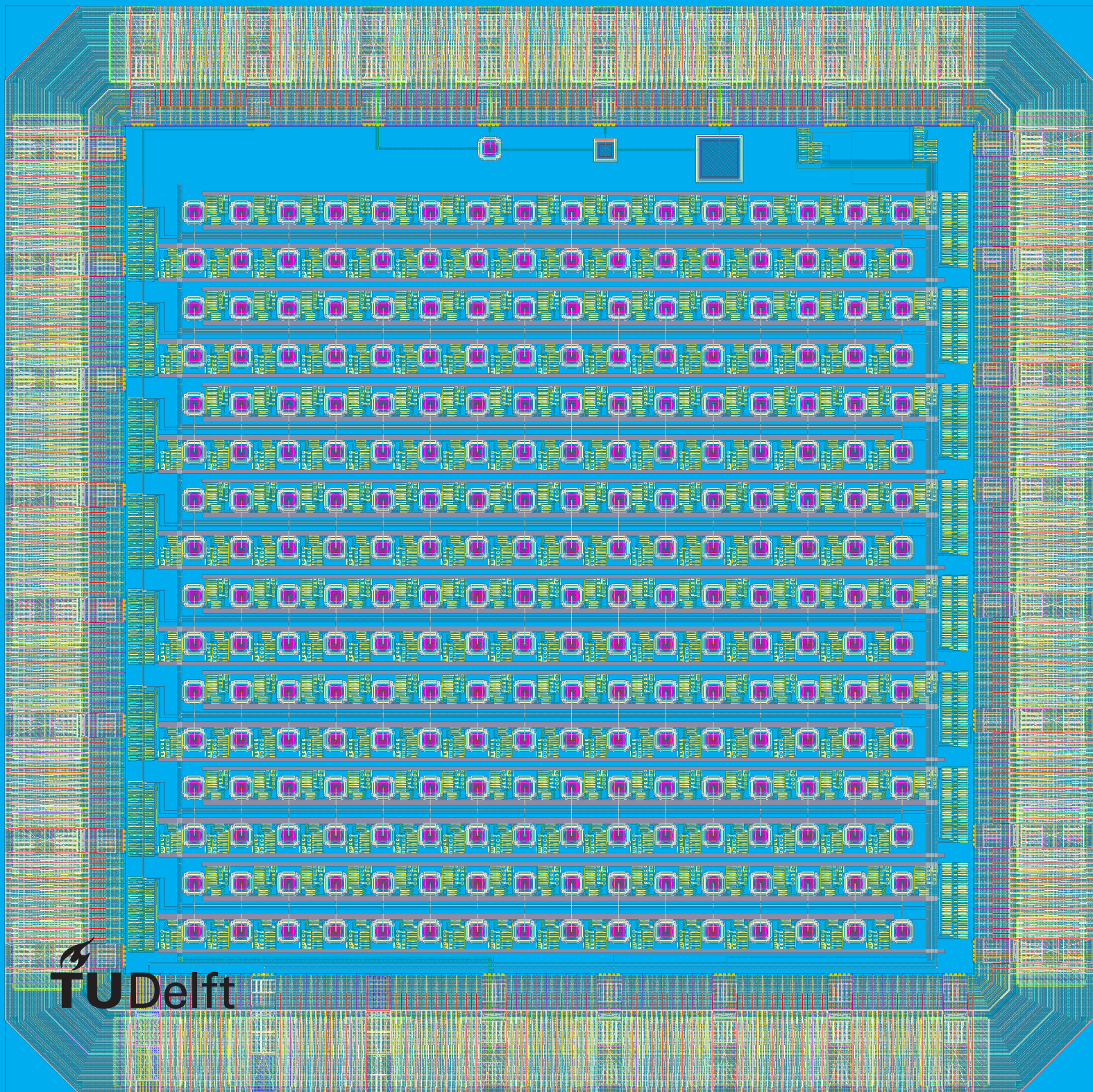
On-chip SPAD array with IO designed for quantum sensing

Samantha van Rijs
Lars Visser
Dagmar Westenbrink

July 6, 2023



TU Delft

On-chip SPAD array with IO designed for quantum sensing

by

Samantha van Rijs
Lars Visser
Dagmar Westenbrink

to obtain the degree of Bachelor of Science
at the Delft University of Technology,
to be defended on Tuesday June 28, 2023 at 15:00 PM.

Student number:	Samantha van Rijs	5077028
	Lars Visser	4359992
	Dagmar Westenbrink	5379954
Project duration:	April 24, 2023 – June 28, 2023	
Thesis committee:	Msc. I. Varveris,	TU Delft, supervisor
	Dr. R. Ishihara,	TU Delft, supervisor
	Prof.Dr.Ir. A.J. van der Veen,	TU Delft
	Dr. O.A. Krasnov,	TU Delft

TU Delft

Abstract

In this Bachelor graduation project, a 16x16 Single Photon Avalanche Diode (SPAD) array is designed in 40nm TSMC CMOS for diamond single Nitrogen Vacancy center array readout. It includes an active quenching and recharge circuit (AQC), a hold-off circuit for controllable dead-time and IO electronics for off-chip communication. The chip is part of a proposed high sensitivity magnetometer based on single NV center readout with a focus on detecting cancer in biological samples.
From the Quantum Integration Technology (QIT) lab, pre-designed SPADs were received to be implemented into the design together with SPAD quenching, recharge and input-output controller electronics. A SPAD model is adapted from literature for simulations of the electric behaviour of the electronics. We present a design and implementation of the Active Quenching Circuit (AQC), a design and implementation of the recharge and hold-off circuits, Input-Output (IO) interface design and implementation and the final top-level implementation ready for the next tape out.
Post-layout simulations show negligible speed slowdown and distortion. The AQC has a 12ns quenching time and a 1ns recharge time, leading to a theoretical maximum count rate of 76Mc/s. The hold-off circuit has a tunable dead-time for afterpulsing reduction and 16, 16 bit parallel-in-serial-out (PISO) modules allow per-row readout of the array.The electronics are co-located with the SPADs and pixel pitch is 24μm. The final chip design meets the single NV fluorescence count rate requirement of above 3 Mc/s, the 1.1x1.1 mm^2 area requirement, the 32-pin IO requirement, the 16x16 SPAD pixel requirement and, steps have been taken to ensure acceptable crosstalk levels in the array.
Finally, the SPAD array chip is designed to run on a 1GHz clock. It should interface with an FPGA that configures the hold-off circuit and reads out the SPAD status bits.

Preface

In these past two months, we have learned how to become multifaceted researchers and designers. Our work spanned from thorough literature reviews, giving recommendations to the Quantum Integration and Technology lab, to designing analog and digital 40nm CMOS circuits on a schematic and layout level. There were many novel concepts that we needed to grasp quickly to move forwards with our work, such as Diamond Nitrogen Vacancy readout, Single Photon Avalanche Diode design and control, mixed signal simulations, mixed analog and digital CMOS layout design and top level chip implementation.
We are grateful for the Quantum Integration and Technology lab for hosting us and giving us access to the necessary software and equipment. Furthermore, for giving us guidance during our weekly meetings so that we stay on track.

We would like to express our gratitude to our supervisor Dr. R. Ishihara for supervising our Bachelor Graduation Project, I, Varveris for being our daily supervisor and answering our many questions. Dr. S. Nur and N. Nitzsche for attending our weekly meetings. Dr. M. Babaie for attending our Greenlight assessment. A special thanks to L. Enthoven for guiding us through the entire chip design process, without him, we would not have achieved this far.
Furthermore, thank you to, Prof. I. Lager, for being an excellent Bachelor Graduation Project coordinator. Finally, much praise for M. Janssen, E. Nootenboom, and S. Britton for being fantastic group members, supporting each other throughout the entire thesis.

Samantha van Rijs
Lars Visser
Dagmar Westenbrink
Delft, June 2023

Contents

1 Introduction

The field of single photon detection is decades old and detectors have been realized using different technologies. Photomultiplier tubes (PMT) and electron-multiplying charge-coupled devices (EMCCD) have historically been used for these applications [14]. In the last twenty years single photon avalanche diodes (SPAD) have emerged as a promising technology for single photon imaging. Single photon avalanche diode (SPAD) arrays are solid state single photon detectors that offer imaging at the single photon level with sub-nanosecond timing accuracy. Specifically, the implementation of SPADs in standard CMOS technology has allowed research into the use of SPAD arrays as digital imagers in single photon applications[1]. As the technology has become more mature SPAD arrays have found use in a variety of fields, such as LIDAR [49], fluorescence lifetime imaging (FLIM) [1], positron emmision tomography (PET) [29] and a host of other biophotonic applications [14].
This work presents an integrated on-chip SPAD array including the readout and control electronics to be used as a detector for single Nitrogen-Vacancy (NV) centers in an integrated sensor. Such a sensor could allow the detection of variations in the magnetic field with a sensitivity of $pT/\sqrt{Hz}$ in an on-chip package. Such a miniaturized highly sensitive magnetometer is desirable for bio-photonic applications [14].
This Bachelor Graduation Project is part of the Quantum Integration Technology (QIT) lab's On-chip quantum sensing project [42]. First, the motivation and objective are given as to why this chip is designed. Then, a brief overview of NV centers and SPADs is given. After that, state-of-the-art analysis of SPAD quenching and recharge electronics is summarized. Finally, an outline of the remaining thesis is provided.

1.1. Motivation

Our SPAD array chip will be part of a package to enable highly sensitive and accurate readout of magnetic fields suitable for biomedical applications such as cancer tissue detection and readout of neural activity in the brain.
An example of how the package would look is in Fig. 1.1, where a certain bio-sample, with the focus on cancer tissue, emits a distinct magnetic field strength, labeled by magnetic nanoparticles, which is detected by the Diamond NV centers [19]. Then, the SPAD, in Geiger mode, detects the number of incoming photons which is then further sent to an FPGA for signal processing. NV centers can distinguish between different types of cells because they have unique magnetic footprints, which are intrinsically weak [57].

1.2. Objective

The aim of this project is to provide a fully integrated on-chip SPAD array including quenching, recharge, and interface electronics ready for the next tapeout which is at the end of July 2023. Initially, the QIT group provided us with a SPAD array fabricated in 40nm TSMC CMOS, which can be seen in Fig. 1.2. The array contains 20 SPADs of varying active areas in size and shape. The total size of the package is $1.21mm^2$. Currently, only the SPADs are integrated on-chip. We build upon this by assembling a new chip, with 40nm TSMC CMOS technology, consisting of a 16x16 SPAD array, including their

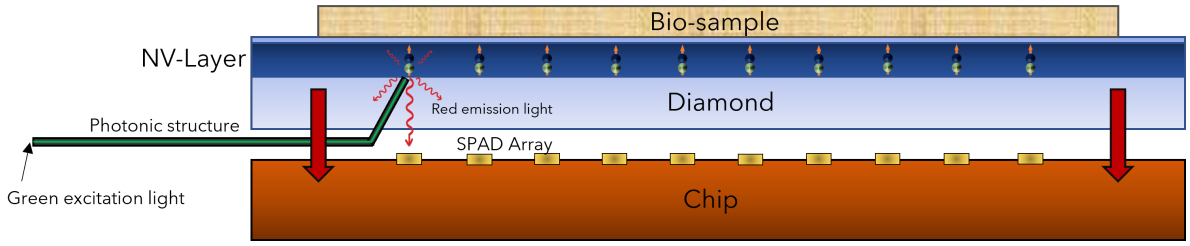


Figure 1.1: Drawing of the package that our chip is going to take part in for biosample magnetometry, courtesy to I. Varveris.

quenching and recharge circuits, and interface electronics. The chip should be capable of handling at least several Mcounts/s from individual NV center emitters. All this will be designed and simulated in Cadence Virtuoso.
Simultaneously, another group, the characterization group, working alongside us will test and characterize the current 20 SPADs and will guide us by choosing the best-performing SPAD for the 16x16 SPAD array.

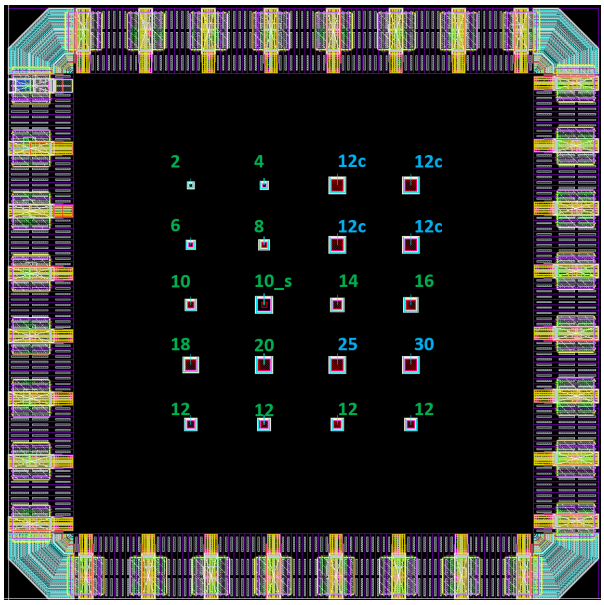


Figure 1.2: Layout of the SPAD array implemented in 40nm TSMC CMOS. The numbers are the dimensions of the active area. Except for the SPADS with subscript c (circular) they are square shaped, A x A (A[μ_m].)

1.3. NV centers

The Nitrogen Vacancy (NV) center is a point defect in a diamond where two adjacent carbon atoms are substituted by a nitrogen atom and a vacancy [15]. Diamond NV centers are sensitive to magnetic fields, electric field, strain and temperature due to the positioning of the NV center's energy level. Furthermore, the diamond NV centers are operable over a wide temperature range with its quantum coherence being stable at room temperature. These properties make the diamond sensor suitable for highly sensitive magnetic field readout applications [23].

1.3.1. Optical properties of the NV center

A simplified energy diagram of the NV center is displayed in Fig. 1.3 It consists of a ground state $|g\rangle$,

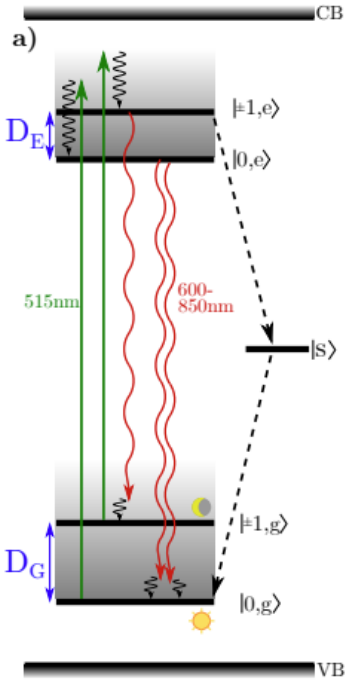


Figure 1.3: An energy diagram of the NV center. [15]

an excited state $|e\rangle$ and a shelving state $|s\rangle$ within the 5.5eV wide diamond band gap. The ground state is split into two magnetic spin levels $|m_s = 0\rangle$ and $|m_s = \pm 1\rangle$ or $|0, g\rangle$ and $|\pm 1, g\rangle$ for brevity. The excited state has the same structure. The decay path via the shelving state almost exclusively ends up in $|0, g\rangle$ [56].

When excited with a green laser electrons are transferred from the ground to the excited state in a spin-conserving process. The electron may then decay back to its ground state by emitting a red photon in the 637-850nm wavelength. A second, non-radiative, decay path exists for electrons in the $|\pm 1, e\rangle$ which end up predominantly in $|0, g\rangle$. This non-spin conserving dark decay path means that the NV center can be initialized in a well-defined spin state. A green laser can be used to cycle through these transitions, transferring the spin population from the $|\pm 1, g\rangle$ to the $|0, g\rangle$ state through this non-radiative decay path. Secondly this allows for an optical readout of the spin state of the system by monitoring the NV fluorescence count-rate [15].

1.3.2. Measuring magnetic fields with the NV center

Under influence of an external magnetic field B the $|-1,e\rangle$ and $|+1,e\rangle$ split, known as the Zeeman effect. This splitting can be seen on an optically detected magnetic resonance (ODMR) spectrum as seen in Fig. 1.4

The top image shows an ODMR spectrum of an NV center without an applied field. The resonance

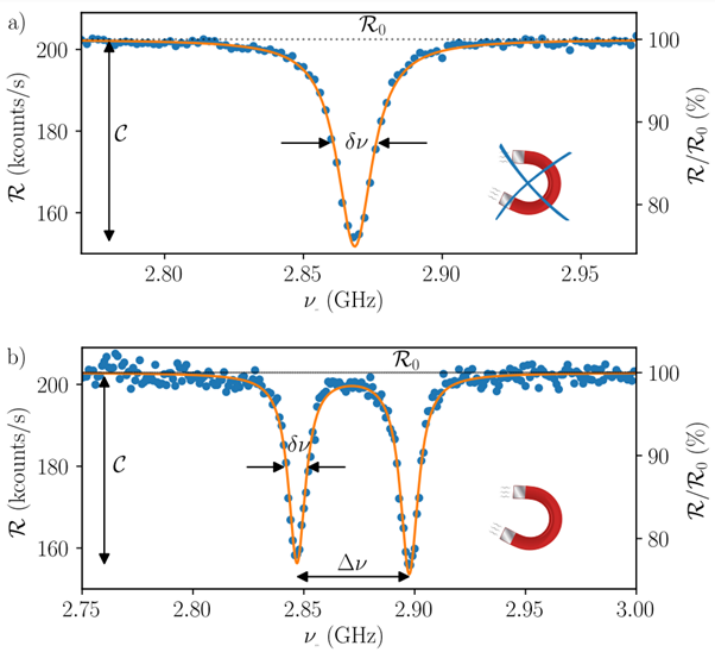


Figure 1.4: Zeeman splitting in an ODMR spectrum. An external field in b) causes a seperation of the resonance frequencies of the $|-1,e\rangle$ and $|+1,e\rangle$ states. [15]

peaks of the $|-1,e\rangle$ and $|+1,e\rangle$ states are at the same frequency ($D_G = 2.87Ghz$). In the second image an applied magnetic field causes a split between the $|-1,e\rangle$ and $|+1,e\rangle$ states. From this splitting Δv, the orientation and amplitude of the applied magnetic field can be derived, thus enabling the NV center to be used as a highly sensitive magnetometer in combination with ODMR. [15]

1.4. SPAD

Single Photon Avalanche Diodes (SPADs) can count single photons when configured in geiger mode. In our chip package, these photons originate from NV centers. As mentioned in section 1.3, by illuminating NV centers with a green laser and using ODMR the applied magnetic field can be determined. [15]. A SPAD can be viewed as a a p-n junction, which is reverse-biased above the breakdown voltage [53], putting it in geiger mode [46]. Since the SPAD is biased above its breakdown voltage with an applied voltage, a strong electric field (>3 × 105 V/cm) is present in the breakdown region. This high electric field means that a single photon can trigger an avalanche of impact ionizations resulting in a macroscopic current. The current avalanche will persist until the the bias voltage is lowered below the breakdown voltage which is called quenching. Then, the SPAD's applied bias voltage must be restored to re-arm the SPAD for the next incoming photon which is called recharging. The SPAD's quenching and recharging is done by external circuits[53]. Fig. 1.5 summarizes the avalanche, quenching and recharging.

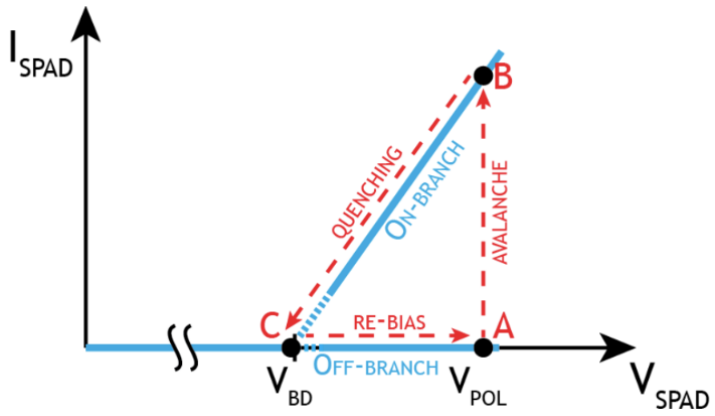


Figure 1.5: The SPAD's I-V characteristics. Before a photon has arrived, the SPAD is biased above the breakdown voltage in point A. When a photon hits the active area, an avalanche is triggered. Then, the current rises (B). Next, the quenching circuit lowers the voltage below the breakdown voltage (C). After which the voltage is restored by a recharge/re-bias circuit (A). [9]

Many different SPAD structures exist. Fig. 1.6 presents a common SPAD stucture, which is also used in Fig. A.4, the SPADs designed by the QIT lab. It has a p+/nwell junction, which is isolated from p- substrate. The guard ring is separated from the p-substrate by using a deep nwell [40].

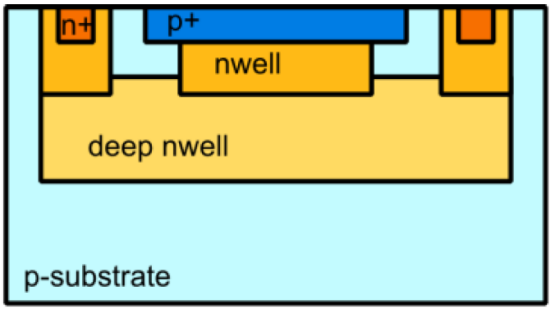


Figure 1.6: Cross-section of a SPAD, this is structure is implemented in our current SPADs designed by the QIT lab [40]

1.5. State-of-the-art analysis

On-chip integrated SPAD arrays have been worked on for more than a decade. The current trend is towards reducing pixel pitch to improve the fill factor leading to higher density arrays for a wide variety of sensing applications such as Fluorescence Lifetime Imaging Microscopy (FLIM) [24], LIDAR [50] and a host of biophotonic applications among others [1]. Quenching and recharge circuits are already capable of dead-times as low as a few nanoseconds, allowing detection rates of hundreds of Mcounts/sec [17]. SPAD arrays are very much the subject of active research and commercial arrays are not available at the time of publishing.

1.5.1. High density SPAD arrays

A mature SPAD array in 40nm technology was published in december 2017 [39], it comprised out of 64 SPADs divided into 4 macropixels of 16 SPADs each with their own dedicated electronics. At around the same time, a group from the University of Edinburgh came out with a 96x40 integrated SPAD array implemented on a 1x1mm^2 chip [5]. More recently, in April 2019, the same group published about a 192x128 array with STMicroelectronic 40nm CMOS technology including FLIM results. The largest array that has been published is the 1 MegaPixel by Morimoto et al. in [36], which proposes an entirely new architecture for the readout of temporal information in large scale SPAD arrays [27].
Although some preliminary work has been done on implementing SPAD arrays in 28nm CMOS [6], there aren't yet publications available on fully integrated SPAD arrays in that technology node.

1.5.2. State-of-the-art biosensing

The sensitivity of NV center's spin energy levels to electromagnetic fields and temperature, combined with the photo-stability at room temperature and the bio-compatibility of diamond, means that NV center based sensing has a wide variety of potential biological applications [41]. A non-exhaustive list of applications includes T_1 relaxometry [21], nuclear magnetic resonance (NMR) [12], magnetic resonance imaging (MRI) [31], measuring neuronal action potentials [28] and, cancer cell identification and tracking nanodiamonds *in vivo* [19]. Accurate determination of 3D magnetic field vectors is at the basis of all these technologies and that is precisely what NV based magnetometry promises. Since the last mentioned application is the project's focus, the current readout schematic is provided in Fig. 1.7. Our chip will build upon the diamond chip by integrating the NV center readout and Input-Output communication, taking over the functionality of the CMOS camera.
An exhaustive analysis of these applications is beyond the scope of this work. We will however compare three quantum-based magnetometry technologies and their sensitivities and limitations.

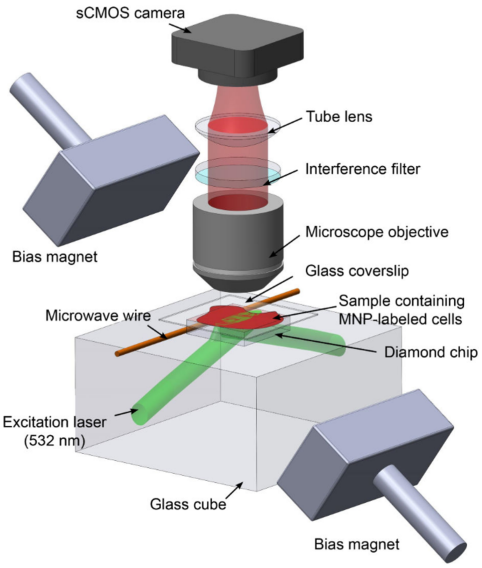


Figure 1.7: Schematic for cancer cell identification utilizing a NV diamond magnetic imaging microscope [19]. Our project would expand on the diamond chip by integrating the NV center readout and communication.

1.5.3. A comparison of three quantum based sensing techniques

In addition to NV-based sensors other devices that can detect weak magnetic fields include superconducting quantum interference device (SQUID) [22] sensors and optically pumped magnetometers (OPM) [20]. SQUID sensors have been employed in Magnetoencephelograhpy (MEG) for decades and can reach sensitivities of $ft/\sqrt{Hz}$. These sensors however require cryogenic operating conditions which comes with significant costs, device complexity and device size. OPM-based magnetometry has already been experimentally tested in MEG applications and although sensitivities are still behind SQUID sensors at $45ft/\sqrt{Hz}$ their room temperature operating conditions allow for placement closer to the tissue under investigation, mitigating this loss of sensitivity. [20]. However, the dynamic range of OPM sensors with fT sensitivity is so small that magnetic shielding is required to suppress local magnetic fields [55]. OPM-based magnetometry is a recently developed field and very much the subject of active researh. It is likely that key parameters such as dynamic range will improve in the coming years. NV-based magnetometry promises both room temperature operating conditions, but also a much higher dynamic range than OPM which could enable $pT/\sqrt{Hz}$ sensitivities in non-shielded environments [55]. MEG usually requires sensitivities of $fT/\sqrt{Hz}$ at a distance of a few mm [37]. However, the high sensitivity per volume for NV magnetometers makes them feasible for intraoperative applications [55]. Biomagnetic fields from muscle fiber bundles, such as the heart, can reach hundreds of pT above the skin and are thus more promising candidates for NV magnetometry. NV magnetometry thus has strong potential for applications in biosensing, with distinct advantages over OPM and SQUID based sensing but the sensitivity of NV magnetometry is several orders of magnitude lower than the alternatives at the moment.

1.6. Thesis synopsis

This thesis starts by providing the design requirements for our chip in Chapter 2. Next, the overview of the total design is given in Chapter 3, after which each part will be described in detail: the SPAD model (Chapter 4), the quenching circuit and hold off circuit (Chapter 5) and the Input Output controller (Chapter 6). The integration of all the components in the chip is described in Chapter 7. In Chapter 8, our results are discussed. Finally, in Chapter 9, a conclusion is made, in which a recommendation is given for future work.

2

Program of requirements

The final deliverable for the quenching group is a layout of a SPAD array with an on-chip quenching and recharge circuit combined with a input-output circuit. There are a number of requirements that dictate the design. Those requirements are laid out in this chapter.

2.1. Photon detection rate

In the final implementation, each SPAD will detect photons emitted from a single NV center. As such, the detection electronics have to be fast enough to detect successive photon emissions from its NV center to not bottleneck the measurement's resolution. In literature we found varying values for the fluorescence rate from single NVs ranging from 200 Kcounts/s [15] to 945 Kcounts/sec [7]. We take the higher limit of 945 Kcounts/sec and add a 3x safety margin. In this way we ensure that we can oversample have a low probability of missing events. We set the requirement at 3 Mcounts/s

- Photon detection rate above the 3 Mcounts/s, or cycle time of 333 ns.

2.2. Quenching and recharge

The previous requirement of 3Mc/s demands a quenching and recharge circuit that can complete a full cycle in less than 333ns. As both afterpulsing and crosstalk are correlated noise sources that depend on the accumulated charge in the breakdown region, it is desirable to limit the avalanche pulse time. A quenching time of several to at most a few 10ns is common in literature [10][45][16]. Therefore, a demand for our circuit is to quench the avalanche in at most 20ns.

- SPAD avalanche quenched in less than 20 ns.

2.3. Dead time

To limit afterpulsing, it is desirable to have a controllable dead-time, where the SPAD is turned off and insensitive to new detection events. Without a controllable dead-time the SPAD can be re-triggered during the recharge phase, leaving it susceptible to correlated noise from previous avalanches (afterpulsing) [53].

- Controllable dead time.

2.4. Crosstalk

With the decreased pixel pitch, crosstalk between the pixels becomes a concern. Based on SPAD array literature review, with varying pixel pitches that results in functional SPAD array, our chip's crosstalk probability between nearest neighboring pixels will be roughly assesed to be under the 2% leading to a pixel pitch between the 14 and 26 μm.

- Crosstalk probability less than 2%.

2.5. Pixels

The current chip has 20 SPADs in a 4x5 array in a $1.1x1.1mm^2$ package. The SPADs have different active areas and geometries. The active areas vary in dimension between 16x16μm to 50x50μm. Most of the SPADs have square active areas but there are some with a circular active area. The current pixel pitch is 100μm. Due to optical crosstalk, the SPAD array's pixel density is limited. With a pixel pitch between the 14 and 26 μm, and our limited chip area, a 16x16 SPAD array is the highest resolution to be achieved while satisfying the previous requirements.

- 16x16 SPAD array.

2.6. Size

The final implementation of the SPAD array will be on a chip with an area of $1.1x1.1mm^2$. Taking into account the IO ring this means that the available area for the SPADs and the electronics is roughly $1mm^2$.

- Chip area of $1.1x1.1mm^2$.

2.7. Input output

The amount of Input-Output (IO) pins at our disposal is 32. As the number of SPADs in the 16x16 array exceeds the number of available IO pins, the data from the SPADs will need to be serialized. An interface controller has to be designed that can communicate the status of the connected SPADs with sufficient speed to not miss any events.

- IO of 32 pins.

2.8. Technology node

As the SPADs are fabricated using the 40nm TSMC process, this is also a requirement for any additional on-chip circuitry. The transistors in this process are rated at a supply voltage of 1.1V.

- 40 nm TSMC process.

3

Chip overview

The chip design consists of 256 SPADs, a 16x16 SPAD array. All these SPADs require their own individual electronics for proper quenching and recharging. Moreover, interface electronics should be added due to our limited IO pins. This chapter provides a high level overview of the chip architecture.

3.1. Top level implementation

Fig. 3.1 shows the top level structure of the chip. The 16-bit SPAD electronics row consists of 16 SPAD units and the SPADs are connected to each other with a bus and a 16-bit Parallel-In-Serial-Out (PISO) circuit. On the top right corner, there are two 5-bit Serial-In-Parallel-Out (SIPO) modules to configure the width and delay of the hold-off circuit's recharge pulse. Next to it, there are three stray SPADs for characterization. They were requested to be added by the QIT group as there was chip area left over. The three SPADs are round with 12 μm diameter, square with 12 μm and square with 32 μm. Fig. A.4, in the Layouts appendix chapter, presents the layout of the whole chip. Section A.5, in the appendix, provides the chip driving requirements and advice.

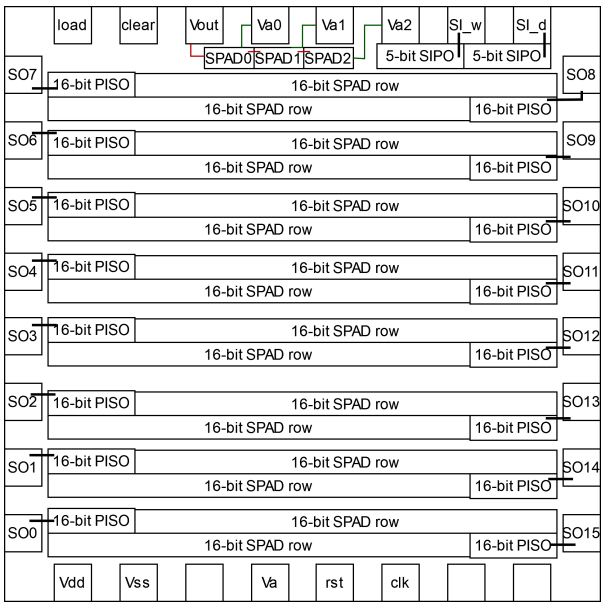


Figure 3.1: Top level strucutre of the SPAD chip including its pinout, 16-bit SPAD pixel row, 16 16-bit PISO, 2 5-bit SIPO, and 3 additional SPADs for characterization.

3.2. SPAD array

The SPADS are laid out in a 16x16 grid. The control electronics (quenching and recharge circuit and readout circuit) for each SPAD are co-located with the SPAD. Each row of 16 SPADs has a PISO to stream data off-chip. This is needed as there are only 32 IO pins. For the communication, 16 lines send out 16 serial SPAD status bits which notify the FPGA if an avalanche has occurred in a given time frame. The interconnection of the 16-bit SPAD row connected to their respective 16-bit PISO is in Fig. 3.2. 'SO' is the serial output to the FPGA. The SPAD unit's 'aqc_out' are parallel connected to the PISO. There is a global bus that propagates the applied voltage 'Va', the width and delay bits, the reset, and, the clock to the row's SPAD units. Then the rows are stacked on top of each other as in Fig. 3.1. The SPAD unit is described in the next section, and the PISO and SIPO modules are described in Chapter 6.

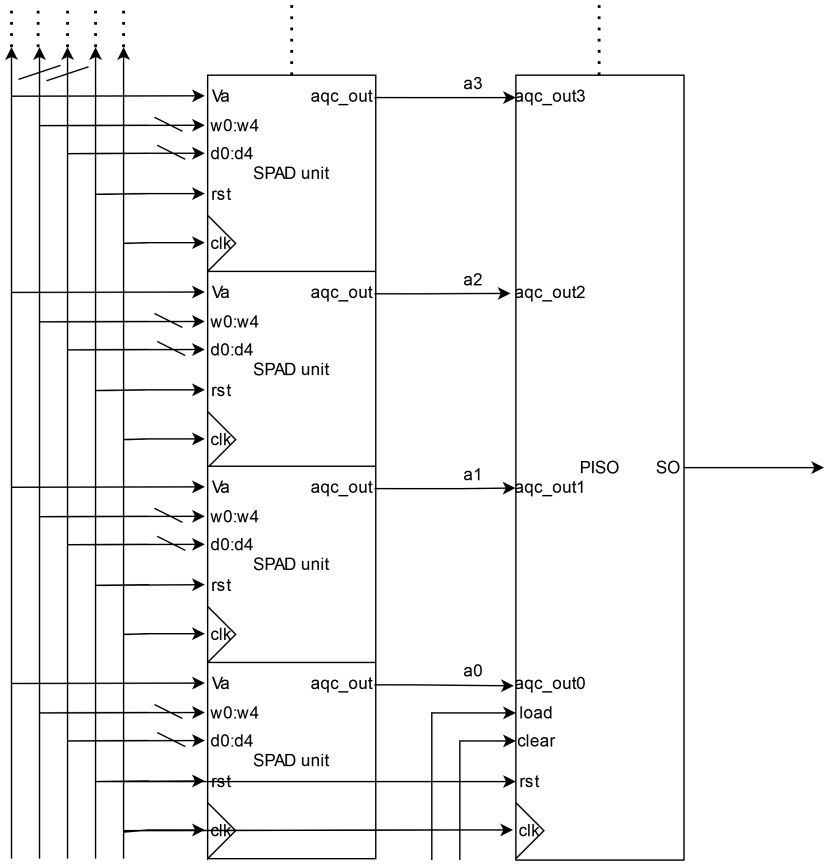


Figure 3.2: Schematic of the first four SPAD units connected to each other to form the row column.

3.3. SPAD unit

Each SPAD has its own active quenching circuit (AQC) and hold-off circuit. These are necessary to quench and recharge the SPAD. The schematic is shown in Fig. 3.3. ‘Va’ biases the SPAD, and ‘pwm_out’ sends the recharge pulse. Furthermore, the inverted AQC output ‘out’ is connected to the hold-off module and is the SPAD unit’s output. Chapter 5 describes the quenching- and hold off circuit in detail.

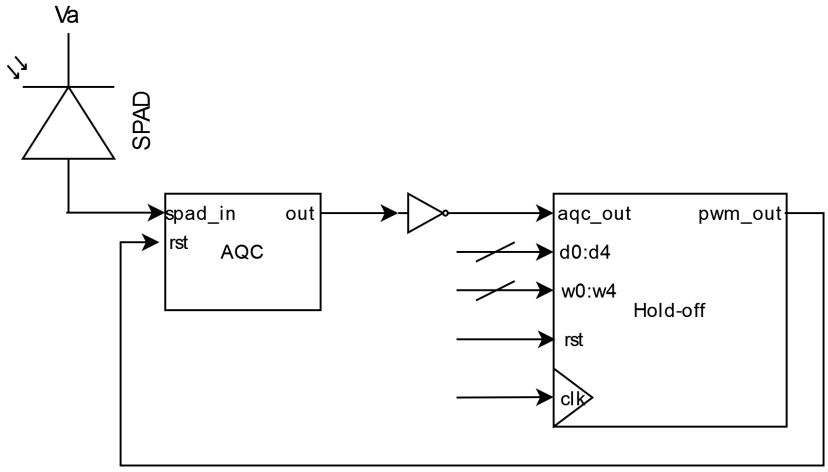


Figure 3.3: The SPAD unit where the SPAD, AQC and hold-off circuit are interconnected to each other.

4

SPAD model

To design the quenching circuit and verify whether it can fulfil the requirements outlined in Chapter 2, a model for the SPAD behaviour is employed. It is a behavioral model written in VerilogA, based on the I-V characteristics of the SPAD. It's an adaptation of the model developed by Mita et al in 2011 [18] which itself expands on their earlier 2008 work [33]. The source code for the model is included in Appendix A.2

4.1. Model's adherence to characterization results

During the project, the characterization group, working alongside us, was unable to experimentally determine an I-V curve and the capacitances for the QIT's SPADs due to an unforeseen IO ring issue in the chip. They did characterize the electrostatic discharge (ESD) diodes in the IO ring. As such, The figures presented in this chapter and throughout the rest of the paper use that I-V characteristic. As such, the predictive value of the models presented is severely constrained with respect to the fabricated SPADs.

4.2. Model selection

In Fig. 4.1, the I-V characteristics of the models [18] and [33] are plotted together. The blue line represent an I-V characteristic of an ESD diode in the IO ring of the chip. The breakdown voltage and series breakdown resistance were experimentally determined at 7.42V and 5650Ω respectively. We worked with reasonably conservative values of 1-2pF for the parasitic capacitances as given by Zappa in [54]. The 2008 model piece-wise approximates the SPAD's I-V characteristic. This approach has the downside that convergence problems arise at the transitions between the piece-wise functions [18]. The pseudo-max function, introduced in the 2011 paper, circumvents that issue and better approaches the slope of the experimental curve in the breakdown region, which is the primary region of interest. However, there exists some discordance between the model and the experimental data near the measured breakdown voltage of 7.42V where the model underestimates the current. Due to the clearly better performance of the I-V curve based on the 2011 model, this is the model that will be used throughout this paper. A normalization voltage V_n of 0.1V was chosen to best fit the experimental curve, several different values for V_n are presented in Fig. A.1.

4.3. Model details

The SPAD is modeled as a voltage controlled current source following the equivalent circuit model in Fig. 4.2. In a quiescent state, only the inverse saturation current I_s flows. During an avalanche, the current is modeled as a function of the diode's breakdown series resistance R_{brk}, the excess voltage ($V_D - V_B$, where V_D is the voltage across the SPAD and V_B is the breakdown voltage.) and a normalization voltage V_n that is used to fit the model to the experimentally determined I-V curve. The current-to-voltage behaviour of the SPAD is described by Equation 4.1.

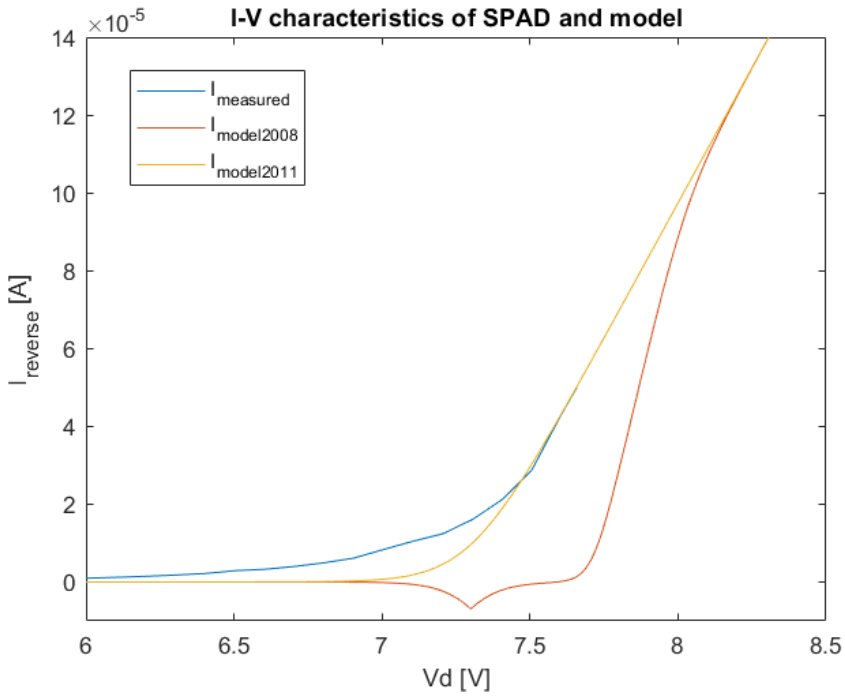


Figure 4.1: IV characteristics of the measured SPAD and the 2008 and 2011 model of Mita [33] [18].

$$I_{SPAD} = \begin{cases} I_s, & \text{if no avalanche} \\ I_s + \frac{V_n}{R_{brk}} ln[1 + e^{(V_D - V_n)/V_n}], & \text{if avalanche} \end{cases} \tag{4.1}$$

The dynamic current behaviour is governed by three parasitic capacitance contributions: the charge

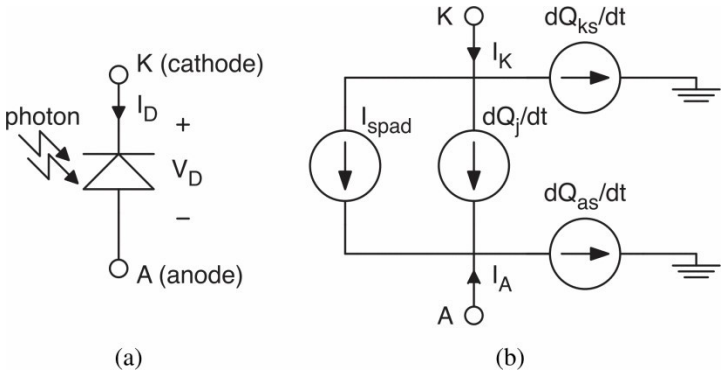


Figure 4.2: The SPAD model. (b) shows the dynamic model. [18]

in the depletion layer Q_j and the charges stored in the cathode-to-substrate Q_{cs} and in the anode-to-substrate Q_{as} stray capacitors. The charges Q_{cs} and Q_{as} are modeled as linearly related to the terminal voltages:

$$Q_{cs} = C_{cs}V_C \; Q_{as} = C_{as}V_A \tag{4.2}$$

Where C_{cs} and C_{as} are constant capacitors that have to be determined experimentally. The charge in the depletion region is given by

$$Q_j = A_D \frac{\phi_i C_{j0}}{1 - m_j} \left(1 + \frac{V_D}{\phi_i}\right)^{1-m_i} \tag{4.3}$$

Where A_D is the device area, ϕ_i the built-in voltage, C_{j0} the zero-bias capacitance per unit area and $m_j = 0.5$ the junction grading coefficient [18]. The parasitic capacitances are important parameters for

the quenching and recharge performance of the SPAD and should be determined experimentally [52].

4.3.1. Photon arrival and turn-off

A photon arrival is modeled by a picosecond duration pulse that has a voltage above a preset threshold. In a real SPAD, the photon detection efficiency (PDE) determines the probability that a photon incident on the SPAD surface [10], however, for the purpose of simplicity in the model, every ”photon” generates an avalanche.
The quenching of the avalanche is determined by the latching current I_{lat}. As the avalanche is a statistical process, there exists a certain threshold current below where not enough impact ionizations occur for the avalanche to be self-sustaining [53]. This latching current is generally taken to $\sim 100\mu A$ [54] and is set at that value in this model. below that value, the avalanche is turned off.

5

Quenching and hold-off circuit

SPADs operate in Geiger mode, where the device is reverse biased at a voltage beyond its breakdown. In this operating regime, the electric field in the breakdown region is so large that a single carrier can trigger a self-sustaining current avalanche by impact ionization in the multiplication region [10]. A quenching circuit is therefore required to terminate the multiplication and return the device back to its operating condition to detect subsequent events. A quenching circuit can be active or passive, each with its own advantages and disadvantages. This chapter lays out the two different approaches and uses simulations, with the program of requirements in mind, to determine that an active quenching circuit (AQC) is required. The proposed topology for that AQC is laid out and verified with simulations. An analog and digital hold-off circuit to tune the dead-time of the AQC are presented and the design decision for the digital circuit is justified with simulations.

5.1. Passive quenching

The simplest way to quench the SPAD's avalanche, is with a large ballast resistor R_L around the tens to hundreds of kΩ in series with the SPAD (see Fig. 5.1). The diode current $I_d(t)$ is given by equation 5.1.

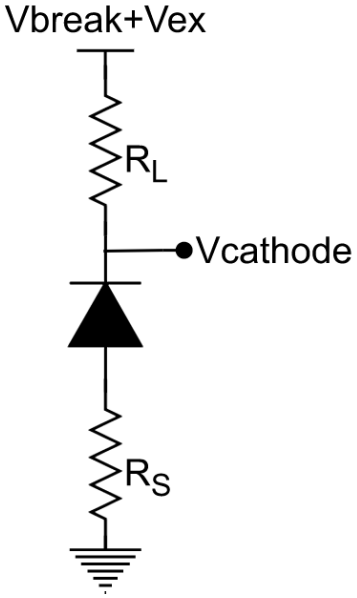


Figure 5.1: Passive quenching circuit. R_S is a sensing resistance of typically 50Ω used to obtain a measurable signal. R_L is the ballast resistor to quench the SPAD.

With $V_d(t)$, the diode voltage, R_d, the diode series resistance in breakdown, and, V_B the breakdown

voltage.

$$I_d(t) = \frac{V_d(t) - V_B}{R_d} = \frac{V_{ex(t)}}{R_d} \tag{5.1}$$

When an avalanche current flows, a voltage drop develops across R_L thus lowering the excess voltage V_{ex} and in turn the diode current I_d. As avalanche multiplication is a statistical process, once the current drops below a threshold value, there are not enough impact ionizations to sustain the avalanche and the current is quenched. A typical value for this latching current I_L is 100μA [54]. The quenching time constant T_q is determined by the SPAD's junction capacitance C_d, the terminal's sum of stray capacitances C_s connected to R_L and by the parallel resistance of R_d and R_L. As the condition $R_d << R_L$ holds, the quenching time can be approximation into equation 5.2.

$$T_q = (C_d + C_s)\frac{R_d R_L}{R_d + R_L} \approx (C_d + C_s)R_d \tag{5.2}$$

When the avalanche is quenched, the capacitances recharge to the applied bias voltage V_A with a time constant in equation 5.3.

$$\tau_r \approx (C_D + C_P)R_L \tag{5.3}$$

A small sense resistor R_S, typically 50Ω [48], converts the current signal into a voltage signal that can drive a comparator to do measurements. This circuit is simulated in Cadence for several R_L values in Fig. 5.2.

From this figure, it can be verified that Equations 5.2 and 5.3 seem to hold. The quenching time only

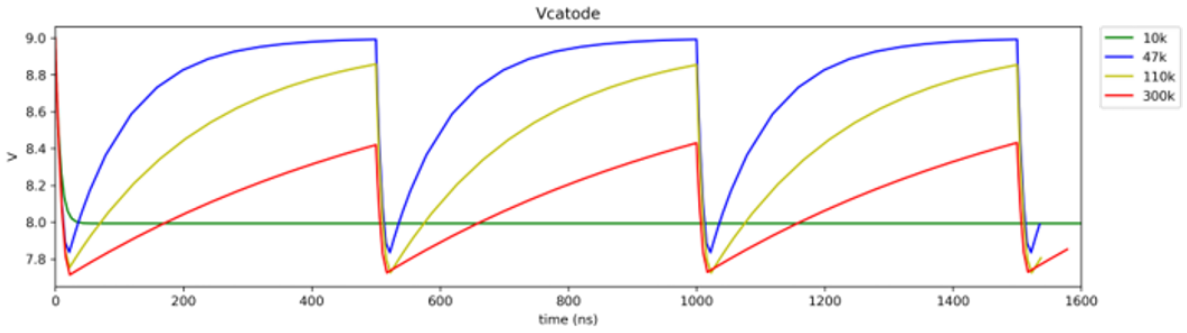


Figure 5.2: Cathode voltage of a passive quenching circuit as depicted in Section 5.1 for various R_L at applied voltage $V_A = 9V$. A photon arives at 0ns, 500ns, 1000ns and 1500ns.

varies slightly for the various R_L whereas large differences can be observed in the recovery time of the SPAD. Note that the 10kΩ resistor is insufficient to quench the SPAD. The Passive Quenching Circuit (PQC), with $R_L = 47k\Omega$, is restored to within 1% of its operating voltage after roughly 380ns whereas for $R_L = 110k\Omega$ and $R_L = 300k\Omega$ the recovery time is well in excess of 500ns. Furthermore, the SPAD is not insensitive to detection events during the recharge. The SPAD can be re-triggered once its voltage is restored above the breakdown voltage, which is much earlier than the recovery to its operating V_E. This ill-defined dead time means that the voltage signal at the sense resistor may only be a fraction of the designed voltage due to premature avalanche triggers.

5.1.1. Conclusion

The program of requirements outlines a required count rate of 3Mc/s and a controllable dead time. A full cycle takes roughly 400ns for $R_L = 47k\Omega$ which is significantly slower than the maximum allowed 333ns. Furthermore the dead time of the passive quenching circuit cannot be controlled, violating another requirement. The passive quenching approach is not suitable for our design purpose.

5.2. Active quenching

A second approach to quenching is commonly referred to as active quenching. Here, a circuit is used to detect the onset of an avalanche and then a positive feedback loop is used to quench the avalanche and

rapidly recharge the SPAD back to its operating excess voltage [11]. The advantage of this approach is that the recharge time is no longer dependent on the value of the large ballast resistor R_L, as in the passive circuit approach, greatly improving the SPAD's recharge speed and allowing much faster photon detection rates up to several 100Mcounts/s [13] [3] [4]. Active quenching can be broken down into three stages: First a sensing stage senses, the onset of the avalanche, then a feedback circuit is used to quench the avalanche by lowering the cathode voltage of the SPAD below its breakdown voltage and, finally the SPAD is reset back to its initial operating condition after a dead-time set by a hold-off circuit. This controllable dead-time allows the user to make a trade-off between dead-time and afterpulsing effects. Afterpulsing is a correlated noise effect caused by local defects in the depletion layer [52]. During an avalanche, carriers can get trapped in these defects and they are released within a time delay. If they are released when the SPAD is sensitive to a carrier then these released carriers can trigger spurious avalanches, compromising the detector dynamic range. A longer hold-off time, the time between quenching the avalanche and recharging it back to its operating condition, allows a portion of these traps to depopulate but this comes at the cost of limiting the maximum achievable count rate [10].

5.3. Active quenching circuit

The chosen active quenching circuit is one proposed by Mita et al. [32] and is depicted in Fig. 5.3. The AQR circuit is of low complexity, comprising only 5 transistors and therefore is a good candidate for an integrated approach, given that it can achieve the required count rates. The hold off circuit, that drives the reset transistor M_R is explored later in this chapter.

In quiescent condition, node **A** is set to ground, transistor M2 is on and the output voltage is high.

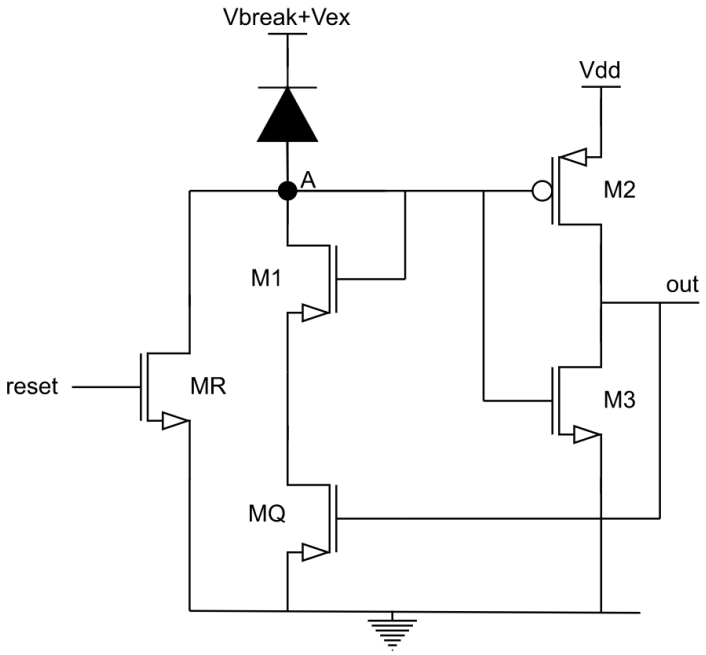


Figure 5.3: Active quenching circuit by [32]

When an avalanche is triggered, the avalanche current flows through M1 and MQ. With the gate of MQ connected to VDD, MQ provides a low-resistance path. As long as this is true, M1 and M3 are in a current-mirror and the avalanche current appears mirrored on the output branch. The subsequent reduction of the output voltage increases the MQ's channel resistance pulling up the voltage at node **A**. In this way, M1, M3 and MQ are a positive feedback loop that accelerates the quenching process. PMOS transistor M2, that forms an inverter with M3, further contributes to the quenching speed up process. At a certain point, the output voltage switches transistor MQ off, breaking the current path for the avalanche current and rapidly quenching the SPAD. Finally, recharging transistor MR is switched on, pulling node **A** back to ground and the SPAD is ready for another detection [32].

5.3.1. Transistor sizing

The AQC's design decisions mainly relate to the sizing of the transistors. Mita already states in the paper that for the best performance in terms of speed, "M1, M2, and MQ should be set with a minimum aspect ratio, thus reducing the overall parasitic capacitances [32]." Theoretically, non-minimal sizing of transistor M3 should reduce the high-low transition at the output. Mita remarks that, although this is true as the switching time of MQ should decrease if M3 is non-minimal, in practice the quenching time is not limited by MQ's switching time but rather by the time constant of the SPAD itself. Indeed, simulations for non-minimal aspect ratios of M3 in Fig. 5.4 show no significant difference in the AQC's quenching time. However, transistor MR, should be of non-minimal size. The transistor's time constant contributes

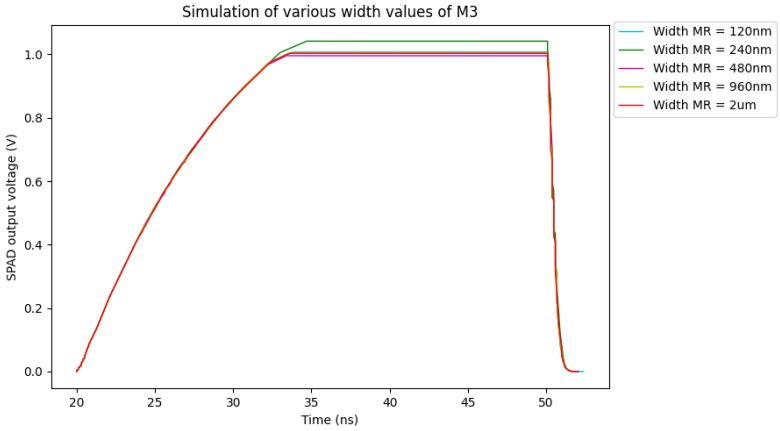


Figure 5.4: A plot of various widths for transistor M3. There is no appreciable different in quenching time, indicating that the quenching is limited by the intrinsic SPAD time constant. $V_A = 9V$

significantly to the speed of the reset phase. As such a wider transistor, lowering the resistance, is desirable [32]. It is clearly seen in Fig. 5.5 that the aspect ratio (AR) of MR has a significant impact

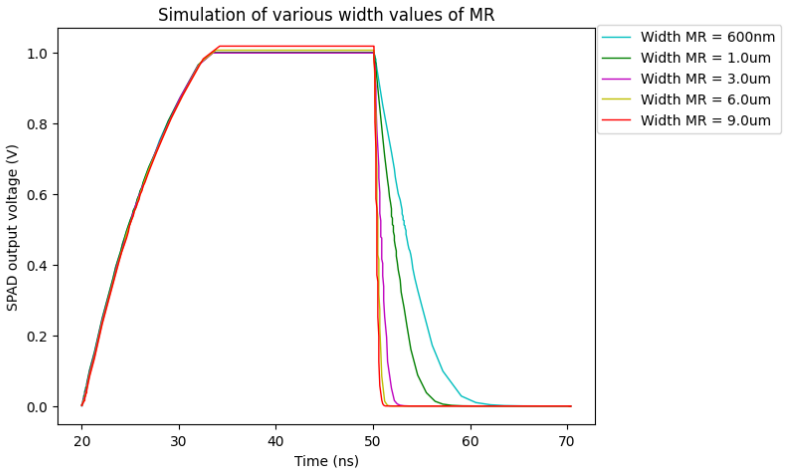


Figure 5.5: A plot of various widths for transistor MR. As the aspect ratio increases the recharging speed gets faster. Length of MR is minimal (40nm). $V_A = 9V$

on the recharge time of the SPAD. For a width of 600nm (AR = 15), the recharge time is more than 10ns. The recharge time for a width of 6μm (AR = 150) is around the 1 ns. The improvement for going beyond an aspect ratio of 150 is marginal.

5.3.2. Conclusion

All transistors are chosen with minimal aspect ratios with the exception of MR. MR is sized with a minimal length of 40nm, but a width of 6μm for an optimal balance between recharging speed and size. With the chosen transistor values, see Table 5.1, the active quenching circuit more than meets the set out requirements. It quenches the avalanche in 12s, being limited by the SPAD time constant and can recharge the SPAD back to its operating excess voltage in around 1ns. Even with the fairly conservative values for the SPAD's parasitic capacitances that were modeled here of 1-2pF. The AQR circuit can execute a full cycle in less than 13ns leading to a maximum count rate larger than 76Mcounts/s. Note that this is the theoretical limit for the modeled SPAD and that in practical applications it is common to insert a delay of up to several nanoseconds before recharging the SPAD to allow traps to depopulate and thus reduce afterpulsing effects. Any change in SPAD capacitance and breakdown-region resistance will neccesarily have an impact on these figures as well.

	MR	MQ	M1	M2	M3
Length (nm)	40	40	40	40	40
Width (nm)	6000	120	120	120	120
Ratio	150	3	3	3	3

Table 5.1: Active quenching circuit transistor sizing.

5.4. Hold off circuit

Active quenching and recharge circuits not directly connected to digital post-processing circuits or a computer for readout need a hold-off circuit to control their dead-time[18]. Several implementations of such circuits exist. Tunable delay lines specifically, based on exploiting the propagation delays of various analog or digital implementations, have seen widespread adoption [2]. Monostable circuits are another candidate for delay circuits in integrated SPAD applications. They are a class of circuits that generate an output pulse of specified width when triggered before returning to their initials stable state. One such implementation is done by Mita in [32] where they used a one-shot generator with a time delay set by an RC circuit.
This section evaluates an analog and digital hold-off circuit. The advantage of analog is its low-complexity architecture. However, the tunability of the delay time and pulse width is less precise than the digital circuit. While the digital implementation is of higher complexity and thus chip area, given a high clock frequency, the delay time and pulse width can be more precisely controlled.

5.5. Analog hold-off circuit

The analog hold-off circuit exploits the transistor's gate capacitance by varying the gate's electric field. It consists of two parts, a delay-, and a pulse width circuit proposed by Tang et al. [47]. It is a 13 transistor circuit and has much lower area requirements compared to the digital implementation discussed later on. The delay transistor circuit is in Fig. 5.6 and the pulse width circuit is provided in the appendix, Fig. A.2. To test out the analog option, the delay circuit is implemented and simulated for different V_{rc}, which tunes the delay length at 'out'. Simulations show that there is a non-linear relation between delay time and the applied analog tuning voltage V_{qc}. In Fig. 5.7, it can be seen that with a ΔVqc of 0.02V, from 0.7 to 0.86 V, the delay time has already varied larger than 10 ns. For our hold-off circuit, we aim for a delay time between 5 and 20 ns. Consequently, this analog circuit is not suitable since the non-linearity falls in that range and minimal voltage swings will already introduce delay time variations in the order of ns.
Additionally, in practise, the complete analog hold-off circuit will require two pins connected to stable low-noise power supplies. This is impractical because we would like the chip to be powered and driven by an FPGA and standard supply.

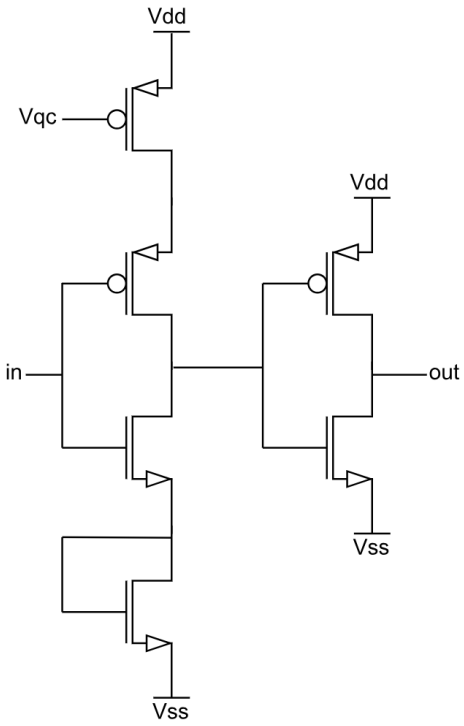


Figure 5.6: The delay part of the analog hold-off circuit.

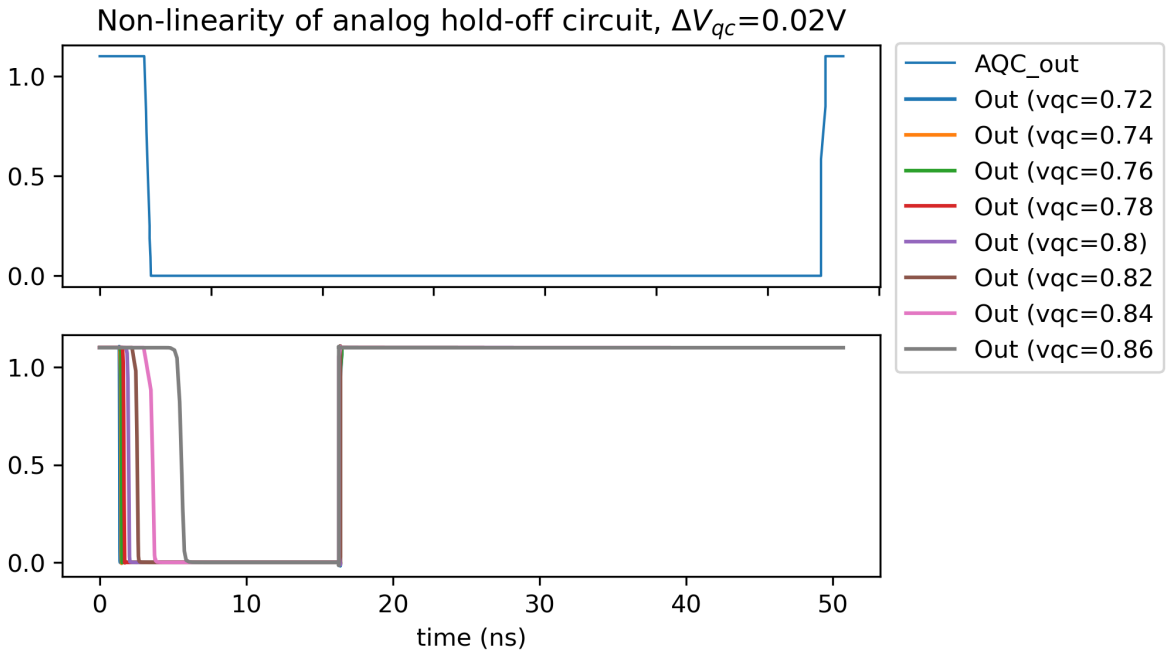


Figure 5.7: Simulation of the analog hold-off circuit's delay part. Observe the non-linearity of the delay time.

5.6. Digital hold-off circuit

In the previous section, we explained why an analog hold-off circuit implementation is impractical despite its low circuit complexity. Therefore, the decision was made to pivot to a digital implementation. A simple architecture for a hold-off circuit is to use a pulse width modulation (PWM) circuit driven by the output of the quenching circuit. The delay and width of the pulse can then be used to tune the dead-time

of the SPAD and drive the reset transistor in well-defined increments. The PWM circuit consists of a 5-bit counter, two 5-bit comparators, and two JK flipflops. See Fig. 5.8.
The JK flipflop, connected to the AQC, acts as a buffer for the 'out' signal. The JK flipflop, connected to

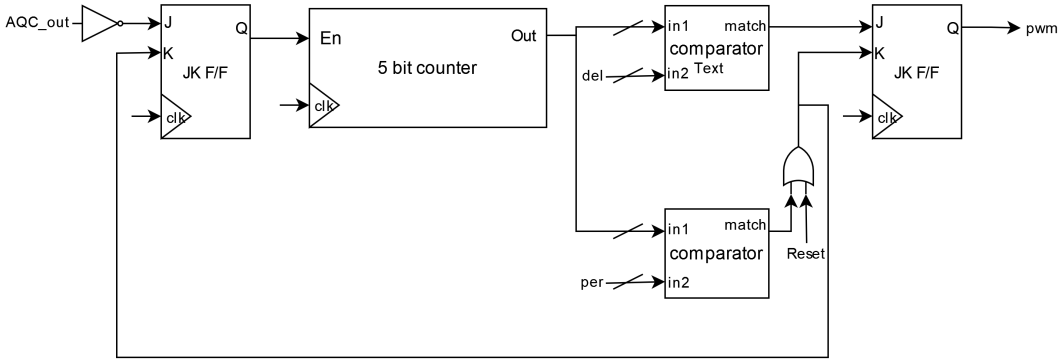


Figure 5.8: Top level schematic of digital hold-off circuit.

the comparators, produces the AQC's reset signal which is a PWM pulse. One comparator determines the pulse delay while the other determines the pulse's period or width.
When an avalanche is detected, the AQC produces an active low 'out' signal. For the hold-off circuit, the 'out' signal is first inverted to active high and stored in the JK flipflop. When the JK flipflop holds a '1', the 5-bit counter is enabled and starts counting. When the 5-bit counter matches with the comparator's delay value, the JK flipflop next to it is set to '1'. As the 5-bit counter keeps on counting, there will be a second match with the period comparator, resetting the JK flipflop back to '0', thus the AQC-reset signal is set to zero again. As the PWM pulse has reset the AQC, the SPAD is biased back to its applied voltage, 'out' will become '1' so '0' at the JK flipflop's input and the 5-bit counter is disabled, the hold-off circuit is off. Moreover, to determine the delay bit vector, use equation 5.4. To determine the width bit vector, use equation 5.5. t_{delay} is the time between the AQC 'out' signal and when the reset pulse should be '1'. t_{width} is the time that the reset pulse should stay high. The delay and width bit vectors should be extended to 5 bits during binary conversion.

$$d[0:4] = (\lfloor t_{delay} t_{clockcycle} \rfloor)_2 \tag{5.4}$$

$$w[0:4] = (\lfloor t_{width} t_{delay} t_{clockcycle} \rfloor)_2 \tag{5.5}$$

Additionally, there is a global reset, connected to the OR-gate, which will always set the AQC reset signal to zero.

5.6.1. JK flipflop

The JK flipflop is the fundamental building block for sequential circuits such as the 5-bit counter. When referring to the JK flipflop, it will always be configured in master-minion configuration to avoid the race around condition[44]. The master flipflop samples and holds the input while the minion flipflop updates its output based on the master's state at clock edge. The JK flipflop's schematic is in Fig. 5.9. When port 'J' is set high, the flipflop holds a one, when port 'K' is set high, the flipflop holds a zero. When 'J' and 'K' are both high, the flipflop toggles. Moreover, the JK flipflop is negatively clock edge triggered. 'Q' and 'Q_bar' hold the JK flipflop's output and inverted output.

5.6.2. 5-bit counter

The 5-bit counter is a chain of toggling JK flipflops activated by signal 'en', the enable signal. The counter is asynchronous, including asynchronous reset, since 40nm pmos and nmos transistors are fast enough to not introduce inter-clock-delay, which is validated with the 5-bit counter's post-layout simulation in the appendix chapter A.4. There is negligible speed slowdown. Moreover, the asynchronous configuration has lower complexity than its synchronous counterpart thus reducing the 5-bit counter size. The synchronous counter contains AND gates between the flipflops. Fig. 5.10 shows the counter's schematic.

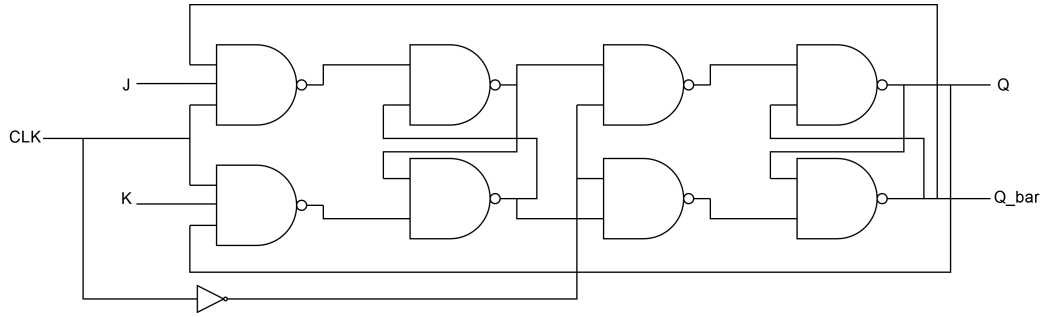


Figure 5.9: The JK flipflop in master-minion configuration.

The counter is five bits, it will count until 31, which is sufficient for the hold-off circuit. Four bits, counting until 15, is not a sufficient resolution for the hold-off circuit. The counting includes the delay and the pulse width, four bits would be sufficient if the delay time is not included. The hold-off circuit's PWM pulse resolution is dependent on the clock frequency and the number of counts. The clock frequency determines the count cycle time, the number of counts determines the range. Validated with simulations, at clock a frequency of 1 GHz, 5 bits is sufficient, $(t_{delay} + t_{width})_{1GHz,max} = 31$ ns, 4 bits, $(t_{delay}+t_{width})_{1GHz,max} = 15$ ns, limits the tunability too much, and 6 bits, $(t_{delay}+t_{width})_{1GHz,max} = 63$ ns, would be unnecessarily complex.

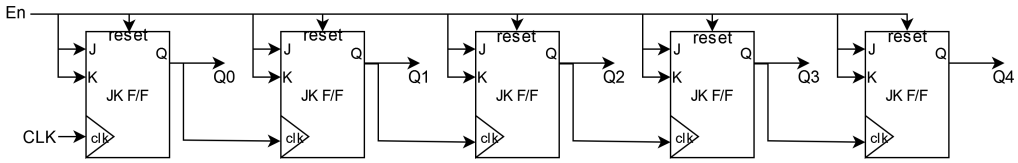


Figure 5.10: The 5-bit counter in asynchronous configuration.

5.6.3. Comparator

The 5-bit comparator matches the 5-bit counter's count with a set 5-bit delay or width value. When there is a match, the comparator propagates a '1'. Its implementation is 5 XNOR gates connected to a 5-terminal AND gate, see Fig. 5.11.

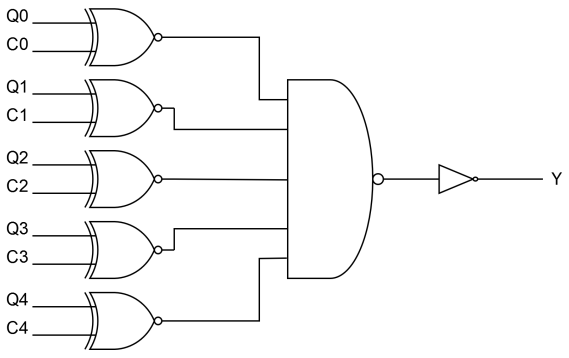


Figure 5.11: The 5-bit comparator.

5.6.4. Simulation results

To verify the digital hold-off circuit, a transient simulation is performed which is in Fig. 5.12. The hold-off circuit is connected to the AQC, which is also connected to the SPAD model. As shown, a photon triggers an avalanche which sets the inverted AQC's output to '1'. This turns on the 5-bit counter and based on the 5-bit period and delay value, a PWM pulse is generated. The PWM signal resets the SPAD and the AQC inverter output is back to '0', ready for a new photon event.

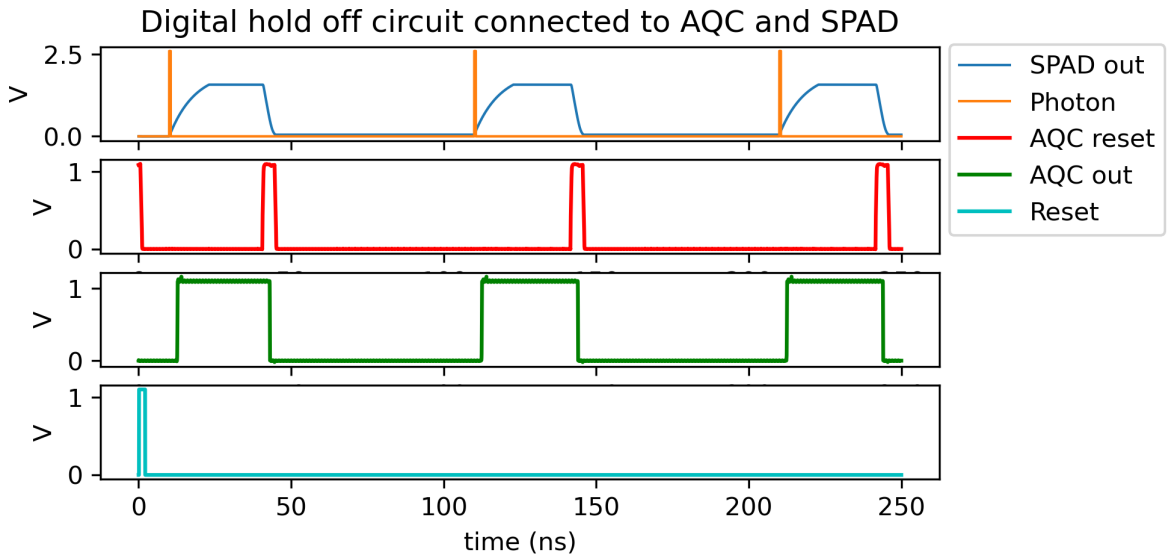


Figure 5.12: Simulation of the digital hold-off circuit together with the AQC and SPAD model. A global reset (teal) initialized the system, a photon (orange) triggers an avalanche at which point the avalanche is quenched by the AQC (blue sloped curve). During the quenching process the AQC output switches to low, which after a preset delay triggers the reset (red) at which point the SPAD is recharged to quiescent conditions.

5.6.5. Conclusion

After inspecting analog and digital options for a tunable hold-off circuit, we opted for a digital implementation. Furthermore, our digital hold-off circuit consists of a 5-bit counter, two comparators, two JK flipflops and additional logic gates. To evaluate that it can drive the AQC, simulations such as Fig. 5.12, have been performed which confirms that the hold-off circuit can accurately reset the SPAD while not sacrificing speed.

6

Input output

For the final layout of the chip, serial communication is necessary when the number of SPADs exceeds the number of input output (IO) pins. In our case, for the 1.21 mm^2 chip area, we have 32 IO pins at our disposal, where three are already reserved for VDD, VSS and the SPAD's applied voltage.
For a fast enough solution, there should be a balance between parallel and serial readout to achieve high-speed communication. Our current proposal is 16 communication lines that send out 16 serial SPAD status bits which notifies the FPGA if an avalanche has occurred in a given time frame. With this, a 16x16 SPAD matrix can be achieved that reaches the speed specifications. Consequently, every 16-bit SPAD column should contain a 16-bit Parallel In Serial Out (PISO) module.
Moreover, to configure the hold-off circuit's delay and period, two extra IO pins need to be reserved for 5-bit bitstreams thus the SPAD chip will also hold two 5-bit Serial In Parallel Out (SIPO) modules.

6.1. Parallel In Serial Out

A 16-bit Parallel In Serial Out (PISO) holds a 17-bit shift register and a 16-bit buffer. The shift register is 17 bits because the first bit in the serial bitstream is a start bit to facilitate communication with the FPGA. There are 16 buffers, that are connected to each AQC 'out' signal, to hold onto the avalanche status bit until it is cleared. To drive the PISO module, it needs a clock, reset, load, and clear signal. The 'load' signal polls the chip to serially send the signal that is currently in its buffer. the 'clear' signal empties the buffer, ready to hold new status bits.
Therefore, to read out the SPAD bits, the FPGA constantly polls the chip via 'load' and 'clear'. Since the AQC and hold-up circuit can reach speeds above the 10 Mcounts per second, oversampling can be achieved to not miss any avalanche events.
Fig. 6.1 gives an example of how a 3-bit PISO module would be implemented, to extend the PISO to 16 bits, the logic gates and flipflop chain will be elongated. In our design, the D flipflop is implemented with a JK flipflop by connecting the JK inputs together with an inverter. Furthermore, to make it positively edge-triggered, the clock is inverted inside de D flipflop.
The leftmost D flipflop is tied to Vdd, which is the start bit. Output 'SO' will hold the 4-bit bitstream.
To demonstrate the serialized concept, the 3-bit PISO is simulated together with 3 SPADs each with their own AQC and hold-off circuit. The result is in Fig. 6.2. Signal 'SO' contains the start bit and the 3-bit SPAD status array. In this simulation, the first and last SPAD have received a photon which is reflected back in the simulation result. Signal 'load' polls the PISO module and 'clear' empties the buffer, which needs to be sent at the end of every communication transaction.

6.2. Serial In Parallel Out

To set up the delay and period of the hold-off circuit, the SIPO module with buffer needs to be implemented. Furthermore, to facilitate communication, the SIPO module automatically detects a start bit and locks the communicated bits correctly into the buffer, which is connected by a bus to all the hold-off circuits. With a reset signal, the buffer is emptied to lock new bit vectors into place.
Fig. 6.3 is the SIPO's schematic. To note, at the end of the flipflop chain, there is a negatively edge triggered D flipflop and JK flipflop. This is to control the select signal 'sel' that will lock in the serialized

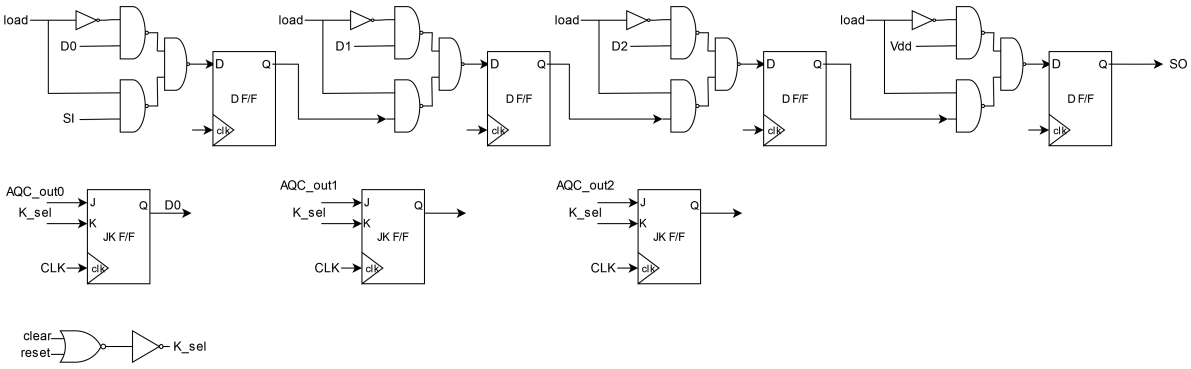


Figure 6.1: Schematic of a 3-bit PISO module. The shift register contains D flipflops, the buffer is a row of JK flipflops and there is an AND gate to accommodate for both clear and reset. Notice that the left-most D flipflop is tied to Vdd since the start bit is always '1'.

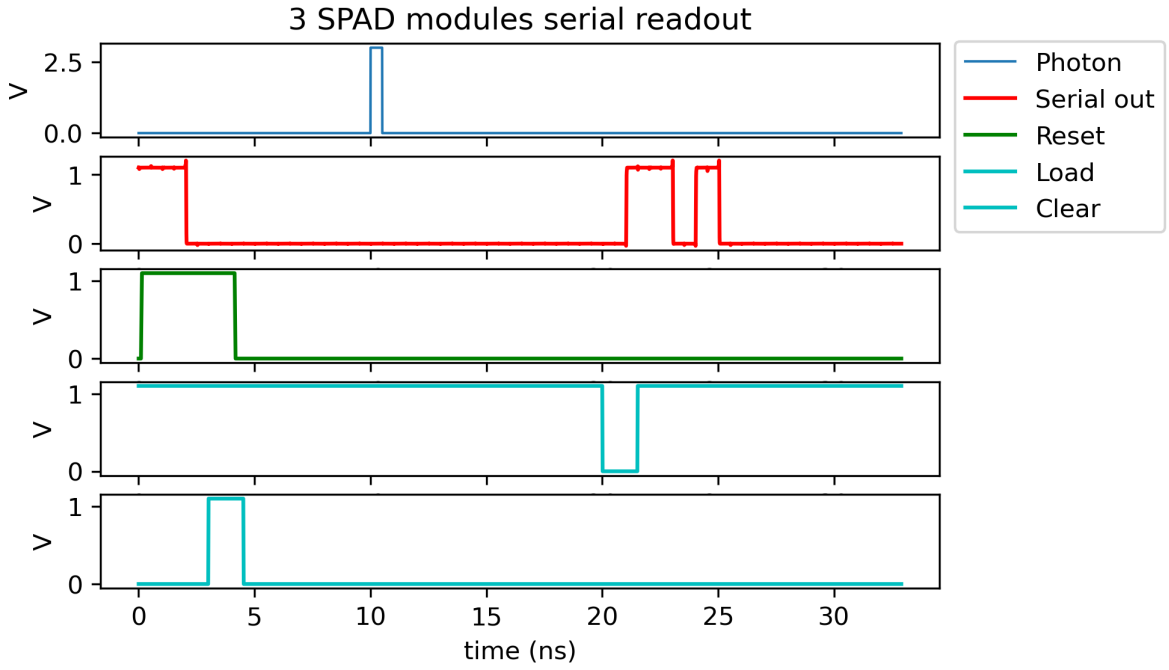


Figure 6.2: Three SPADs, with their AQC, serial readout. In the simulation, SPAD[0] and SPAD[2] received a photon and SPAD[1] did not. Moreover, for FPGA readout, a high start bit is sent.

bits into the correct buffer. The buffers, the JK flipflops not connected to the flipflop chain, will lock in the value when the 'sel' becomes '1'. Signals D0 to D4 will hold the parallelized bits.

Fig. 6.4 holds the SIPO module simulation. First, to make sure the buffer is empty, a reset signal is sent. Afterwards, serial-in communication can commence with a high start bit, then the other configuration bits are sent directly after. In this simulation, 'D0' and 'D1' is set to '1' while the others remain '0'. In the top-level implementation, two 5-bit SIPO modules will be included to set the delay and width of the hold-off circuit's reset pulse.

6.3. Conclusion

This chapter outlines the PISO and SIPO modules necessary for external communication. The interface modules are designed in a way that facilitates communication with FPGA such as utilizing a start, load, clear and reset bit. In the chip, there will be 16 16-bit PISO modules and 2 5-bit SIPO modules. The variation of reading out 16 16-bit bitstreams causes our chip design to meet our speed and IO requirement well within reach.

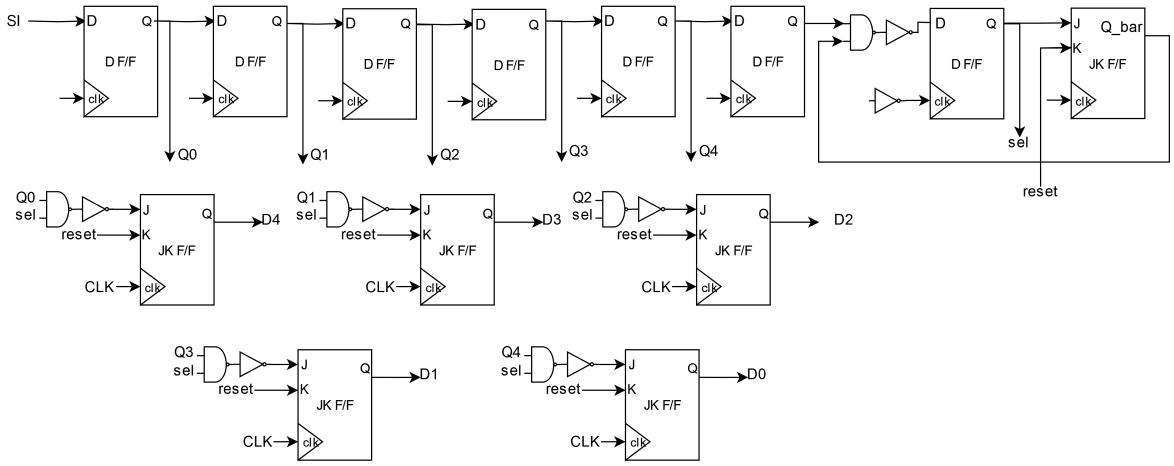


Figure 6.3: Schematic of the 5-bit SIPO. There is a chain of flipflops with at the end a negatively edge-triggered D flipflop and JK flipflop to detect a start bit and lock all the serialized bit correctly in place in its parallelized counter part.

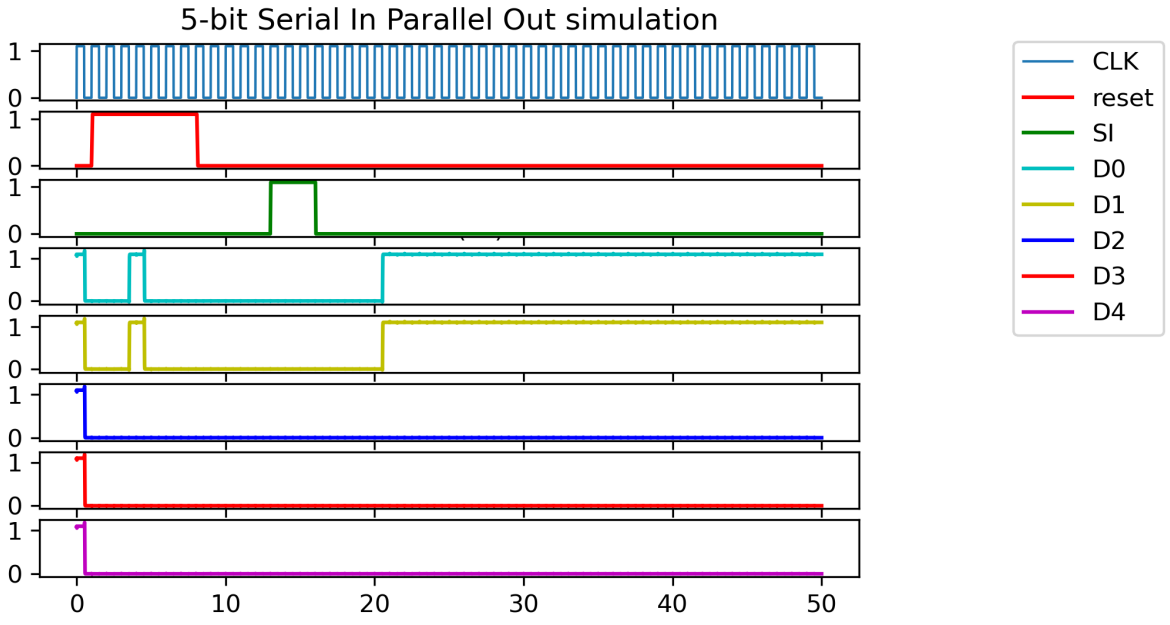


Figure 6.4: Simulation of the SIPO module. Input ‘SI’ sends a high start bit and directly two high bits afterwards so that ‘D0’ and ‘D1’ will store a ‘1’.

7

Top level design

The previous chapters discussed the various components that went into our chip design. In this chapter the distinct parts are assembled together in layout and a validation is done of the entire system. Every layout has been validated via a Design Rule Check (DRC) and Layout Versus Schematic (LVS) check. All the layouts are in the appendix, Chapter A.6. Additionally, the layout's resistances, capacitances and, coupling capacitances are extracted via PEX, and post-layout simulations are performed. Finally a literature based review is done of the expected crosstalk effects in the array.

7.1. Validation

To validate that the SPAD, quenching, and hold-off circuit together work as intended (Figure 3.3), a PEX has been performed on its layout and simulated afterward. The results are in Figure 7.1, there is negligible speed and signal attenuation. Additionally, the reset pulse width and delay are tuned to the SPAD's quenching and recharge time, achieving a total cycle of around 20ns.

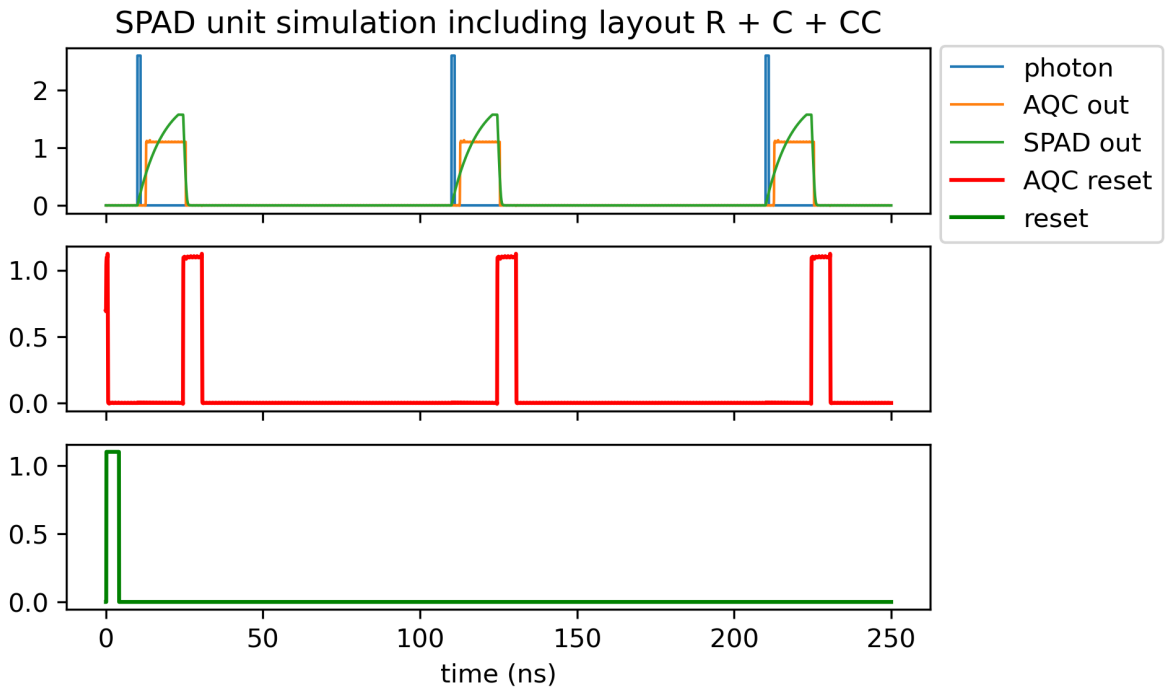


Figure 7.1: Post-layout simulation of the SPAD unit including the extracted resitances, capacitances and, coupling capacitances via PEX. Moreover, the hold-off circuit's reset pulse is tuned to the SPAD's quenching and recharge time

The 16-bit SPAD electronics row consists of 16 SPAD units and SPADs are connected to each other with a bus and 16-bit PISO (Figure 3.2), which is simulated post layout. To confirm that the row implementation works, a PEX operation is performed to extract all the layout's resistances, capacitances, and coupling capacitances and is then simulated to validate its performance. The simulation is in Figure 7.2. The SPAD row's performance has a complete readout cycle of 30 ns, 33 Mcounts/s, from photon to last SPAD avalanche bit, which is well above the required 3 Mcounts/s to be implemented in the chip. Moreover, the resistances, capacitances, and coupling capacitances barely attenuate and slow down the signals. Finally, it correctly produces a bitstream starting with the start bit then the fifteenth SPAD, ending with the first SPAD, the SPAD closest to the PISO module. After every transaction, the FPGA needs to send a '1' over pin 'clear' to empty the PISO buffers for the next 16-bit polling operation.

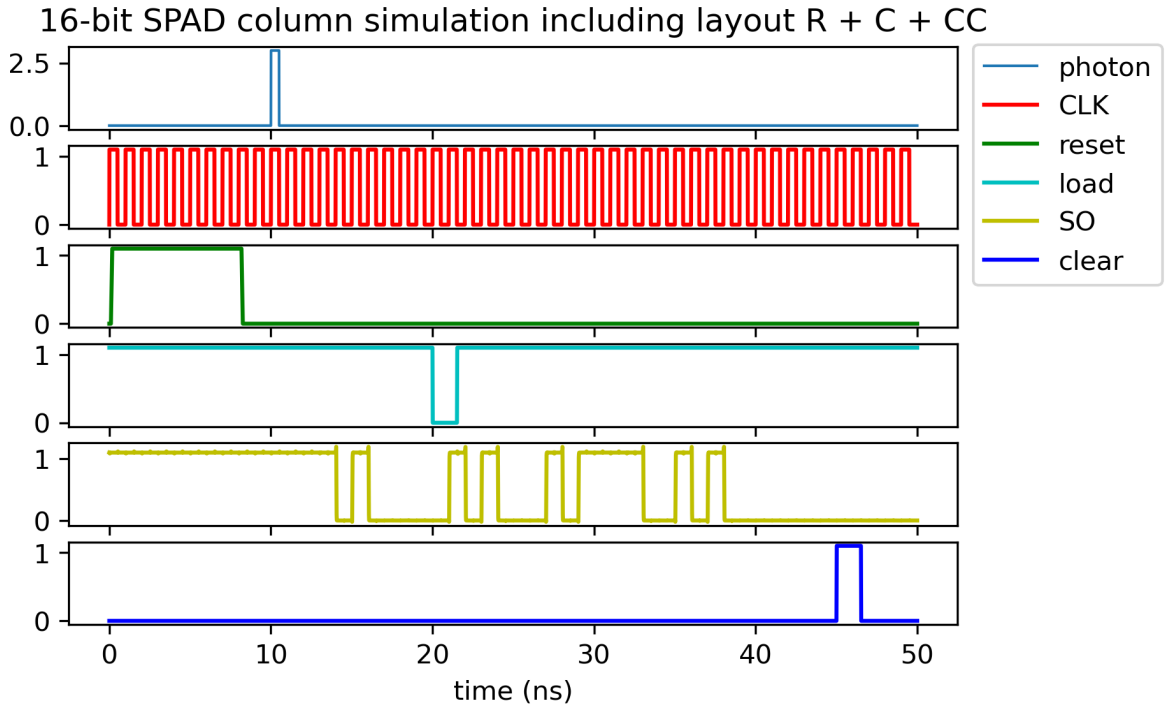


Figure 7.2: The ultimate post-layout simulation of the chip, the 16-bit SPAD column including the layout's resistances, capacitances and coupling capacitances extracted via PEX. 'SO' holds the 16-bit SPAD avalanche bits from a row.

7.2. Shape and size of the SPADs

For the final layout, the round SPAD is used with an active diameter of 12μm. Round SPADs are generally preferred as they are less prone to premature edge breakdown, which is concentrated at the corners of the junction as that is where the electric field is highest [26]. The three SPADs on top of the chip (Figure 3.1), have an active diameter of 12 and 32 μm. One of the 12 μm SPADs is round and the other SPADs are square.

7.3. IO ring

The SPAD characterization group, who worked alongside us, gave us a requirement to detach the IO ring's internal ESD diodes. This is because, during testing, they discovered that the diodes redirect signals above the 3V away from the SPAD so the SPADs can never be biased correctly, since $V_{SPAD,breakdown} > 3V$.

7.4. Crosstalk

In order to give a recommendation for the pixel pitch, it is necessary to obtain an idea about the expected pixel crosstalk. Crosstalk in a SPAD array is a correlated noise effect in one device that are caused

Technology Node	Pixel Pitch (μm)	Crosstalk probability (%)
0.35μm HV CMOS	14μm	0.3% [25]
0.35μm HV CMOS	26μm	2% [8]
0.35μm HV CMOS	32μm	0.11% [38]
40nm CMOS	8.4μm	Not reported [5]
150nm CMOS [51]	11μm	<1%
180nm CMOS	4μm	3.57% [35]

Table 7.1: An overview of crosstalk in literature for various SPAD arrays.

by avalanches in another, neighbouring, device [25] leading to false detection events. Three different mechanisms for crosstalk are considered: direct optical, indirect optical and electrical crosstalk. During an avalanche, photons are emitted by a SPAD as a result of the relaxation of hot carriers generated in response to the current passing through the high electrical field in the breakdown region [25]. These photons travel in direct and indirect, via the substrate, optical paths to neighbouring pixel and can trigger avalanches there leading to optical crosstalk. Crosstalk can also be electrical in nature, where carriers generated during the avalanche may diffuse laterally to depletion regions of neighbouring devices.

7.4.1. A model for optical crosstalk

The crosstalk probability depends on the number of avalanche-generated carriers N_{ava}. The crosstalk probability P_c is proportional to N_{ava} in a first order approximation [34]. Another factor is the distance between pixels, as the emitted photons decay exponentially with distance travelled. The emitter-to-receiver distance dependence of crosstalk can be approximated by [43]

$$P_c = B\frac{e^{-\alpha r}}{r^2}$$

α is the effective decay length of the emitted light, which is determined by the doping concentration of the silicon and the wavelength of the light. r corresponds to the pixel pitch L_p. The coefficient B depends on both emitter and receiver characteristics. It is dependent on the total number of photons emitted, the sensitivity of the receiver and the probability of detecting an emitted photon. The parameter B is an experimentally determined parameter and is unavailable for our device. As such we cannot rely on this mathematical scaling model to approximate our crosstalk probability. A literature based approach is used to determine a suitable pixel pitch.

7.4.2. Literature based recommendation for pixel pitch

In order to make a reasonable estimate for the pixel pitch to limit the crosstalk probability to 1% we've done research on comparable arrays. Such an estimate is necessarily prone to a significant margin of error as technology node, active area, excess bias voltage, SPAD structure, crosstalk mitigation measures, the location of the pixel electronics and well-sharing may all impact cross talk figures. Experimental mapping of the crosstalk probabilities must be performed to gain an accurate insight. Table 7.1 contains various crosstalk figures from literature on which we based our recommendation for the pixel pitch.

Of particular interest are the publications on 40nm CMOS and the paper on 150nm CMOS. The 40nm arrays from the University of Edinburgh do not include crosstalk figures, but operate at significantly lower pixel pitch than we aim to (8 microns versus 24 microns). They have already applied their sensors in FLIM and LIDAR and it may therefore be reasonably assumed that their arrays have acceptable levels of crosstalk. The paper on the array in 150nm CMOS cites crosstalk figures well below 1% for a pitch of 15 μm for a device with an active area of 15.6x15.6μm at 3V excess bias. The 0.35μm HV implementations have pixel pitches in the neighbourhood of our array as well and report crosstalk figures ranging from well below 1% to at most 2%. Moreover E.Charbon published about a 4μm pitch pixel in 180nm CMOS and found a crosstalk probability of 3.57% [35].

7.4.3. Conclusion

From the experimental data from these previous papers, given that our pixels do not share their deep N-well, reducing the crosstalk probability, we feel reasonably confident that a 16x16 SPAD array with

a pixel pitch of 24 μm has a good chance of meeting the stated requirement of the sub 2% crosstalk probability for nearest neighbouring pixels.

8

Discussion

In this thesis, a layout is designed for a 16x16 SPAD array, including quenching, recharge, and interface electronics.
Unfortunately, the existing 4x5 SPAD chip could not be characterized by the SPAD characterization group, working alongside us, because the ESD diodes in the guard ring were designed to protect the chip from voltages over 3V. However, the breakdown voltage of the SPADs is higher than that. The IO ring is thus incompatible with the chip requirements and the SPADs could not be tested in their breakdown region. This has several consequences for us.
The fact that the I-V behaviour and the capacitance of the SPADs are unknown, as they could not be measured, means that the simulations for quenching and recharge time in this work can at best provide an indication of the eventual SPAD behaviour as the SPAD model wasn't able to be experimentally fit to measurement data. Reasonably conservative values of 1-2pF, as given by Zappa in [54] for the parasitic SPAD capacitances, are used and a fairly conservative 5.4kΩ resistance for the diode in the breakdown region. This leads to a larger intrinsic time constant for the SPAD, which is most obvious in the quenching time of approximately 20ns for the AQC. If the practical capacitances are smaller, that will affect a speed up in the quenching and recharge times as well. The hold-off circuit has a tunable delay and pulse width which should be able to accommodate reasonable values for the quenching and recharge time.
Secondly, as the SPADs could not be characterized, the characterization group could not come to a recommendation for the optimal SPAD for the final array. We chose a round SPAD with a 12 μ_m active area for the final design. This was primarily based on size restrictions, which made fitting a bigger SPAD on the chip a challenge in a 16x16 array and literature research that indicated round SPADs have fewer issues with premature edge breakdown of the active area. However, a qualitative analysis of this effect could not be done and it is not known whether another SPAD might have more desirable properties.
The hold-off circuit has a pulse resolution of 1 ns which is suitable for our speed requirements. A faster clock with a larger bit counter could have been selected but that would add too much area, reserved for the SPAD array.
The IO circuits, the 16-bit PISO and the 5-bit SIPO, all have additional flipflop implemented for start bits and holding onto states. This route is chosen to facilitate the programming at the FPGA side for the QIT lab. The chip driving requirements is included in Appendix A.5.
The layout for all the designed electronics is validated with the DRC, LVS and post-layout simulations. The top hierarchy validated with the DRC and LVS is at the 16x16 SPAD array and the top hierarchy that is validated with post-layout simulation is until the 16-bit SPAD row. All the checks and simulations are as expected and correct which leads us to believe that the chip is ready for the dry-run and tape out.

9

Conclusion and recommendation

To summarize, the designed and validated chip, with an area of $1.1x1.1mm^2$, which houses a 16x16 SPAD array including active quenching circuits, recharge and hold-off circuits, PISO, and SIPO circuits. Therefore, we have met all the mandatory requirements which were:

- Photon detection rate above the 3Mc/s, or every 333 ns. The AQC and digital hold-off circuit are tuned to reach a theoretical count rate in excess of 76Mc/s for the modeled SPAD.
- SPAD Avalanche quenched in less than 20 ns. With our current AQC sizing, the avalanche is quenched in 12ns.
- Controllable dead time. After the quenching, the SPAD is inoperable, the hold-off circuit's delay and pulse width can be configured to determine when the SPAD is armed again thus the dead time is controllable.
- Crosstalk probability less than 2%. Our current pixel pitch is 24 μm and contains no shared N-well for pixel isolation. Therefore, looking at the literature, there is evidence to suggest that the crosstalk will be less than 2%.
- 16x16 SPAD array. Our current 16x16 SPAD array houses the round variant with an active area diameter of 12μm
- A chip area of $1.1x1.1mm^2$. The entire SPAD array, electronics, and IO ring fit on an area of $1.1x1.1mm^2$.
- IO of 32 pins. The chip houses 16 16-bit PISOs and 2 5-bit SIPOs to accommodate for the IO limitation of 32 pins.
- 40 nm TSMC process. The chip layout is designed for TSMC 40nm CMOS technology.

Moreover, the chip's layout has been validated with a Design Rule Check, Layout Versus Schematic Check, and post-layout simulations.

9.1. Future work

Some future work, which is beyond the scope of the thesis, is redesigning the SPAD for bigger round active areas and experimental metal grating tuned to NV centers. Additionally, three additional 16x16 SPAD arrays including electronics will be added to tape-out to reach an eventual chip area of $2x2mm^2$. Those SPADs will need to be characterized as well.
When the chip is fabricated and bond wired, VHDL or Verilog code for FPGA, running on a 1 GHz clock, needs to be developed to drive and control the chip. Please contact the authors for assistance when the time comes.
Further studies could be done into optimal array design. A design that incorporates deep N-well sharing and houses the pixel electronics separately from the SPADs may achieve a significant increase in fill-factor as evidenced by the 96x40array in 40nm CMOS on the same 1x1mm area in [5]. In order to

mitigate increased crosstalk effects from this well-sharing and decreased pixel pitch the guard ring structure and pixel isolation would need to be re-examined. Another consideration is the use of metal layers on top of the inactive areas to mitigate unwanted avalanche triggers by photons traveling through the device structure to the active areas as done in [30]. Another area of interest is using Time-to-Digital (TDC) converters to precisely mark the onset of avalanche pulses. Readout circuits based on TDCs have already been discussed in SPAD array literature and could greatly increase the temporal resolution of the photon detection [8][24][5]. This would require a ground-up redesign of the readout and communication electronics.

A

Appendix

A.1. IV characteristic model

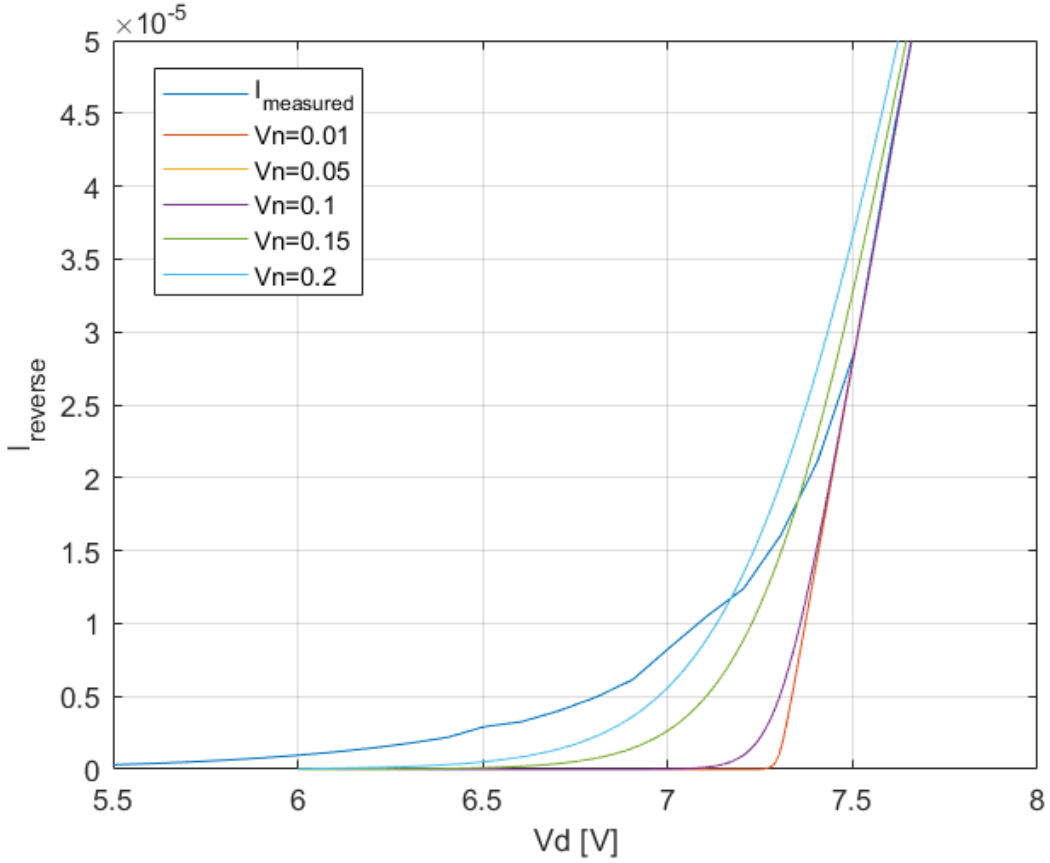


Figure A.1: The IV measurements vs the 2011 model of Mita [18] with different values for Vn.

A.2. Verilog-A model

```
// VerilogA for SPAD_lib, SPADv3, veriloga



`include "constants.vams"
`include "disciplines.vams"



module SPADv3 (a,c,photon);



    inout a,c,photon;
    electrical a,c,photon,gnd;
    ground gnd;
    branch (c,gnd) Ccap;
    branch (a,gnd) Acap;
    branch (c,a) spad,cap;



    parameter real VB = 7.41994; // Breakdown voltage of the SPAD
    parameter real Rbreak = 5650.419; // Series resistance of SPAD in breakdown region
    parameter real Ilatchup=100e-6; // Latch current, below which the avalanche is quenched.
    parameter real Is = 80e-12; // Reverse current
    parameter real Cjo = 1e-12; // Used to model junction capacitance
    parameter real Cjop1=2e-12; // Junction - Cathode capacitance
    parameter real Cjop2=2e-12; // Junction - Anode capacitance
    parameter real phi=0.65; // Built in voltage
    parameter real m = 0.5; // Junction grading coefficient from 0.5 abrupt junctions - 1/3
    for graded junctions
    parameter real n = 1e-5;
    parameter real Vn=0.1; // Normalization voltage used to fit I-V curve to experimental
    curve



    real Ibreak,INL;
    real Vd,Vex;
    real Cjc,Cja,Cjd;
    real Qjc,Qja,Qj;
    real avalanche;
    real Vac_max, Vac_min;



analog begin
    @(initial_step)
    begin
        Vac_max = 16;
        Vac_min=VB;
        Cjc = Cjop1;
        Cja = Cjop2;
        Cjd = Cjo*(pow(phi,m))*(pow(phi-Vac_min,1-m)-pow(phi-Vac_max,1-m))/(1-m)/(Vac_max-
    Vac_min);
        avalanche=0.0;
    end

    Vd = V(c)-V(a);
    Vex = Vd - VB;



if (V(photon)>2.5) begin // 2.5 is an arbitrary threshold for testing purposes.
        avalanche = 1.0;
    end


```

```
67
68      Ibreak = (Vn/Rbreak)*ln(1 + limexp(Vex/Vn));
69      INL = Is + avalanche*Ibreak;
70
71
72
73      if(INL < Ilatchup) begin
74          avalanche = 0.0;
75      end
76
77
78
79      Qj = Cjd*Vd;
80      Qjc = Cjc*V(Ccap);
81      Qja = Cja*V(Acap);
82
83
84
85      I(Acap) <+ ddt(Qja);
86      I(Ccap) <+ ddt(Qjc);
87      I(cap) <+ ddt(Qj);
88      I(spad) <+ INL;
89
90
91
92  end
93  endmodule
```

veriloga.va

A.3. Analog hold-off circuit

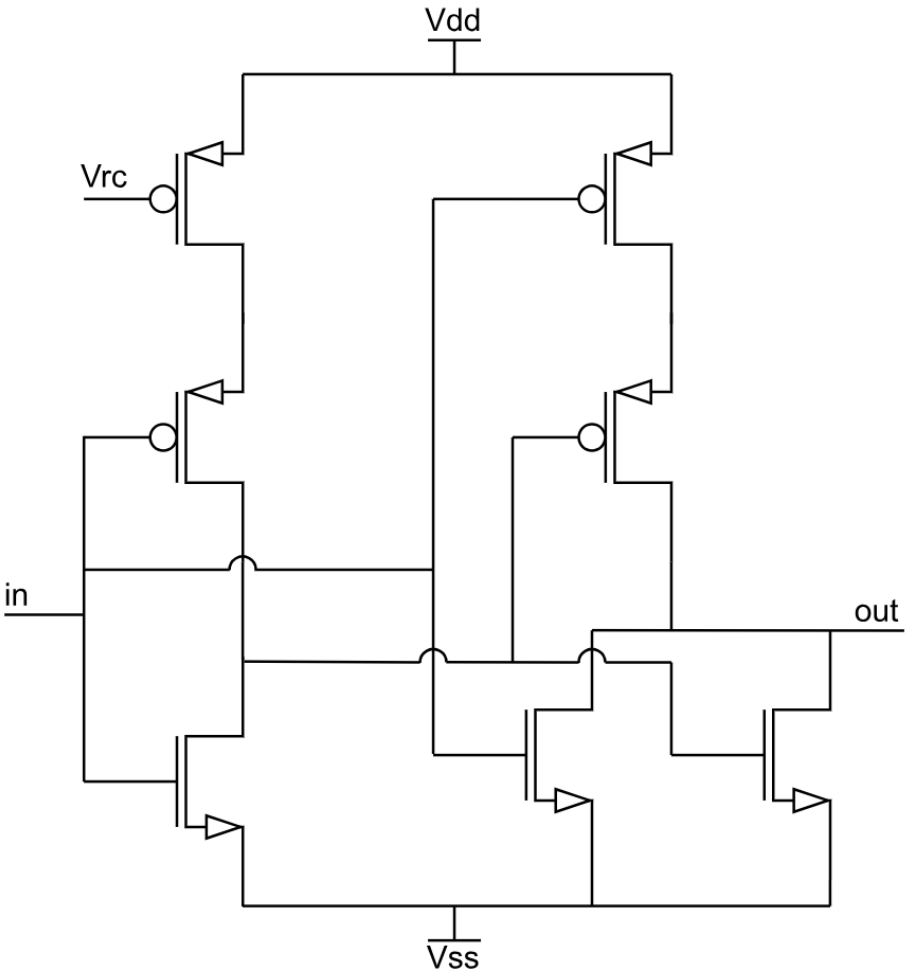


Figure A.2: The pulse width part of the analog hold-off circuit.

A.4. 5-bit counter post layout simulation

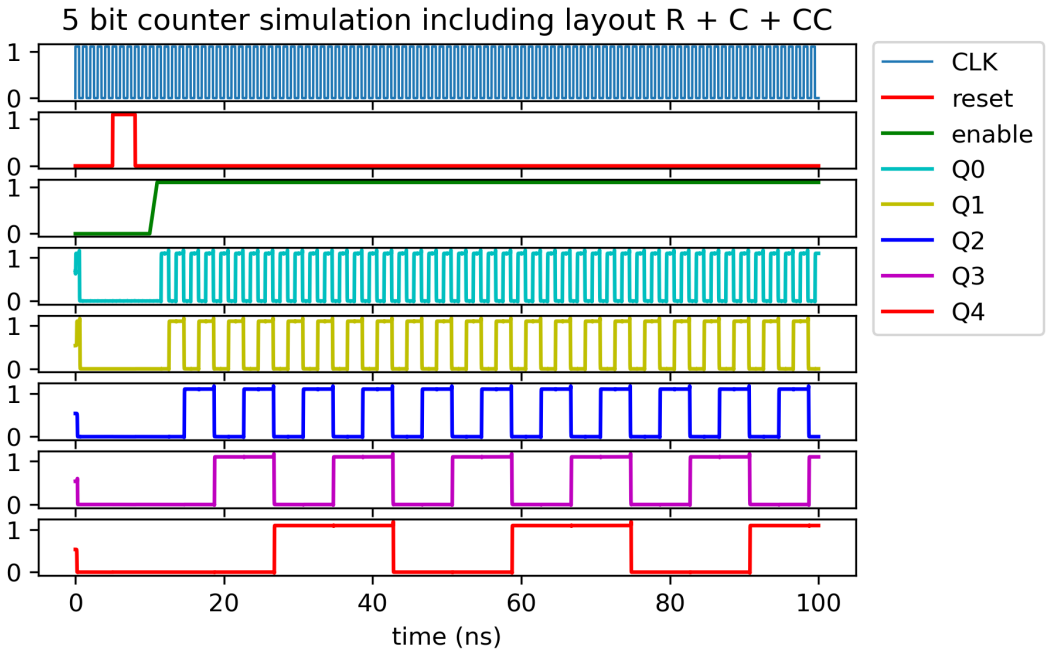


Figure A.3: The 5-bit counter's post-layout simulation including resistances, capacitances, and coupling capacitances.

A.5. Chip driving requirements and advice

- Vdd = 1.1 V
- Vss = 0 V
- Clock frequency = 1 GHz
- Reset pulse > 7 ns
- Load pulse > 1.1 ns
- Clear pulse > 1.1 ns

Reading out pins SO0 to SO15, to acquire the SPAD array avalanche status bits:

- Send a Load pulse
- At negative clock edge trigger, detect the start bits and start counting until 16 SPAD avalanche status bits have passed. The order of SPAD bits is S15, at the end of the row, to S0, nearest to the PISO module.
- Send a clear pulse
- Restart the cycle within a set time window

Sending in the 5-bit delay and width configure bits through SI_w and SI_d:

- Send reset pulse, to clear the buffers
- Send in 6-bit bitstream, starting with the start bit, then D0 to D4.

A.6. Layouts

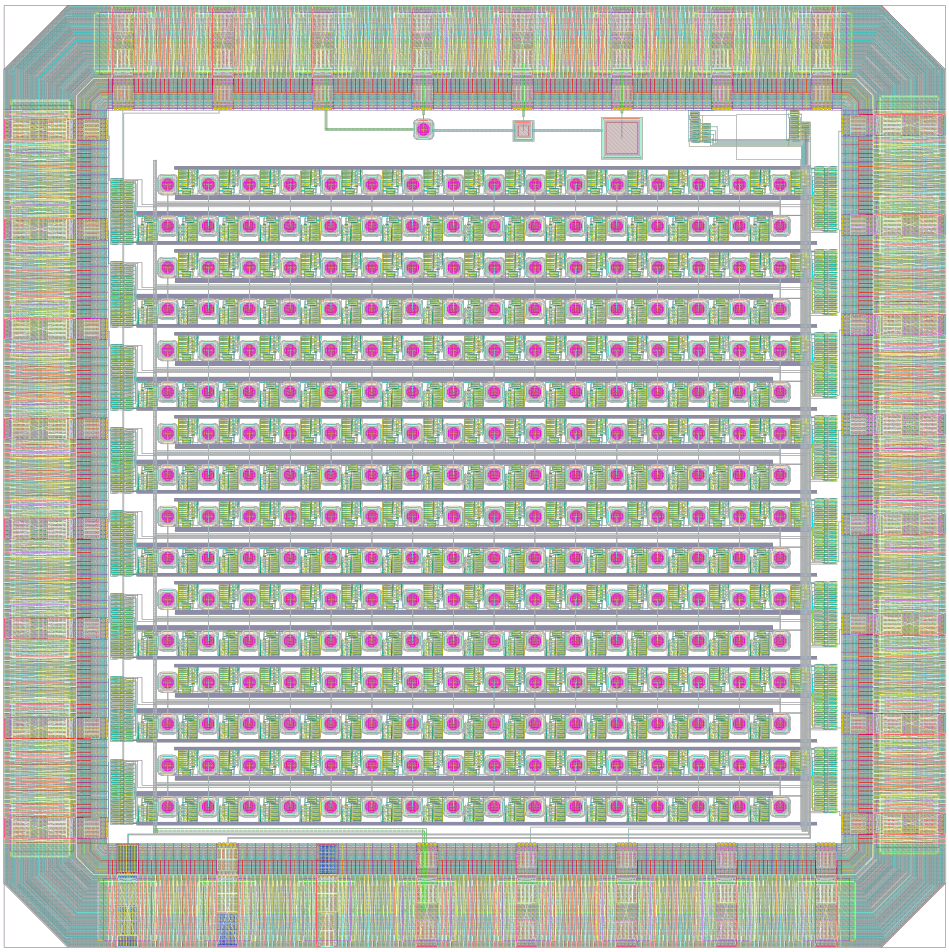


Figure A.4: The complete layout of the SPAD chip.

Figure A.5: SPAD column layout including 16 SPAD unit electronics and 16-bit PISO.

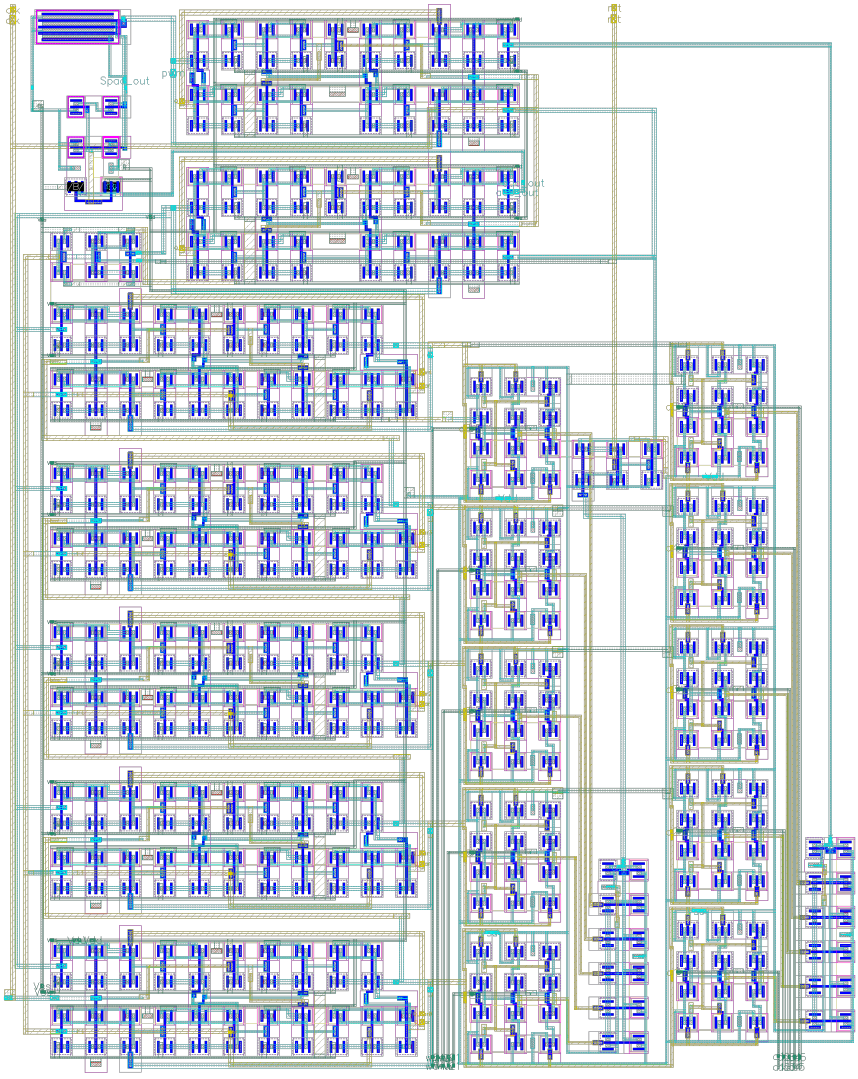


Figure A.6: SPAD unit layout including the AQC, two JK flipflops, 5-bit counter, and two comparators.

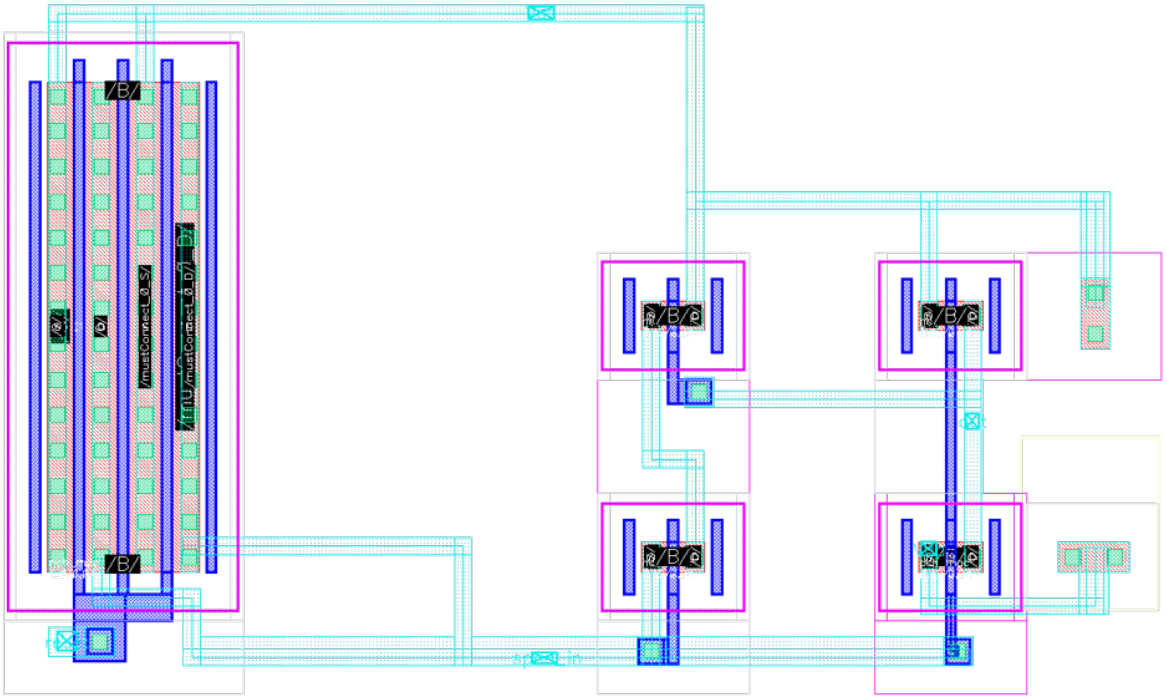


Figure A.7: Active Quenching Circuit layout.

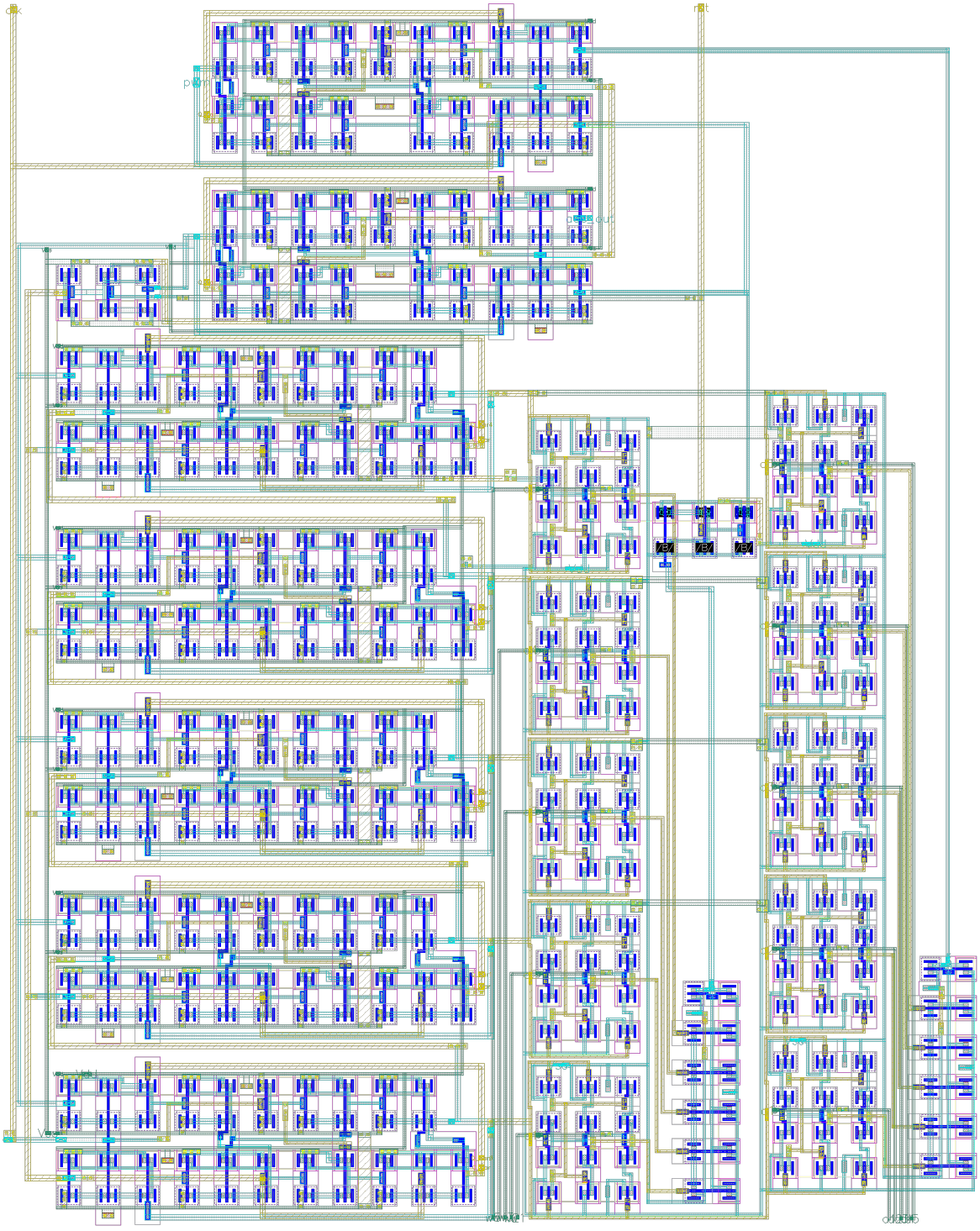


Figure A.8: Digital hold-off circuit layout.

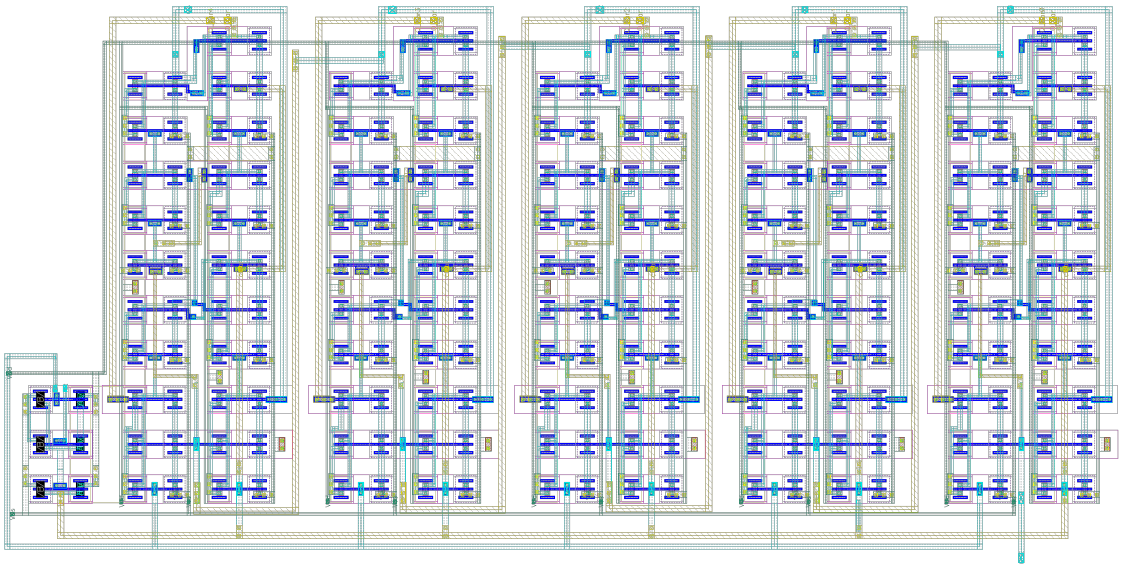


Figure A.9: 5-bit counter layout.

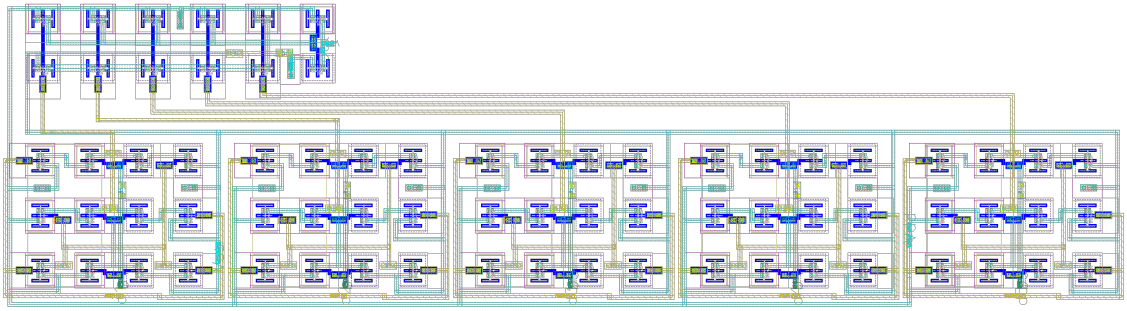


Figure A.10: 5-bit comparator layout.

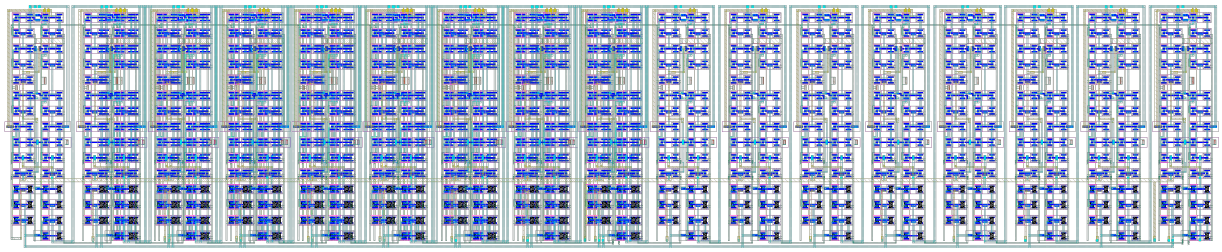


Figure A.11: 16-bit PISO layout.

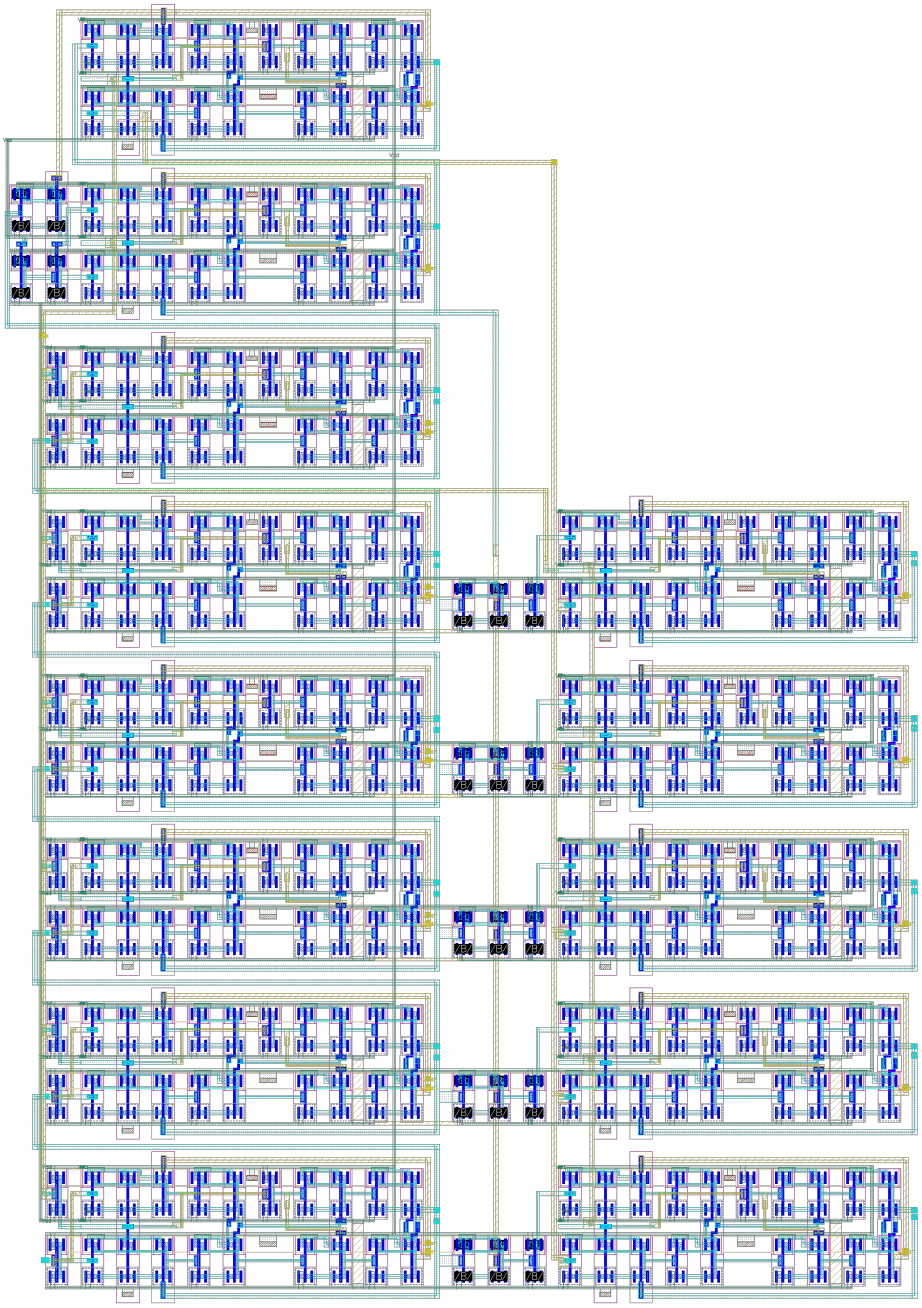


Figure A.12: 5-bit SIPO layout.

Bibliography

[1] Claudio ”Bruschini et al. “”Single-photon avalanche diode imagers in biophotonics: review and outlook””. ”en”. In: *”Light Sci. Appl.”* 8.1 (Sept. 2019), ”87”.

[2] Bilal I. Abdulrazzaq et al. *A review on high-resolution CMOS delay lines: Towards sub-picosecond Jitter Performance - SpringerPlus*. Apr. 2016. URL: https://doi.org/10.1186/s40064-016-2090-z.

[3] G. Acconcia et al. “Fast fully integrated active quenching circuit for single photon counting up to 160 Mcounts/s”. In: *Advanced Photon Counting Techniques XIII*. Ed. by Mark A. Itzler, Joshua C. Bienfang, and K. Alex McIntosh. Vol. 10978. International Society for Optics and Photonics. SPIE, 2019, p. 109780D. DOI: 10.1117/12.2518883. URL: https://doi.org/10.1117/12.2518883.

[4] G. Acconcia et al. “Note: Fully integrated active quenching circuit achieving 100 MHz count rate with custom technology single photon avalanche diodes”. In: *Review of Scientific Instruments* 88.2 (Feb. 2017). 026103. ISSN: 0034-6748. DOI: 10.1063/1.4975598. eprint: https://pubs.aip.org/aip/rsi/article-pdf/doi/10.1063/1.4975598/13576563/026103_1_online.pdf. URL: https://doi.org/10.1063/1.4975598.

[5] Tarek Al abbas et al. “8.25µm Pitch 66% Fill Factor Global Shared Well SPAD Image Sensor in 40nm CMOS FSI Technology”. English. In: International Image Sensor Workshop 2017, IISW 2017 ; Conference date: 30-05-2017 Through 02-06-2017. June 2017.

[6] T. Albuquerque et al. “Integration of SPAD in 28nm FDSOI CMOS technology”. In: Sept. 2018, pp. 82–85. DOI: 10.1109/ESSDERC.2018.8486852.

[7] A. Batalov et al. “Temporal Coherence of Photons Emitted by Single Nitrogen-Vacancy Defect Centers in Diamond Using Optical Rabi-Oscillations”. In: *Phys. Rev. Lett.* 100 (7 Feb. 2008), p. 077401. DOI: 10.1103/PhysRevLett.100.077401. URL: https://link.aps.org/doi/10.1103/PhysRevLett.100.077401.

[8] Michele Benetti et al. “Characterization of a CMOS SPAD sensor designed for fluorescence lifetime spectroscopy”. In: *2011 7th Conference on Ph.D. Research in Microelectronics and Electronics*. 2011, pp. 185–188. DOI: 10.1109/PRIME.2011.5966248.

[9] Danilo Bronzi et al. “Large-area CMOS SPADs with very low dark counting rate”. In: *Proceedings of SPIE - The International Society for Optical Engineering* 8631 (Feb. 2013). DOI: 10.1117/12.2004209.

[10] S Cova et al. “Avalanche photodiodes and quenching circuits for single-photon detection”. In: *Applied optics* 35 (Apr. 1996), pp. 1956–76. DOI: 10.1364/AO.35.001956.

[11] S. Cova, A. Longoni, and G. Ripamonti. “Active-Quenching and Gating Circuits for Single-Photon Avalanche Diodes (SPADs)”. In: *IEEE Transactions on Nuclear Science* 29.1 (1982), pp. 599–601. DOI: 10.1109/TNS.1982.4335917.

[12] Stephen J DeVience et al. “Nanoscale NMR spectroscopy and imaging of multiple nuclear species”. en. In: *Nat. Nanotechnol.* 10.2 (Feb. 2015), pp. 129–134.

[13] Mohammadreza Dolatpoor Lakeh et al. “An Ultrafast Active Quenching Active Reset Circuit with 50% SPAD Afterpulsing Reduction in a 28 nm FD-SOI CMOS Technology Using Body Biasing Technique”. In: *Sensors* 21.12 (2021). ISSN: 1424-8220. DOI: 10.3390/s21124014. URL: https://www.mdpi.com/1424-8220/21/12/4014.

[14] Matthew W. Fishburn. “Fundamentals of CMOS Single-Photon Avalanche Diodes”. In: 2012.

[15] “Fundamentals of magnetic field measurement with NV centers in diamond”. In: *Qnami* (Dec. 2020). URL: https://qnami.ch/wp-content/uploads/2020/12/2020-12-07-Qnami-TN1-The-NV-center-1.pdf.

[16] M. Ghioni et al. “Compact active quenching circuit for fast photon counting with avalanche photodiodes”. In: *Review of Scientific Instruments* 67.10 (Oct. 1996), pp. 3440–3448. ISSN: 0034-6748. DOI: 10.1063/1.1147156. eprint: https://pubs.aip.org/aip/rsi/article-pdf/67/10/3440/8810386/3440_1_online.pdf. URL: https://doi.org/10.1063/1.1147156.

[17] Andrea Giudici et al. “Fast fully-integrated active quenching circuit driving an external SPAD up to 250Mcounts/s”. In: *Advanced Photon Counting Techniques XVI*. Ed. by Mark A. Itzler, Joshua C. Bienfang, and K. Alex McIntosh. Vol. PC12089. International Society for Optics and Photonics. SPIE, 2022, PC1208902. DOI: 10.1117/12.2618500. URL: https://doi.org/10.1117/12.2618500.

[18] G. Giustolisi, R. Mita, and G. Palumbo. “Behavioral modeling of statistical phenomena of single-photon avalanche diodes”. In: *International Journal of Circuit Theory and Applications* 40.7 (2012), pp. 661–679. DOI: https://doi.org/10.1002/cta.748. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cta.748. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cta.748.

[19] David Glenn et al. “Single-cell magnetic imaging using a quantum diamond microscope”. In: *Nature Methods* 12 (June 2015), pp. 736–738. DOI: 10.1038/nmeth.3449. (Visited on 04/28/2023).

[20] Tjerk P. Gutteling et al. “A New Generation of OPM for High Dynamic and Large Bandwidth MEG: The 4He OPMs—First Applications in Healthy Volunteers”. In: *Sensors* 23.5 (Mar. 2023), p. 2801. ISSN: 1424-8220. DOI: 10.3390/s23052801. URL: http://dx.doi.org/10.3390/s23052801.

[21] L T Hall et al. “Detection of nanoscale electron spin resonance spectra demonstrated using nitrogen-vacancy centre probes in diamond”. en. In: *Nat. Commun.* 7.1 (Jan. 2016), p. 10211.

[22] Riitta Hari and Riitta Salmelin. “Magnetoencephalography: From SQUIDs to neuroscience: Neuroimage 20th Anniversary Special Edition”. In: *NeuroImage* 61.2 (2012). NEUROIMAGING: THEN, NOW AND THE FUTURE, pp. 386–396. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2011.11.074. URL: https://www.sciencedirect.com/science/article/pii/S1053811911013565.

[23] Mutsuko Hatano. “Potential of diamond solid-state quantum sensors”. In: *2022 6th IEEE Electron Devices Technology Manufacturing Conference (EDTM)*. 2022, pp. 42–43. DOI: 10.1109/EDTM53872.2022.9797922.

[24] Robert K. Henderson et al. “A 192×128 Time Correlated SPAD Image Sensor in 40-nm CMOS Technology”. In: *IEEE Journal of Solid-State Circuits* 54.7 (2019), pp. 1907–1916. DOI: 10.1109/JSSC.2019.2905163.

[25] Sahba Jahromi and Juha Kostamovaara. “Timing and probability of crosstalk in a dense CMOS SPAD array in pulsed TOF applications”. In: *Opt. Express* 26.16 (Aug. 2018), pp. 20622–20632. DOI: 10.1364/OE.26.020622. URL: https://opg.optica.org/oe/abstract.cfm?URI=oe-26-16-20622.

[26] Ehsan Kamrani, Frédéric Lesage, and Mohamad Sawan. “Premature edge breakdown prevention techniques in CMOS APD fabrication”. In: June 2012, pp. 345–348. ISBN: 978-1-4673-0857-1. DOI: 10.1109/NEWCAS.2012.6329027.

[27] Yan Kang et al. “High-resolution depth imaging with a small-scale SPAD array based on the temporal-spatial filter and intensity image guidance”. In: *Opt. Express* 30.19 (Sept. 2022), pp. 33994–34011. DOI: 10.1364/OE.459787. URL: https://opg.optica.org/oe/abstract.cfm?URI=oe-30-19-33994.

[28] Mürsel Karadas et al. *Opto-magnetic imaging of neural network activity in brain slices at high resolution using color centers in diamond*. 2017. arXiv: 1708.06158 [q-bio.NC].

[29] Mohammad Azim Karami and Misagh Ansarian. *Neural imaging using single-photon avalanche diodes*. Jan. 2017. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5396169/.

[30] Mohammad Azim Karami et al. “A New Single-Photon Avalanche Diode in 90nm Standard CMOS Technology”. In: *Optics express* 18 (Oct. 2010), pp. 22158–66. DOI: 10.1117/12.859435.

[31] I Lovchinsky et al. “Nuclear magnetic resonance detection and spectroscopy of single proteins using quantum logic”. en. In: *Science* 351.6275 (Feb. 2016), pp. 836–841.

[32] Rosario Mita and Gaetano Palumbo. “High-Speed and Compact Quenching Circuit for Single-Photon Avalanche Diodes”. In: *IEEE Transactions on Instrumentation and Measurement* 57.3 (2008), pp. 543–547. DOI: 10.1109/TIM.2007.911691.

[33] Rosario Mita, Gaetano Palumbo, and G. Fallica. “Accurate model for single-photon avalanche diodes”. In: *Circuits, Devices Systems, IET* 2 (May 2008), pp. 207–212. DOI: 10.1049/iet-cds:20070180.

[34] Kazuhiro Morimoto and Edoardo Charbon. “A Scaling Law for SPAD Pixel Miniaturization”. In: *Sensors* 21.10 (May 2021), p. 3447. ISSN: 1424-8220. DOI: 10.3390/s21103447. URL: http://dx.doi.org/10.3390/s21103447.

[35] Kazuhiro Morimoto and Edoardo Charbon. “High fill-factor miniaturized SPAD arrays with a guard-ring-sharing technique”. In: *Opt. Express* 28.9 (Apr. 2020), pp. 13068–13080. DOI: 10.1364/OE.389216. URL: https://opg.optica.org/oe/abstract.cfm?URI=oe-28-9-13068.

[36] Kazuhiro Morimoto et al. “Megapixel time-gated SPAD image sensor for 2D and 3D imaging applications”. In: *Optica* 7.4 (Apr. 2020), pp. 346–354. DOI: 10.1364/OPTICA.386574. URL: https://opg.optica.org/optica/abstract.cfm?URI=optica-7-4-346.

[37] Dmitry Murzin et al. “Ultrasensitive magnetic field sensors for biomedical applications”. en. In: *Sensors (Basel)* 20.6 (Mar. 2020), p. 1569.

[38] I. Nissinen et al. “Cross talk measurements of a time-gated 4×128 SPAD array for pulsed Raman spectroscopy”. In: *2014 NORCHIP*. 2014, pp. 1–4. DOI: 10.1109/NORCHIP.2014.7004711.

[39] S. Pellegrini et al. “Industrialised SPAD in 40 nm technology”. In: *2017 IEEE International Electron Devices Meeting (IEDM)*. 2017, pp. 16.5.1–16.5.4. DOI: 10.1109/IEDM.2017.8268404.

[40] Matteo Perenzoni, Lucio Pancheri, and David Stoppa. “Compact SPAD-Based Pixel Architectures for Time-Resolved Image Sensors”. In: *Sensors* 16.5 (2016). ISSN: 1424-8220. DOI: 10.3390/s16050745. URL: https://www.mdpi.com/1424-8220/16/5/745.

[41] Giulia Petrini et al. “Is a Quantum Biosensing Revolution Approaching? Perspectives in NV-Assisted Current and Thermal Biosensing in Living Cells”. In: *Advanced Quantum Technologies* 3.12 (2020), p. 2000066. DOI: https://doi.org/10.1002/qute.202000066. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qute.202000066. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/qute.202000066.

[42] *Quantum Integration Technology*. TU Delft. URL: https://www.tudelft.nl/en/eemcs/the-faculty/departments/quantum-computer-engineering/sections/quantum-circuits-architectures-and-technology/groups/quantum-integration-technology (visited on 06/14/2023).

[43] Ivan Rech et al. “Optical crosstalk in single photon avalanche diode arrays: a new complete model”. In: *Opt. Express* 16.12 (June 2008), pp. 8381–8394. DOI: 10.1364/OE.16.008381. URL: https://opg.optica.org/oe/abstract.cfm?URI=oe-16-12-8381.

[44] “Sequential Circuit Design”. In: *Digital VLSI Design and Simulation with Verilog*. John Wiley Sons, Ltd, 2021. Chap. 2, pp. 25–42. ISBN: 9781119778097. DOI: https://doi.org/10.1002/9781119778097.ch2. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781119778097.ch2. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119778097.ch2.

[45] Mario Stipčević et al. “Advanced active quenching circuit for ultra-fast quantum cryptography”. In: *Opt. Express* 25.18 (Sept. 2017), pp. 21861–21876. DOI: 10.1364/OE.25.021861. URL: https://opg.optica.org/oe/abstract.cfm?URI=oe-25-18-21861.

[46] P. Sun. “Flexible CMOS Single-Photon Avalanche Diode Image Sensor Technology”. In: (2016). DOI: https://doi.org/10.4233/uuid:11717f7d-51c9-471b-8f9e-ee1a70e7f032. URL: https://repository.tudelft.nl/islandora/object/uuid:11717f7d-51c9-471b-8f9e-ee1a70e7f032?collection=research.

[47] Chao Tang et al. “A Hybrid Quenching and Active Reset Circuit for High-Speed SPAD Detectors”. In: *2022 7th International Conference on Integrated Circuits and Microsystems (ICICM)*. 2022, pp. 76–79. DOI: 10.1109/ICICM56102.2022.10011317.

[48] S. Tisa et al. “Electronics for single photon avalanche diode arrays”. In: *Sensors and Actuators A: Physical* 140.1 (2007), pp. 113–122. ISSN: 0924-4247. DOI: https://doi.org/10.1016/j.sna.2007.06.022. URL: https://www.sciencedirect.com/science/article/pii/S0924424707004979.

[49] Federica Villa et al. “SPADs and SiPMs Arrays for Long-Range High-Speed Light Detection and Ranging (LiDAR)”. In: *Sensors* 21.11 (June 2021), p. 3839. ISSN: 1424-8220. DOI: 10.3390/s21113839. URL: http://dx.doi.org/10.3390/s21113839.

[50] Federica Villa et al. “SPADs and SiPMs Arrays for Long-Range High-Speed Light Detection and Ranging (LiDAR)”. In: *Sensors* 21.11 (June 2021), p. 3839. ISSN: 1424-8220. DOI: 10.3390/s21113839. URL: http://dx.doi.org/10.3390/s21113839.

[51] Hesong Xu et al. “Crosstalk Characterization of Single-photon Avalanche Diode (SPAD) Arrays in CMOS 150nm Technology”. In: *Procedia Engineering* 87 (2014). EUROSENSORS 2014, the 28th European Conference on Solid-State Transducers, pp. 1270–1273. ISSN: 1877-7058. DOI: https://doi.org/10.1016/j.proeng.2014.11.417. URL: https://www.sciencedirect.com/science/article/pii/S1877705814025326.

[52] F. Zappa et al. “Monolithic active-quenching and active-reset circuit for single-photon avalanche detectors”. In: *IEEE Journal of Solid-State Circuits* 38.7 (2003), pp. 1298–1301. DOI: 10.1109/JSSC.2003.813291.

[53] F. Zappa et al. “Principles and features of single-photon avalanche diode arrays”. In: *Sensors and Actuators A: Physical* 140.1 (2007), pp. 103–112. ISSN: 0924-4247. DOI: https://doi.org/10.1016/j.sna.2007.06.021. URL: https://www.sciencedirect.com/science/article/pii/S0924424707004967.

[54] F. Zappa et al. “SPICE modeling of single photon avalanche diodes”. In: *Sensors and Actuators A: Physical* 153.2 (2009), pp. 197–204. ISSN: 0924-4247. DOI: https://doi.org/10.1016/j.sna.2009.05.007. URL: https://www.sciencedirect.com/science/article/pii/S0924424709002581.

[55] Chen Zhang et al. “Optimizing NV magnetometry for Magnetoneurography and Magnetomyography applications”. In: *Frontiers in Neuroscience* 16 (2023). ISSN: 1662-453X. DOI: 10.3389/fnins.2022.1034391. URL: https://www.frontiersin.org/articles/10.3389/fnins.2022.1034391.

[56] Jixing Zhang et al. *Diamond Nitrogen-Vacancy Center Magnetometry: Advances and Challenges*. 2020. DOI: 10.48550/arXiv.2010.10231. arXiv: 2010.10231.

[57] Ning Zhang et al. “Magnetic Sensing of Magnetization in Magnetotactic Bacteria with Nitrogen Vacancy Centers in Dimond”. In: (Sept. 2017), pp. 18–21. DOI: 10.1109/ES.2017.10. URL: https://ieeexplore.ieee.org/document/8119321 (visited on 06/14/2023).