the cross coupled pair is 2.5 times the transconductance g_{m2} of transistor M_2 . Therefore, the overall differential resistance is negative and acts as positive feedback, increasing the speed during a transition. The complementary amplifier consisting of M_6 and M_{12} increases the output swing of this comparator. The bias current is made adjustable through the terminal *Bias*. To protect the gate of M_{11} the diode-connected MOSFET M_{10} is added. With this transistor M_{10} the bias current can be adjusted via a voltage at or via constant current into terminal *Bias*. With this option the bias current can be conveniently adjusted with a potentiometer off chip.

2.6 Level Shifter

Another important part of the design is the level shifter, which is necessary for the M_Q gate signal to translate from the high-side voltage domain to the low-side voltage domain. The task of the level shifter is to shift a 1.8 V logic signal in the high-side voltage domain to a 3.3 V logic signal in the low-side voltage domain. The literature offers a wide range of different level shifter designs optimized for various purposes from which a suitable one has to be selected. The important parameter is the delay because the faster the quenching transistor can be switched on, the sooner the active quenching phase begins, thereby reducing the avalanche charge in the SPAD. A second concern for selecting a level shifter is its power consumption, because for such high voltages to conduct even low currents cause a high power dissipation. The used technology poses also limitations to the possible level shifter design. For example Lutz et al. [22] proposes a low power level shifter for 50.0 V by using capacitors to support the high voltage. But the used process LF15A only offers metal insulator metal (MIM) capacitors with a maximum voltage of 16.0 V. Therefore, this design is not usable. In the following, two tested level shifter designs for the high voltage AQC are reviewed.

The first one, shown in Fig. 2.11, is based on a design proposed by Declerq et al. [23]. It is a very simple level shifter design with only 6 transistors, 2 of them are high voltage transistors to support the high voltage. The main part of the circuit is the combination



Fig. 2.11: Declerq [23] based Level shifter design.

of M_1 and M_3 called "voltage mirror" by Declerq et al. [23]. This voltage mirror, mirrors the input voltage $V_{dd,HS} - V(In_{P,HS})$ to $V(Out_{N,LS}) - GND$. Except in this implementation $V(Out_{N,LS}) - GND$ is scaled up to a 3.3 V logic signal by appropriately choosing the width of the transistor M_3 . The transistor M_4 is connected in parallel to the diode connected MOSFET M_3 as pull-down transistor to drive the output low during the LOW-state. Two of these voltage mirrors are cross coupled via the gates of the pull-down transistor to form a high-voltage flip-flop [23]. Thereby it can be seen that one high voltage transistor is always switched on and will dissipate power. The value of the output voltage is equal to the gate-source voltage of the diode-connected MOSFET M_3 . When keeping the channel length fixed, the output voltage can be set by the width W according to following equation.

$$U_{gs} = \sqrt{\frac{2I_D L}{\mu C_{OX} W}} + U_{th} \tag{2.3}$$

 I_D in Equation 2.3 is determined by the dimensions of the high voltage transistor M_1 . A short delay can be achieved when using the minimum channel width for the high voltage transistor, according to [24]. For the used LF15A the minimum channel width is given to be 20.0 µm, which already places a large load for the driving logic circuit. Another important reason, which speaks for a small channel width W_{pLDMOS} is the power dissipation. The high voltage, the LDMOS has to support is the dominating factor at the power dissipation. One reason for a larger channel width W_{pLDMOS} for the LDMOS is the output drive strength, which is an interesting consideration especially for this AQC design. A higher drive strength of the level shifter can save an inverter in the driver chain for the quenching transistor and therefore reduce the overall delay. In [24], it is suggested to size the transistors M_3 and M_4 equally. The sizing of the pull-down transistors is a compromise between delay and output drive strength. Obviously, a larger size of the pull-down transistor increases the drive strength at the cost of a longer delay and vice versa. The final sizes of the transistors in Fig. 2.11 were found with the help of the optimizer tool from *Cadence*. For the HIGH-level output at the lowside $V(Out_{N,LS})$ and $V(Out_{P,LS})$ a value of around 3 V was chosen, at typical process parameters, this leaves a margin for process, voltage and temperature (PVT)-variations to the technological limit of 4.0 V for 3.3 V-MOSFETs. This margin was confirmed via Monte Carlo simulation as shown in Fig. 2.12. With these sizes of the transistors



Fig. 2.12: Histogram for changes in the peak output voltage of the level shifter due to PVT-variations, evaluated with Monte Carlo simulation.

the resulting delay from $In_{P,HS}$ to $Out_{N,LS}$ is about 100 ps and the constant current flowing through the active branch is around 2.4 mA. With the maximum quenching voltage of 35 V this current will cause a continuous power dissipation of around 90 mW in the active branch of the level shifter. The fast response of the level shifter in Fig. 2.11 comes at the cost of a rather large power consumption.

The second level shifter design tested for this high voltage AQC shown in Fig. 2.13 is based on [25], [26]. This level shifter design was tested due to concerns regarding the



power consumption of the level shifter design in Fig. 2.11 based on [23]. Therefore, this

Fig. 2.13: Moghe [25], [26] based Level shifter design.

design was evaluated, because it has no static power consumption and it is especially suitable for asymmetric thin-oxide LDMOS [25] as are available in the PDK. This circuit is a modified version of the basic differential cascode voltage switch low voltage level shifter for high voltage applications. The main part are the four LDMOS M_1 to M_4 and the cross coupled low voltage transistors M_5 and M_6 taken from [25]. In this design the n-channel LDMOS are responsible for clamping the low-side voltage to $V_{dd,LS}$ and sustaining the high voltage in combination with the p-channel LDMOS. The inverters in the cross coupled feedback are a modification to improve the transition speed as suggested in [25], [26]. The inverters I_1 and I_4 are specially sized to have a very low threshold voltage, so they can react on the initial rise of the output voltage, the reason for that is as follows: Consider the initial state when the input $In_{P,HS}$ is HIGH and $In_{N,HS}$ is LOW. This input state requires the output to have following logic levels $Out_{N,LS}$ is LOW and $Out_{P,LS}$ is HIGH. For the following considerations only the left branch of the circuit in Fig. 2.13 is observed. This means M_5 is switched on by the HIGH level from the regular inverter I_3 and the impedance from $Out_{N,LS}$ to GND is the ON-resistance of M_5 . Thus the source of M_3 is pulled to GND and therefore M_3 is also in the linear region. To flip the state the logic level LOW is applied to the gate of M_1 and the logic level HIGH to M_2 . Now the drain current from M_1 needs to generate a high enough voltage drop across the drain-source resistance of M_5 to trigger the inverter

 I_1 and toggle the latched state of the cross coupled transistors M_5 and M_6 . To react faster on the initial rise of this voltage drop the inverters I_1 and I_4 are custom designed to have a low threshold voltage as already mentioned. During this transition the rising voltage at $Out_{N,LS}$ reduces the gate voltage of M_3 and therefore limits the current in the branch. This limited current, since both LDMOS M_1 and M_3 are in saturation, is one of the reasons why this type of level shifter is known to be one of the slowest [24]. Because during a transition the voltage at node V_D has to go through a full swing from $V_{dd,HS}$ to GND, which is equal to the applied high voltage and node V_{DM} has to do the transition in the opposite direction. Thus the parasitic capacitance, mostly caused by the high voltage transistors, associated with this nodes need to recharge to fulfill this voltage swing. While in the left branch both high voltage transistors are active and contribute to recharging the parasitic capacitance, in the right branch only the n-channel LDMOS is active in the mentioned state transition. Leading to a shallower voltage slope at node V_{DM} compared to node V_D resulting in different delays in the two branches during a transition. Due to these different delays the outputs exhibit a stair step rising or falling edge with an intermediate state during transition which is known as "Miller plateau" [26] (see simulation results in Fig. 2.14). Simulations also showed asymmetrical delays for the rising and the falling edge and a delay for the rising edge at the output of around 220 ps at best. The best achievable delay is twice as long as for the Declerq-based level shifter in Fig. 2.11 and only for the rising edge. The stair step edge of the output makes it also harder to predict when a following stage will interpret this as a state change especially in context of PVT-variations. Therefore, and because of the longer delay time this level shifter design was not used and the Declerq-based design was used due to its faster response.

2.7 Control Logic

The control logic is the brain of the AQC, where all the input signals are processed and the gate and output signals are generated. The control logic is responsible for switching the quenching MOSFET and the transistor of the VMS and reset stage in the correct sequence with the right timing as explained in Section 2.3.2. This sequence is captured in the state diagram in Fig. 2.15. The control logic is usually in the *READY*-state when waiting for a photon triggered avalanche. When the comparator from Section 2.5 detects an avalanche it is signalized via the *Photon detect*-signal. The control logic immediately generates the correct gate signals for the MOSFETs and transits into the state *QUENCH*. The control logic is also responsible to generate the time delays for the hold-off time and reset time for the time dependent state changes. When the hold-off



Fig. 2.14: Output signals for the Moghe based level shifter with the stair step like edge clearly visible.

time is over the control logic changes into the RESET-state and produce the right gate signals. After a delay of the adjustable reset time the state changes again into the READY-state and the circuit is ready again to detect another photon.

An additional *Gate*-signal is available, for transitioning into the *QUENCH*-state manually thereby deactivating the SPAD by discharging it down to or below the breakdown voltage. With this bias voltage V_{Bias} the SPAD is essentially blind for photons. The *Gate*-input is especially useful during the start-up phase when applying the bias voltage to the SPAD to protect the circuit.

The control logic needs to react fast on an asynchronous input impulse from the comparator. Therefore, it cannot be built as sequential logic, which is synchronous to a clock signal. Instead, the control logic is build as asynchronous logic without any clock signal and the storage behavior is realized via delay elements and feedback.

Both delays, the hold-off time and the reset time, are made variable in the range of 1.0 ns–100.0 ns and can be set via an external resistor. The largest delay time is twice the hold-off time used in [20] for their passive quenching circuit (PQC), which should be sufficient for an AQC.

An inverter with adjustable fall time is designed as shown in Fig. 2.16 [12], which are required for the adjustable delay elements. When the input signal A goes HIGH the MOSFET M_1 is switched off and the metal-oxide semiconductor (MOS)-capacitance



Fig. 2.15: State diagram for the control logic

 M_4 is discharged via the constant current source M_3 . This current can be adjusted by the voltage applied to the terminal *Delay* or with a potentiometer connected between $V_{dd,1V8}$ and the *Delay*-terminal. The voltage at the *Delay*-terminal must be between 0.0 V and 1.8 V, referenced to $V_{ss,1V8}$. When 1.8 V is applied to the *Delay*-terminal the shortest delay of about 1.0 ns is selected. When no potentiometer or voltage source is connected at the *delay*-terminal, resistor R_1 sets the delay to about 100.0 ns. The adjustable current through M_3 and the gate capacitance of M_4 define the fall time of the output Q.

In combination with an AND-gate, like for example U_4 in Fig. 2.17, the inverter with adjustable fall time forms an edge detection circuit with adjustable pulse length for rising edges, which is used as central storage and delay element in the control circuit in Fig. 2.17. The control circuit is built around the variable delay elements U_4 and U_8 which form pulse generators with the AND-gates U_5 , U_9 and the *feedback*-line. After the *photon detect* input goes HIGH the pulse generator U_4 and U_5 starts to generate the hold-off time pulse. Because the input pulse at *photon detect* is shorter than the hold-off time, the feedback network with the OR-gates U_{10} , U_3 and the *feedback*-line is necessary to latch the detection state. This means the *feedback*-line is HIGH as long as the control logic is in state QUENCH or RESET. The inverted *feedback*-signal is



Fig. 2.16: Schematic of a inverter with adjustable fall time



Fig. 2.17: Schematic of the actual implemented control logic.

also used as gate-signal for the sensing-MOSFET M_{sense} . When the hold-off time is over or the gate-signal goes LOW, the control logic transits into the *RESET*-state and U_8 , U_9 generate the reset time pulse. After the reset time is over the feedback-signal goes LOW thereby changing into *READY*-state and enabling the circuit to detect another avalanche. During *QUENCH* and *RESET* state the control logic is blocked for detecting further avalanches due to the feedback-input at U_3 is HIGH and thereby no changes on photon detect input can alter the output. This is necessary because the comparator also produces an output pulse during the reset-phase, which would otherwise interfere with the ongoing quenching sequence. The gate-signal Quench ON for the quenching-MOSFET M_{quench} could be tapped of the hold-off signal but this signal has a longer delay compared to gate-signal generated through U_1 and U_2 . To speed up the quenching, the *photon detect* signal is fed forward to the selection circuit U_1 and U_2 to generate a gate signal with the shortest possible delay. The same is done for the $M_{protect}$ gate-signal *Protect OFF* to rapidly turn off the transistor. Turning off the quenching-MOSFET M_Q and switching on the protection-MOSFET $M_{protect}$ when transitioning from the *QUENCH*- into the *RESET*-state happens theoretically at the same time. In reality the level shifter introduces a delay onto the gate signal for M_Q . To prevent cross conduction between the quenching MOSFET M_{quench} and the protection MOSFET $M_{protect}$, an additional constant delay U_{12} is inserted. U_{12} delays the falling edge when switching on $M_{protect}$ by approximately 1.0 ns to compensate for the delay of the level shifter.



Fig. 2.18: One quenching sequence simulated with equal hold-off and reset time.

2.8 Output Driver

Additionally, to all this circuitry for detection and quenching an avalanche, an output driver with enough drive strength is required to signalize an avalanche event to off-chip circuitry. Therefore, an output driver with the capability to drive a coaxial cable with 50Ω characteristic impedance and the 50Ω input of an oscilloscope is needed. For this task, a simple driver built of cascaded inverters with increasing drive strength was



Fig. 2.19: Schematic for the 50Ω output driver.

designed as shown in Fig. 2.19. The numbers in each inverter give their drive strength in multiple of unit standard cells. Further details on how such inverter chains are designed will be discussed in Section 3.1 when the sizing of the gate driver is explained.

2.9 Thermal Considerations

The level shifter selected for the final design is based on the design from Declerq et al. [23], because it provides the fastest response. As already mentioned in Section 2.6, this fast response comes at the cost of a high power dissipation. The power consumption of the level shifter is the dominating contribution to the overall power consumption of the whole chip, as will be shown in Chapter 4. For experimental verification of the design, the fabricated circuit will be glued with its backside onto a printed circuit board (PCB). In this setup, the generated heat needs to dissipated through the silicon die to its backside to keep the junction temperature within reasonable limits. Due to its thermal conductivity λ of 148 W K/m [27] only the silicon die is taken into account for following approximated thermal considerations. The contributions of bond wires and the surrounding air to the heat flow are neglected. During the quiescent state the

active one of the two LDMOS transistors from the level shifter in Fig. 2.11 contributes mainly to the heat generation. This means the heat sources are not distributed over the entire chip surface but instead concentrated in one point. To estimate the thermal resistance between this heat generating transistor and the backside of the chip, this heat source is approximated as a sphere with a radius of half the channel width and the center positioned on the top surface of the chip. The thermal resistance from this spherical heat source to the bottom side of the chip will be approximated by assuming that the heat will flow in radial direction into a lower half spherical part of the chip with a radius equal to the thickness of the chip. The size of this sphere in horizontal direction is much smaller than the horizontal dimension of the overall chip. Therefore, the remaining silicon outside the sphere is assumed to have equal temperature. The silicon chip can be assumed to have a homogeneous and isotropic thermal conductivity and Fourier's law [28] can be written in the stationary case as

$$\dot{q}(r) = -\lambda \frac{\partial T(r)}{\partial r}.$$
(2.4)

Where T(r) is the local temperature at a radial distance r from the center of the sphere. The local heat flux density $\dot{q}(r)$ is equal to

$$\dot{q}(r) = \frac{P_q}{2\pi r^2} \tag{2.5}$$

with P_q the quiescent electrical power of the transistor which will be spread in radial direction over half the surface of the sphere. Integration of Equ. 2.4 over the radius from $r_S = 6.25 \,\mu\text{m}$, which is half the gate width of a LDMOS, to $r_t = 200 \,\mu\text{m}$, which is equal to the final thickness of the chip, will give a thermal resistance of

$$R_{th} = \frac{T(r_S) - T(r_T)}{P_q} = \frac{1}{2\pi r_S \lambda} - \frac{1}{2\pi r_T \lambda} = 166.7 \,\mathrm{K/W}.$$
 (2.6)

That means an electrical power consumption of 90 mW will increase the local temperature of the active transistor by 15 K. If on the other hand a more distributed heat source over the whole chip is assumed, which is the case during the quenching and resetting phase when M_Q or $M_{protect}$ are active. The thermal resistance in this case is more closely approximated as the thermal resistance of a silicon body with a cross section equal to the area A of the circuit and a thickness equal the final chip thickness $t = 200 \,\mu\text{m}$.

$$R_{th} = \frac{t}{A\lambda} = 33.8 \,\mathrm{K/W} \tag{2.7}$$

For the area A the final dimensions of the circuit layout in Fig. 3.2 are used. Overall, these are rough approximations of the thermal resistance and only be used to check if there is a potential danger of exceeding the maximum allowed junction temperature of 150 °C, according to the PDK. The real thermal situation will be much more complex due to numerous environmental factors, which were not taken into account for this rough approximation. Such environmental factors are the metallization on top of the device like the number of metal layers and their shape and also the location of the pads the number of bond wires.

Additionally, the *hisim*-model [13] used for the LDMOS-transistors provides support for thermal considerations by modeling the self-heating effect and calculating the local temperature change. For the transistor in question with a quiescent power consumption P_q of 90 mW the simulation gives a temperature rise ΔT_{local} of around 37.4 K. Combining both results gives a local junction temperature T_j of

$$T_j = \Delta T_{local} + P_q R_{th} + T_{amb} = 82.4 \,^{\circ}\mathrm{C}$$

$$(2.8)$$

for an ambient temperature T_{amb} of 30 °C, which is the temperature of the backside of the chip. For the thermal resistance the value from the spherical approximation is used, which should cover the worst case. This means it is preferable to provide sufficient cooling when the fabricated circuit will be evaluated.



Chapter 3

Implemented AQC Design

3.1 Implemented Design

This chapter will describe the final implemented design of the complete high-voltage active quenching circuit (AQC) shown in Fig. 3.1. The *CONTROL LOGIC* block implements the control circuitry from Fig. 2.17. The three control inputs *Gate*, *Reset time* and *Hold-off time* are connected to Pads, so that these parameters can be adjusted off-chip. For the *LEVEL SHIFTER* block the circuit in Fig. 2.11 is used because of its fast response. To detect an avalanche and resetting the SPAD to its ready state bias voltage, the combined VMS and reset stage from Section 2.3.2 is used. The corresponding circuit elements are enclosed in the green dashed rectangle in Fig. 3.1.

The three main transistors M_{sense} , $M_{protect}$ and M_Q are rather large in comparison to the transistors of the *CONTROL LOGIC* which provides the gate signals for them. Therefore, driver circuits for the gates are necessary, usually built as a chain of inverters with increasing size. Proper selection and design of the gate driver is important, because they introduce additional delay which increases the overall delay from detection until active quenching begins. The used design kit LF15A provides libraries with digital gates for the voltage domains of 1.8 V, 3.3 V and 5.0 V. These libraries include inverters and buffers with different drive strength, specified as multiple of unit standard cells they can drive. The design kit offers two kinds of inverters, regular inverters *INV* and clock inverters *CLKINV* with a drive strength range of x1 to x20.

In literature there are a few solutions for the optimal sizing ratio between consecutive inverters in an inverter chain. Linholm et al. [29] has found an optimum sizing ratio f to be 2 whereas in a lecture note from Berkeley [30] the optimal sizing ration is found to be 3.6 when including the self-loading of the inverters. In the thorough design methodology from Cherkauer et al. [31] ratios between 4 to 5 are mentioned to be the optimum. For the required inverter chains a ratio of 4 was chosen due to the stepping of the drive strength of the provided inverters in powers of two. Simulations of inverter



Fig. 3.1: Schematic of the implemented high-voltage AQC.

chains with a sizing ratio of f = 2 and f = 4 with the same load showed a similar edge rise and fall time for both. But the overall delay is shorter for the inverter chain with f = 4 because only half as many inverters are needed to drive the same load. The delay of the inverter chain for driving the gate of M_Q is crucial, to keep the overall delay from detection to onset of active quenching, low. Thus, this inverter chain is designed with a ratio of four except for the first inverter. This first inverter only would need a drive strength of x8 for a sizing ratio of f = 4. Due to the fact that the level shifter can drive a x16 inverter and the following inverters are x16 which eases the layout due to the same physical size, the given configuration was chosen. One may argue, if the first inverter is larger than required why not stay consistent with the sizing ratio. For this argument would speak to have twice the drive strength for the gate of M_Q , but simulations showed only minor improvements on the rise time. Against this argument speaks an increased area consumption and coming with it, longer signal traces in the physical layout. So an overall drive strength of x128 for the M_Q is sufficient. In the physical layout the inverter chain can be fitted very close to the quenching transistor.

For the $M_{protect}$ and M_{sense} gate drivers, the goal was that both transistors switch at the same time and before M_Q switches. It is important that the protection transistor $M_{protect}$ switches off very soon after an avalanche is detected so that the initial voltage drop due to the passive quenching phase is supported by the OFF-resistance of this high-voltage LDMOS transistor. Therefore, only two inverters are used in the driver to minimize the delay. Ideally, the sensing transistor M_{sense} should switch on simultaneously when the protection transistor switches off so that at no point in time the SPAD-cathode is connected through a low impedance path to $V_{dd,HS}$. The transistors M_{sense} and $M_{protect}$ are different types with different nominal gate voltages and gate oxide thicknesses, therefore they have dissimilar switching speed. For that reason the drive chain for the sense transistor deviates from sizing ratio of f = 4 to reduce the drive strength and to delay the switching point of M_{sense} to get it closer to $M_{protect}$'s switching point. Nevertheless, the sense transistor switches a little faster than $M_{protect}$, which causes the mentioned current spike in Fig. 2.9 at 0.4 ns.

The two complementary inputs of the level shifter are driven by two x16 inverters which are necessary to drive the gates of the high-voltage p-channel LDMOS in the level shifter circuit. The last inverter in this chain drives also the 50 Ω output buffer U_3 necessary to signalize an avalanche event to off chip equipment like a oscilloscope or counter with 50 Ω input impedance.

In addition to the principle structure discussed in Chapter 2 the current source built by M_1 and M_2 was added to the circuit to compensate for the leakage current of M_Q . If not compensated, this leakage current of about 470 pA would discharge the SPAD- capacitance and thereby decrease V_{Bias} by a voltage of two times the reference voltage in about 2.12 ms. This change of the bias voltage V_{Bias} would trigger the comparator U_1 and start the whole quenching and resetting sequence described previously. The self-induced quenching by the AQC only happens when no avalanche by means of thermal energy or photon is triggered during a period of around 2.12 ms. A self-induced quenching cannot be distinguished in the output signal from a real avalanche triggered ideally by a photon. Therefore, the mentioned compensation circuit was added to eliminate self-induced quenching operation, especially during dark count measurements. If not needed, the compensation circuit can be deactivated by connecting the terminal *Leakage Compensation* to $V_{dd,HS}$.

A 40 k Ω pull up resistor is added to the *Gate*-input of the control logic which keeps the control logic in the *QUENCH*-state if no explicit gate signal is applied. Thus, the transistor M_Q is switched on thereby protecting the VMS and reset stage, for example during the startup or other deviating operating conditions. In this state, the SPAD-detector is essentially blind for photons because the bias voltage V_{Bias} is below the breakdown voltage V_{Break} . This means no avalanches can be triggered. Thereby the detector is also protected because no sustaining avalanche can occur which could thermally destroy the detector.

It is important to note that all control inputs and the output signal are referenced to the high-side supply voltage or in other words to the potential at the terminal $V_{ss,HS}$.

3.2 Layout

3.2.1 AQC-Core Layout

The physical layout of the AQC-core of Fig. 3.1 is shown in Fig. 3.2, with major circuit blocks surrounded with a black rectangle and labeled accordingly. For this layout all 5 available metal layers (metal1 to metal5) and the final metal layer (metalF) of the PDK are used and also n- and p-channel LDMOSs for 40 V and MIM-capacitors. The high-side logic block in the deep n-well isolation is closely fitted around the protection transistor $M_{protect}$ in a L-shaped arrangement to reduce circuit parasitics. The driver circuits for the three main transistors are arranged in such a way that their gate fingers are driven from both ends. This will help to distribute the gate signal more evenly onto the 20 gate fingers and reduce the delaying influence of parasitic resistance and capacitance from the gate metal trace and gate poly-silicon. Furthermore, is the metal plane for the supply of $V_{dd,HS}$ split into two parts. One metal trace connects all the driver circuitry and another metal trace for connecting the control logic, comparator



and level shifter to the supply pin. Both traces are connected to a star point for $V_{dd,HS}$ in the vicinity of M_{sense} from where to the connection the supply pad is made.

Fig. 3.2: Layout image of the implemented high-voltage AQC-core from Fig. 3.1.

3.2.2 Custom Pads and Padring

Due to the high voltages used in this AQC-design, the provided IO-pads in the PDK can not be used right away because of the electro-static discharge (ESD)-protection circuitry integrated in these pads. Therefore, an IO-pad for analog and RF signals from the PDK was taken and customized by removing all the ESD-protection circuitry to fit the required needs. For supply pads, the trace width was increased to accommodate for the high current during the active quenching phase. To give a minimum amount of protection the custom signal pads have a 300Ω resistor between the bond pad and the

internal circuit. Special attention was given to the bond pad SPAD cathode at which the external SPAD will be connected. To reduce the parasitic capacitance contribution from the pad, only the two top metal layers are used, thereby increasing the distance to substrate which is kept at the potential of $V_{ss,LS}$, which is the lowest potential in the circuit. Furthermore, all metal connections which are part of the supply rings are removed. The supply rings are the four surrounding metal traces in Fig. 3.3, which enclose the AQC-core circuit and the bypass capacitors. Only the connection to the supply ring which carries the substrate potential $V_{ss,LS}$ is kept. The placement and position of the bond pad for the SPAD can be seen in the layout image of the full chip in Fig. 3.3 at the lower left side of the pad ring. The SPAD bond pad is separated from the remaining pads through a higher distance to reduce parasitic capacitance on the one hand and to have a higher isolation gap for the high quenching voltages. The remaining pads are arranged in such a way that the pads for the high-side logic are separated with larger distance at the lower right corner and around the SPAD pad, SPAD cathode, on the left side, from the pads for the low-side logic block. This makes the design of the PCB, on which the chip will be mounted for testing, easier in terms of isolation gap between the traces for the different voltage domains. This separation is marked via the orange dashed line in the chip layout image in Fig. 3.3.

The core circuit is placed in the lower left corner inside the pad ring to keep the connection between the bond pad for the external SPAD and the quenching transistor M_Q as short as possible. The remaining area, which is limited by the number of required pads, is filled with bypass capacitors for the high-side supply voltage of 1.8 V and the low-side supply voltage of 3.3 V.

The testbench for evaluating the performance of the developed high-voltage AQCchip is shown in Fig. 3.4. This testbench reproduces the real environment as close as possible, in particular the PCB, in which the fabricated chip will be tested and evaluated. Therefore, also bypass capacitors are placed on positions where they will be needed despite the fact they have no influence in the simulation when ideal voltage sources are used. The default value for all bypass capacitors is 100 nF unless there is a different value mentioned in Fig. 3.4. The inductors in Fig. 3.4 model the behavior of the bond wires. For the dimensions of a bond wire, a diameter of 25.4 µm and a length of 2 mm are assumed, with these dimensions their inductance is 2 nH and their series resistance is 100 m Ω , according to [32].

The quenching voltage will be applied between $V_{dd,HS}$ and $V_{ss,LS}$ as already mentioned. In Chapter 4 will be shown that the current into the *SPAD cathode*-pad peaks at around 300 mA during quenching. This high current pulse must also flow through the $V_{ss,LS}$ -connection back to the anode of the SPAD. Therefore, the $V_{ss,LS}$ -connection is realized



Fig. 3.3: Layout of the overall high-voltage AQC-chip with the padring and custom designed pads.



Fig. 3.4: Schematic of the testbench circuit for the high-voltage AQC to evaluate the design.

via three bond pads, as well as the connection to $V_{dd,HS}$, as can be seen in the layout image in Fig. 3.3. With three bond pads and three bond wires to connect the substrate terminal $V_{ss,LS}$ to the PCB, the resulting parasitic resistance can be reduced to one third of one bond wire. More important, the remaining inductance can be reduced to one third of a single bond wire when neglecting the mutual inductance. Taking into account the mutual inductance the remaining inductance of three bond wires will be between one half and two third of one bond wire. This is especially important during the short high current pulse while quenching, since simulation showed that increasing the number of bond wires from one to two reduces ringing notably. Shorter bond wires will even further reduce ringing and oscillations caused by the bond wires. Due to similar considerations two bond pads are used for the remaining logic supply voltage terminals. The same consideration would also hold for the SPAD cathode terminal but on the other hand more bond pads would increase the parasitic capacitance at this sensitive node, therefore only one bond pad is used. The resistor R_{limit} in Fig. 3.4 is a safety measure to limit the current during startup when C_{SPAD} is charged up to its ready state bias voltage or in case the developed AQC malfunctions. To protect the SPAD, the value of R_{limit} should be at least $\frac{V_{EX}}{I_{max}}$ with I_{max} as the maximum allowed avalanche current for the particular used SPAD.

Chapter 4

Results and Comparison

The typical electrical characteristics of the designed AQC from Chapter 3 are summarized in Table 4.1. The control input voltages for *Reference*, *Comparator Bias* and *Leakage Compensation* are all referenced to $V_{ss,HS}$. This means with a *Reference*-voltage of 1.7 V the threshold until the comparator detects an avalanche is 100 mV. If not otherwise mentioned the typical parameters are used for the post-layout simulation in the following Section 4.1.

parameter	typical	range
quenching voltage V_q	35 V	$5\mathrm{V}35\mathrm{V}$
SPAD-capacitance C_{SPAD}	5.0 pF	$0.1\mathrm{pF}{-}5.0\mathrm{pF}$
SPAD intrinsic resistance R_{SPAD}	800 Ω	$\geq 500\Omega$
hold-off time	$5\mathrm{ns}$	$5.0\mathrm{ns}160\mathrm{ns}$
Reference voltage	1.7 V	$0.0\mathrm{V} ext{}1.8\mathrm{V}$
Comparator Bias voltage	1.4 V	$0.0 \mathrm{V}{-}1.8 \mathrm{V}$
Leakage Compensation voltage	1.0 V	$0.0{ m V}{-}1.8{ m V}$
Output load	$ 50\Omega 5.0\mathrm{pF}$	

Tab. 4.1: Electrical characteristic of the designed AQC.

4.1 Post-Layout Simulation Results

4.1.1 Exemplary Results for a Quenching Voltage of 35 V

Post-layout simulation results for a quenching voltage V_q of 35 V and the shortest possible hold-off time are shown in Fig. 4.1. The top graph shows the voltage waveform at the *SPAD cathode* terminal in blue. Additionally, the different phases during quenching are annotated. The encircled part denoted with M_{sense} to $M_{protect}$ switch over marks the short voltage spike which happens when the switch states of M_{sense} and $M_{protect}$ change as discussed in Section 2.3.2. This voltage spike, which is more clearly visible in Fig. 4.8, arises in the short time when both transistor are conducting during the switch over due to not perfectly synchronized gate signals and different switching delays. The output signal, which synchronously indicates the detection of an avalanche to an off-chip circuitry, is shown as orange waveform in the top graph in Fig. 4.1. The length of the output pulse is equal to the time the quenching transistor M_Q is switched on and therefore is equal to the hold-off time. The lower graph shows the corresponding



Fig. 4.1: SPAD-cathode voltage waveform and corresponding *Out*-signal (top), and SPAD-current into cathode (bottom).

current flowing into the SPAD cathode during quenching. The current waveform is a composition of the ideal current flow and superimposed high frequency oscillations arising from the inductance of the bond wires. In the bottom graph of Fig. 4.1, it can be seen that during passive quenching, current is flowing into the SPAD. When the quenching transistor turns on it rapidly discharges the SPAD-capacitance C_{SPAD} as can be seen on the large negative current spike and thereby lowering V_{Bias} below the breakdown voltage and finally stopping the avalanche current. This rather large current spike, which must flow also through the $V_{ss,LS}$ -pads and through a closely placed bypass capacitor into the SPAD-anode is the reason for the three supply pads at the $V_{ss,LS}$ and $V_{dd,HS}$ terminals. The parasitic inductance of the bond wire and the voltage drop, caused by the current spike, can be reduced with this effort. At the $V_{ss,LS}$ and $V_{dd,HS}$ terminals the additional parasitic capacitance of the bond pads does not degrade the performance of the circuit.



Fig. 4.2: Power consumption of the AQC-chip during one quenching and reset phase.

The corresponding transient power consumption of the quenching and resetting sequence of Fig. 4.1 is shown in Fig. 4.2. The power consumption prior to t = 0.0 ns is the quiescent power of about 100 mW for the circuit in its *READY*-state as discussed in Section 2.6 and 2.9. The average power during the quenching and resetting sequence from 0.0 ns to 16.5 ns is about 606 mW. Taking the thermal resistance result from Section 2.9 of $R_{th} = 33.8$ K/W for bulk silicon the temperature rise will be 20.5 K when continuously quenching with the highest possible count rate. This means for consecutive quenching sequences following immediately after each other the steady power consumption will settle to 606 mW and cause a temperature rise of about 20.5 K.

The performance of this AQC at different process corners is shown in Fig. 4.3. The passive and active quenching phase vary only minor from PVT-variations due to the fast reaction of the *CONTROL LOGIC* onto an avalanche. On the contrary, the resetting phase is more influenced by PVT-variations, in part due to PVT-dependent changes

of the delay elements in the CONTROL LOGIC. Other parts which are influenced by PVT-variations are the cascoded transistors $M_{protect}$ and M_{sense} . In the corner when the parameters for p-channel transistors are in the slow regime, both transistors $M_{protect}$ and M_{sense} are similarly affected. In this case, their reduced drain current density slows down the reset phase, since the recharge current must flow through both p-channel transistors, and the cascode connection will to some extent intensify the slow down even more.



Fig. 4.3: SPAD-cathode voltage at different process corners for a temperature of 0 °C (top) and 80 °C (bottom). The typical waveform in blue in both diagrams was simulated with a temperature of 30 °C.

The behavior of this AQC-design at different hold-off times is shown in Fig. 4.4, which is achieved by increasing the resistance of $R_{holdoff}$ in Fig. 3.4. The minimal hold-off time is limited by the internal resistor R_2 and the MOS-capacitance M_4 in the circuit of Fig. 2.16 when $R_{holdoff}$ is set to 0Ω . Post-layout simulation give a minimal hold-off time of 5 ns due to additional parasitic capacitance and resistance. The longest hold-off time is limited to 160 ns by the internal resistor R_1 in Fig. 2.16, when using an external resistor for the adjustment. There is however the option to adjust the hold-off time via an external voltage source at pin *Hold-off time* in Fig. 3.1. This allows to further increase the hold-off time, but it should be noticed that for even longer hold-off times the internal current which charges the MOS-capacitance M_4 is very small and the superimposed noise could make the resulting hold-off time inaccurate.



Fig. 4.4: SPAD-cathode voltage at different hold-off times by adjusting the resistor $R_{holdoff}$ accordingly.

4.1.2 Capability of Using different SPAD-Types

This AQC was initially designed for an external thick SPAD such as SAP500 from Laser Components [16]. Therefore, the simulation parameters of Table 2.1 were used for the SPAD-model of Section 2.2 to design the quenching circuit. These simulation parameters are derived from existing characterizations of this SPAD [20], [33] and its data sheet [16]. As already mentioned in Section 2.2, the SPAD capacitance C_{SPAD} was chosen to be larger than given in the data sheet to account for additional parasitic capacitance, which will arise from the physical layout of the circuit. This section now evaluates the ability of the designed circuit to successfully quench the targeted SPAD with its nominal capacitance and other SPADs. In the simulation the characteristic of a SPAD is captured in two parameters, which are its capacitance C_{SPAD} and its intrinsic resistance R_{SPAD} as discussed in Section 2.2. The capacitance C_{SPAD} mainly depends on the size and structure of the SPAD, for example the SAP500 has a nominal capacitance of 3.3 pF. Whereas the intrinsic resistance R_{SPAD} depends mainly on doping concentration of the involved silicon layers and the layer stack up. Its value for thick SPADs is around 500 Ω and for thin SPAD around a few kiloohms according to [19].

The main reason that limits the range of usable SPADs is the voltage at the V_{sense} node in the detection circuit of Fig. 2.8. Due to its connection via the protection transistor $M_{protect}$ to the SPAD-cathode, the voltage at the V_{sense} -node could potentially get too low and destroy M_{sense} by gate oxide breakdown. The V_{sense} -node voltage must not go below $V_{dd,HS} - 2.2$ V, which is the technological limit for the sensing transistor M_{sense} . Two typical voltage waveforms at the V_{sense} -node for different SPADs, which are represented via their capacitance C_{SPAD} and intrinsic resistance R_{SPAD} , are shown in Fig. 4.5. It can be seen that the voltage at V_{sense} exhibits a high negative voltage



Fig. 4.5: V_{sense} -node voltage during an avalanche and subsequent quenching phase for two different SPADs and the voltage limit for M_{sense} marked in magenta.

peak immediately after an avalanche occurs, due to the fast rising avalanche current. This high voltage peak is used to detect the avalanche via the comparator. The blue waveform shows the V_{sense} voltage for the default SPAD from Section 2.2 used to design the circuit. Whereas the orange V_{sense} -voltage waveform reaches the technological limit for a SPAD with a lower capacitance and intrinsic resistance. The capacitance C_{SPAD}


Fig. 4.6: Safe area for the usage of different SPADs characterized by their capacitance C_{SPAD} and intrinsic resistance R_{SPAD} at a quenching voltage of 35 V.

of 3 pF used to get the orange waveform is below but closer to the nominal capacitance of the photon detector SAP500 [16]. The area in the R_{SPAD} - C_{SPAD} space for pairs of R_{SPAD} - C_{SPAD} -values for which the V_{sense} -voltage stays within the limit is depicted as green shaded area in Fig. 4.6 and labeled *safe area*. This *safe area* was evaluated for a quenching voltage of 35 V. The orange dots in Fig. 4.6 mark simulation results for SPADs where the voltage peak at the V_{sense} -node reaches the technological limit of 2.2 V at the drain to source voltage of M_{sense} . Similar to the orange waveform in Fig. 4.5 only with a different combination of C_{SPAD} and R_{SPAD} representing a different SPAD. The blue crosses mark a selection of SPADs, characterized via their pair of C_{SPAD} and R_{SPAD} , which are within the *safe area*. The SPADs marked with blue crosses and orange dots were used to evaluate the *safe area*.

Although this simulated limitation of the range of usable SPADs may be too stringent for two reasons. First of all the duration of the voltage peak at node V_{sense} is limited to approximately 1 ns due to the active protection of M_{sense} from high quenching voltages accomplished by the VMS and reset stage and the CONTROL LOGIC. The delay until the CONTROL LOGIC detects an avalanche and interferes to protect M_{sense} depends to some extent on the actually used chip sample and the actual PVT deviation. Furthermore, M_{sense} is passively protected in the DC-regime since it is connected in series with $M_{protect}$. In the worst case, when no interfering through the CONTROL LOGIC happens and the gate signals stays unchanged as in the READY-state. The increasing voltage across source to drain of M_{sense} would reduce the source to gate voltage of $M_{protect}$, thereby switching $M_{protect}$ off and limiting the maximum DC-voltage across M_{sense} to the supply voltage of 1.8 V. Thus, only during a rapidly changing voltage at the SPAD cathode terminal a higher voltage peak due to capacitive coupling can occur.

Second, the worst case over-voltage peak that can occur is 5.5 V below the lower limit, with SPAD-parameters $R_{SPAD} = 100 \Omega$ and $C_{SPAD} = 100 \text{ pF}$. Studies of CMOS-circuits with similar DC over voltages in the range of a few volts give a cumulative probability of 0.7% for a gate oxide breakdown below a time to breakdown t_{BD} of 10s [34], [35]. This means the probability for a gate oxide breakdown in less than 10 s after applying the over voltage is 0.7 %. This probability agrees with the reliability data given in the PDK user guide for the LF15A-process. It also is worth mentioning that the safe area in Fig. 4.6 was evaluated with a quenching voltage of 35 V. This means lowering the quenching voltage V_q will also increase the safe area towards lower R_{SPAD} and C_{SPAD} values. Furthermore, a real SPAD is only to a certain extent realistically modeled with the used SPAD-model in Section 2.2. Due to some approximations in the SPAD-model the current at the trigger point of an avalanche changes abruptly, which will not happen as abruptly in reality. Therefore, it can be concluded that the designed AQC will work even if in some of the cases where the voltage across M_{sense} at the onset of an avalanche exceeds the technological limit. In case of exceeding the technological limit, these recurring voltage peaks at onset of an avalanche may reduce the projected lifetime of 10 years for the used technology significantly. These peaks will also increase the failure rate of the designed circuit which is acceptable since the purpose of the circuit is to be used for scientific research rather than as commercial product.

The ability to use this designed AQC with different quenching voltages V_q is demonstrated in Fig. 4.7. For this simulation the default SPAD-parameters as mentioned in Table 2.1 were used. The hold-off time and reset time are set to an arbitrary time via a 10 k Ω resistor for $R_{holdoff}$ and R_{reset} . With this setup of the testbench of Fig. 3.4 the results in Fig. 4.7 can be achieved by changing the $V_{dd,HS}$ -potential in 5V-steps. It can be seen that for a quenching voltage of $V_q = 35$ V to $V_q = 10$ V the circuit exhibits a nearly identical quenching time of about 2.2 ns. For 5V quenching voltage



Fig. 4.7: Quenching pulses at the SPAD-cathode for different quenching voltages V_q .

the quenching time is a bit longer due to the shallower voltage slope during the passive quenching phase. Thus delaying the detection of the avalanche by the comparator. This goes along with the previous statement that lower quenching voltages reduce also the voltage peak at the V_{sense} -node at onset of an avalanche and therefore allows to use a wider range of different SPADs.

4.2 Comparison to State of the Art

Finally, in this section the characteristic parameters for the State of the Art AQCs presented in Section 1.4 are compared with this design. An overview of their characteristic parameters is listed in Table 4.2. It should be noted that all the given parameters are for the combination of the AQC and the used SPAD. This means only comparing the AQCs should be treated with some caution due to the fact they face different conditions on the input terminal where the SPAD is connected. A second point which should be considered in this comparison is the fact that for this AQC-design at the moment of writing this thesis only simulation results exist. In contrary, the values for the other designs in Table 4.2 are taken from experimental measurements.

For evaluating the maximum photon count rate of this design a lower quenching voltage of $V_q = 10$ V was selected. At this operating condition the circuit has the same

Tab. 4.2: Comparison of characteristic AQC-parameters (second parameter gives the used quenching voltage when it deviates from the maximum quenching voltage given in the second column.)

design	max. quench- ing voltage / V	count rate / Mcounts/s	quenching time / ns	reset time / ns
this design	35	90 @ 10 V	2.21	10.3
Fig. 1.8 [3]	50	80 @ 5V	60 @ 30 V	40 @ 30 V
Fig. 1.9 [6]	9.9	129	2.2	2.4
Fig. 1.10 [10]	20	20	25	20



Fig. 4.8: SPAD-cathode voltage for a SPAD with $R_{SPAD} = 800 \,\Omega$ and $C_{SPAD} = 5 \,\mathrm{pF}$ for evaluation of the maximum count rate at a quenching voltage of $V_q = 10 \,\mathrm{V}$

quenching time as for higher voltages but the shortest dead time of 11 ns due to a faster resetting time, which gives a maximum count rate of 90 Mcounts/s. The corresponding SPAD cathode voltage waveform is depicted in Fig. 4.8. The dead time at a quenching voltage of 35 V is around 17.8 ns, which gives a maximum photon count rate of about 56.2 Mcounts/s. This quencher design is based strongly on the design of Acconcia's AQC [3], as already mentioned. For comparison the behavior of Acconcia's AQC at high quenching voltages during quenching is displayed in Fig. 4.9. For Acconcia's AQC-design it is also worth to notice that active quenching of an avalanche starts 2.3 ns



Fig. 4.9: Waveform of the anode voltage for a RE-SPAD at two different quenching voltages [3, Fig. 7].

after its detection due to internal delays in its logic circuit, which controls the quenching and resetting sequence. Whereas this design fully quenches an avalanche in 2.21 ns according to simulation. This should have a positive impact on the APP because of less avalanche charge passing through the space charge region during an avalanche. The achieved quenching time is similar to the results of Dervić's design [6] except that this design achieves it at a quenching voltage of 35 V. Further, is the quenching voltage in this design adjustable compared to Dervić's design [6] where the quenching voltage is fixed due to the cascoded switch structure.

Overall, this AQC-design accomplishes similar performance as the low voltage AQCdesign from Dervić et al. in terms of quenching time and photon count rate. Whereas the resetting is not as fast due to the compromise in the sizing of the quenching and protection transistor in favor of a short quenching time.



Chapter 5

Conclusion and Outlook

In this thesis a high voltage AQC was designed in the LF15A process from the manufacturer *LFoundry* by utilizing the included high voltage LDMOS transistors. These LDMOS transistors give the AQC the possibility to apply quenching voltages up to 35 V onto an external SPAD. According to recent works such high quenching voltages should allow to further increase the photon-detection probability compared to AQCs with 3.3 V or 9.9 V quenching voltages, as mentioned in Section 1.2. Therefore, this designed AQC in combination with an appropriate SPAD should be able to increase the PDP up to 90 %.

The designed AQC shows similar time performance for quenching an avalanche in post layout simulations as a recent low voltage quencher design, with a quenching time of 2.2 ns. This performance is achieved with the use of a rather unique method for detecting an avalanche which uses two MOSFETs in dual purpose for sensing an avalanche and recharging the SPAD. The short quenching time minimizes the avalanche charge and thereby helps to decrease the afterpulsing probability (APP) which comes at the cost of a rather high quiescent power consumption of the chip. This high power consumption is mainly caused by the necessary level shifter described in Section 2.6. Therefore, a future improvement of the circuit could be the adaption of a level shifter with no static power consumption, like the design in [26], to achieve a similar response time as the used level shifter.

In retrospect, the suitability of this AQC-design to work with a variety of SPADs was a bit overlooked during the design phase. As discussed in Section 4.1.2 the range of usable SPADs at a quenching voltage of 35 V has a lower limit in terms of the combination of its SPAD-capacitance and its intrinsic resistance, which is shown in Fig. 4.6. The characterization of the lower limit in Section 4.1.2 may be too stringent as already mentioned due to a certain tolerance of the CMOS-circuit against short overvoltage spikes and a less abrupt current rise in reality as modeled. A first possible future improvement would be to adjust the size of the sensing transistor to better

cope with SPADs with lower intrinsic resistance. This approach would lower the avalanche detection sensitivity for SPADs with high intrinsic resistance. Another future improvement could be to cascade a few sensing MOSFETs, for example two or three, with a selectable number of MOSFETs which take part in the sensing, thereby changing the sense resistance of the *SENSE STAGE*. By this approach the circuit is adjustable for SPADs with high and low intrinsic resistance. Another advanced improvement of the *SENSE STAGE* could be the implementation of current-mode sensing via the usage of a customized current comparator or transimpedance amplifier. Their low input impedance should help to guide most of the avalanche current at its onset into the *SENSE STAGE* for faster detection. Thereby circumventing the current distribution issue mentioned in Section 2.3.1. Also, it was confirmed via PVT-simulations that all transistors operate inside their technological limits when the circuit is connected to a SPAD with electrical characteristic like in Section 2.2. The change of the *safe area* in Fig. 4.6 due to PVT-variations should be evaluated prior to production of the chip.

Overall post-layout simulation results for this high voltage AQC-design show very promising performance compared to State of the Art high voltage AQC-designs. Thereby fulfilling the goal of this thesis, improving the performance of the inspiring design from Acconcia, Rech, Gulinatti, *et al.* [3].

Bibliography

- [1] A. Gulinatti, I. Rech, F. Panzeri, C. Cammi, P. Maccagnani, M. Ghioni, and S. Cova, "New silicon SPAD technology for enhanced red-sensitivity, high-resolution timing and system integration", *Journal of Modern Optics*, vol. 59, no. 17, pp. 1489–1499, Oct. 10, 2012, ISSN: 0950-0340. DOI: 10.1080/09500340.2012.701340. [Online]. Available: https://doi.org/10.1080/09500340.2012.701340 (visited on 11/22/2018).
- [2] A. Gallivanoni, I. Rech, and M. Ghioni, "Progress in Quenching Circuits for Single Photon Avalanche Diodes", *IEEE Transactions on Nuclear Science*, vol. 57, no. 6, pp. 3815–3826, Dec. 2010, ISSN: 0018-9499. DOI: 10.1109/TNS.2010.2074213.
- G. Acconcia, I. Rech, A. Gulinatti, and M. Ghioni, "High-voltage integrated active quenching circuit for single photon count rate up to 80 Mcounts/s", *Optics Express*, vol. 24, no. 16, pp. 17819–17831, Jul. 27, 2016, ISSN: 1094-4087. DOI: 10.1364/0E.24.017819. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5946911/ (visited on 10/05/2018).
- [4] R. H. Hadfield, "Single-photon detectors for optical quantum information applications", *Nature Photonics*, vol. 3, no. 12, pp. 696-705, Dec. 2009, ISSN: 1749-4893. DOI: 10.1038/nphoton.2009.230. [Online]. Available: https://www.nature.com/articles/nphoton.2009.230 (visited on 10/30/2019).
- [5] F. Zappa, S. Tisa, A. Tosi, and S. Cova, "Principles and features of single-photon avalanche diode arrays", Sensors and Actuators A: Physical, vol. 140, no. 1, pp. 103-112, Oct. 1, 2007, ISSN: 0924-4247. DOI: 10.1016/j.sna.2007.06.021.
 [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0924424707004967 (visited on 09/25/2019).
- [6] A. Dervić, B. Steindl, M. Hofbauer, and H. Zimmermann, "High-voltage active quenching and resetting circuit for SPADs in 0.35 µm CMOS for raising the photon detection probability", *Optical Engineering*, vol. 58, no. 4, p. 040 501, Apr. 2019, ISSN: 0091-3286, 1560-2303. DOI: 10.1117/1.0E.58.4.040501. [Online]. Available: https://www.spiedigitallibrary.org/journals/Optical-Engineering/ volume-58/issue-4/040501/High-voltage-active-quenching-and-resetting-

circuit - for - SPADs - in / 10.1117 / 1.0E.58.4.040501.short (visited on 04/30/2019).

- [7] M. Ghioni, S. Cova, F. Zappa, and C. Samori, "Compact active quenching circuit for fast photon counting with avalanche photodiodes", *Review of Scientific Instruments*, vol. 67, no. 10, pp. 3440–3448, Oct. 1996, ISSN: 0034-6748, 1089-7623. DOI: 10.1063/1.1147156. [Online]. Available: http://aip.scitation.org/doi/10.1063/1.1147156 (visited on 09/27/2019).
- [8] G. A. Reider, *Photonik: Eine Einführung in die Grundlagen*, 3rd ed. Wien: Springer-Verlag, 2012, ISBN: 978-3-7091-1520-6. [Online]. Available: https:// www.springer.com/de/book/9783709115206 (visited on 09/25/2019).
- H. Zimmermann, *Integrated silicon optoelectronics*, 2. ed, ser. Springer series in optical sciences 148. Heidelberg: Springer, 2010, 385 pp., OCLC: 845539542, ISBN: 978-3-642-01520-5.
- [10] F. Zappa, A. Lotito, A. C. Giudice, S. Cova, and M. Ghioni, "Monolithic activequenching and active-reset circuit for single-photon avalanche detectors", *IEEE Journal of Solid-State Circuits*, vol. 38, no. 7, pp. 1298–1301, Jul. 2003, ISSN: 0018-9200. DOI: 10.1109/JSSC.2003.813291.
- [11] H. Zimmermann, B. Steindl, M. Hofbauer, and R. Enne, "Integrated fiber optical receiver reducing the gap to the quantum limit", *Scientific Reports*, vol. 7, no. 1, p. 2652, Jun. 1, 2017, ISSN: 2045-2322. DOI: 10.1038/s41598-017-02870-2. [Online]. Available: https://www.nature.com/articles/s41598-017-02870-2 (visited on 04/30/2019).
- R. Enne, B. Steindl, M. Hofbauer, and H. Zimmermann, "Fast Cascoded Quenching Circuit for Decreasing Afterpulsing Effects in 0.35-µm CMOS", *IEEE Solid-State Circuits Letters*, vol. 1, no. 3, pp. 62–65, Mar. 2018, ISSN: 2573-9603. DOI: 10.1109/LSSC.2018.2827881.
- [13] T. Ohguro, T. Iizuka, M. Taguchi, S. Kumashiro, S. Miyamoto, R. Inagaki, Y. Furui, and N. Fudanuki, "HiSIM HV 1.0.2, version 1.02 user's manual", p. 89, 2008.
- M. R. Duncan, J. M. Robertson, R. J. Holwill, and R. Rodrigues, "CMOScompatible High-voltage Complementary LDMOS Devices", in *ESSDERC '89:* 19th European Solid State Device Research Conference, Sep. 1989, pp. 535–538. DOI: 10.1007/978-3-642-52314-4_110.

- S. C. Sun and J. D. Plummer, "Modeling of the on-resistance of LDMOS, VDMOS, and VMOS power transistors", *IEEE Transactions on Electron Devices*, vol. 27, no. 2, pp. 356–367, Feb. 1980. DOI: 10.1109/T-ED.1980.19868.
- [16] (). SAP500: Silicon APDs for Photon Counting Applications, [Online]. Available: https://www.lasercomponents.com/us/news/sap500-silicon-apds-forphoton-counting-applications-1/ (visited on 10/31/2019).
- [17] (). DTS_C30902_C30921.pdf, [Online]. Available: http://www.excelitas.com/ downloads/DTS_C30902_C30921.pdf (visited on 02/13/2019).
- [18] A. D. Mora, A. Tosi, S. Tisa, and F. Zappa, "Single-Photon Avalanche Diode Model for Circuit Simulations", *IEEE Photonics Technology Letters*, vol. 19, no. 23, pp. 1922–1924, Dec. 2007, ISSN: 1041-1135. DOI: 10.1109/LPT.2007.908768.
- [19] S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa, "Avalanche photodiodes and quenching circuits for single-photon detection", *Applied Optics*, vol. 35, no. 12, pp. 1956–1976, Apr. 20, 1996, ISSN: 2155-3165. DOI: 10.1364/A0.35.001956.
 [Online]. Available: https://www.osapublishing.org/ao/abstract.cfm?uri=ao-35-12-1956 (visited on 11/22/2018).
- [20] M. Stipčević, H. Skenderović, and D. Gracin, "Characterization of a novel avalanche photodiode for single photon detection in VIS-NIR range", Optics Express, vol. 18, no. 16, pp. 17448–17459, Aug. 2, 2010, ISSN: 1094-4087. DOI: 10.1364/0E.18.017448. [Online]. Available: https://www.osapublishing.org/oe/abstract.cfm?uri=oe-18-16-17448 (visited on 10/21/2019).
- [21] W. M. C. Sansen, Analog design essentials, ser. SECS. Dordrecht: Springer, 2006, 777 pp., ISBN: 978-0-387-25746-4.
- [22] D. Lutz, A. Seidel, and B. Wicht, "A 50V, 1.45ns, 4.1pJ High-Speed Low-Power Level Shifter for High-Voltage DCDC Converters", in ESSCIRC 2018 - IEEE 44th European Solid State Circuits Conference (ESSCIRC), Sep. 2018, pp. 126–129. DOI: 10.1109/ESSCIRC.2018.8494292.
- [23] M. J. Declerq, M. Schubert, and F. Clement, "5 V-to-75 V CMOS output interface circuits", in 1993 IEEE International Solid-State Circuits Conference Digest of Technical Papers, Feb. 1993, pp. 162–163. DOI: 10.1109/ISSCC.1993.280014.
- [24] D. Liu, S. J. Hollis, H. C. P. Dymond, N. McNeill, and B. H. Stark, "Design of 370-ps Delay Floating-Voltage Level Shifters With 30-V/ns Power Supply Slew Tolerance", *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 63, no. 7, pp. 688–692, Jul. 2016, ISSN: 1549-7747. DOI: 10.1109/TCSII.2016. 2530902.

- [25] Y. Moghe, T. Lehmann, and T. Piessens, "Nanosecond Delay Floating High Voltage Level Shifters in a 0.35µm HV-CMOS Technology", *IEEE Journal of Solid-State Circuits*, vol. 46, no. 2, pp. 485–497, Feb. 2011, ISSN: 0018-9200. DOI: 10.1109/JSSC.2010.2091322.
- [26] X. Lai, L. Zhong, D. Xu, H. Wang, B. Yuan, Q. Li, R. Ding, and J. Zhao, "A novel low delay high-voltage level shifter with transient performance insensitive to parasitic capacitance and transfer voltage level", *Circuits, Systems, and Signal Processing*, vol. 36, no. 9, pp. 3598–3615, Sep. 2017, ISSN: 0278-081X, 1531-5878. DOI: 10.1007/s00034-016-0488-z. [Online]. Available: http://link.springer.com/10.1007/s00034-016-0488-z (visited on 12/14/2018).
- [27] C. Baltacı and Y. Leblebici, "Thermal aware design and comparative analysis of a high performance 64-bit adder in FD-SOI and bulk CMOS technologies", *Integration*, vol. 58, pp. 421-429, Jun. 2017, ISSN: 01679260. DOI: 10.1016/j.vlsi.2017.03.001. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S016792601730144X (visited on 06/12/2019).
- [28] S. Kemmetmüller, "Modellbildung SS 2019", p. 272, 2019.
- [29] L. W. Linholm, "An optimized output stage for MOS integrated circuits", *IEEE Journal of Solid-State Circuits*, vol. 10, no. 2, pp. 106–109, Apr. 1975, ISSN: 0018-9200. DOI: 10.1109/JSSC.1975.1050569.
- [30] (). Lecture7-invsize.pdf, [Online]. Available: http://bwrcs.eecs.berkeley.edu/ Classes/icdesign/ee141_s02/Lectures/Lecture7-invsize.PDF (visited on 03/01/2019).
- B. S. Cherkauer and E. G. Friedman, "A unified design methodology for CMOS tapered buffers", *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 3, no. 1, pp. 99–111, Mar. 1995, ISSN: 1063-8210. DOI: 10.1109/92. 365457.
- [32] (). MOSIS Packaging Options, [Online]. Available: https://www.mosis.com/ pages/products/assembly/index#wire (visited on 05/28/2019).
- [33] S. Gaspar, "Diplomarbeit: Vergleich von Quencherschaltungen und SPADs", PhD thesis, TU Wien, Wien, Dec. 31, 2016. DOI: 10.7767/9783205202240-005. (visited on 12/11/2018).
- [34] B. Kaczer, R. Degraeve, M. Rasras, K. Van de Mieroop, P. Roussel, and G. Groeseneken, "Impact of MOSFET gate oxide breakdown on digital circuit operation and reliability", *IEEE Transactions on Electron Devices*, vol. 49, no. 3, pp. 500– 506, Mar. 2002, ISSN: 1557-9646. DOI: 10.1109/16.987122.

B. E. Weir, M. A. Alam, J. D. Bude, P. J. Silverman, A. Ghetti, F. Baumann,
P. Diodato, D. Monroe, T. Sorsch, G. L. Timp, Y. Ma, M. M. Brown, A. Hamad,
D. Hwang, and P. Mason, "Gate oxide reliability projection to the sub-2 nm regime", *Semiconductor Science and Technology*, vol. 15, no. 5, pp. 455-461,
May 1, 2000, ISSN: 0268-1242, 1361-6641. DOI: 10.1088/0268-1242/15/5/304.
[Online]. Available: http://stacks.iop.org/0268-1242/15/i=5/a=304?key=
crossref.9c258254563a7bfef11591fcb66545b4 (visited on 11/22/2019).



Code of Conduct

Hiermit erkläre ich, dass die vorliegende Arbeit gemäß dem Code of Conduct – Regeln zur Sicherung guter wissenschaftlicher Praxis (in der aktuellen Fassung des jeweiligen Mitteilungsblattes der TU Wien), insbesondere ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel, angefertigt wurde. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Die Arbeit wurde bisher weder im In– noch im Ausland in gleicher oder in ähnlicher Form in anderen Prüfungsverfahren vorgelegt.

Martin Jungwirth Wien, am 13. Januar 2020