UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE Mathématiques, Science de l'information et de l'ingénieur

ICube UMR 7357

THÈSE présentée par :

Mohammadreza DOLATPOOR LAKEH

soutenue le : **10 Octobre 2023**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline : Electronique, Microélectronique, photonique

Spécialité : Micro et Nanoélectronique

Design of a Single Photon Sensor in a 28 nm FD-SOI CMOS technology

THÈSE dirigée par :

M KAMMERER Jean-Baptiste	Docteur HDR, université de Strasbourg, France
M UHRING Wilfried	Professeur, université de Strasbourg, France

RAPPORTEURS :

M BRUSCHINI Claudio	Docteur HDR, EPFL, Suisse
M MAGNAN Pierre	Professeur, ISAE SUPAERO, Toulouse, France

AUTRES MEMBRES DU JURY :

Mme Cathelin Andreia	Docteur HDR, STMicroelectronics, France
M Calmon Francis	Professeur, INSA Lyon, France
M Dartigues Alexandre (invité)	Docteur, STMicroelectronics, France
M Schell Jean-Baptiste (invité)	Docteur, CNRS, France

Acknowledgments

This research was conducted at the ICube laboratory of the University of Strasbourg and CNRS, UMR 7357, and was funded by a grant from Agence Nationale de la Recherche (ANR-18-CE24-0010).

I wish to express my profound gratitude to my dedicated mentors, Professor Wilfried Uhring and Dr. Jean-Baptiste Kammerer, whose support and guidance have been instrumental at every stage of my doctoral journey. Working under their supervision was not only a scientific privilege but also a valuable source of life lessons. They have profoundly shaped my scientific identity, exemplifying dedication, and passion for the pursuit of knowledge.

I would also like to extend my heartfelt appreciation to my colleagues at ICube, Dr. Fabrice Aguenounon, Pascal Leindecker, and Nicolas Collin, for their invaluable technical support, especially during the chip characterization phase. Special thanks are owed to Dr. Jean-Baptiste Schell for his consistent availability in answering my queries.

Gratitude is also extended to ST Microelectronics and CMP for granting access to their technology and providing essential IC fabrication services. I offer special thanks to Andreia Cathelin for her assistance.

I am deeply thankful to Professor Francis Calmon for leading the SPADFDSOI project and for supplying the SPADs vital to this research.

I owe a debt of gratitude to Professor Luc Hebrard and Dr. Morgan Madec for affording me the opportunity to start my doctoral journey in France.

To my family and friends, I extend my heartfelt thanks for your encouragement and belief in me. Your unwavering emotional support has been an eternal source of motivation.

Lastly, I reserve a special place of honor and gratitude for my wife, who stood by my side throughout my years of PhD pursuit, offering her eternal love and support.

Table of contents

Acknowledgments 1
Table of contents 1
List of Acronyms 0
Introduction 1
Chapter 1 3
SPAD and Quenching Circuits 3
1.1. SPAD Operation Principals 3
1.2. SPAD Performance Parameters 5
1.2.1. Photon Detection Probability (PDP) 5
1.2.2. Timing Jitter 6
1.2.3. Dark Count Rate (DCR) 7
1.2.4. Afterpulsing 9
1.2.5. Fill Factor & Photon Detection Efficiency (PDE) 10
1.2.6. Crosstalk 12
1.3. Quenching Circuits State of the Arts 13
1.3.1. Passive Quenching 13
1.3.2. Active Quenching 17
1.3.3. State-of-the-Art Active Quenching Circuits 19
1.4. Thesis Objective 31
1.4.1. Monolithic 3D SPAD Pixel in FD-SOI CMOS Technology 31
Chapter 2 34
2.1. Introduction 34
2.2. Low Threshold Inverter as an Avalanche Detector 35
2.2.1. Delay Reduction Through Inverter Sizing 36
2.2.2. Delay Reduction Through Body Biasing 38
2.3. Body Biased Inverter Based AQAR Circuit 41
2.4. Results and Discussions 44
2.4.1. Chip and Test Bench 45
2.4.2. Simulation Results 50
2.4.3. Experimental Results 52
Chapter 3 58

3.1. Introduction 58
3.2. Ticklish Inverter for Avalanche Detection 59
3.2.1. Ticklish Avalanche Detector 60
3.2.2. Ticklish Active Quenching Active Reset Circuit 63
3.3. Native 3D SPAD Active Pixel 65
3.3.1. Native 3D Pixel Structure 66
3.3.2. Placing the Circuit Over SPAD 67
3.3.3. Mixed Sensing and Indirect Sensing Pixels 68
3.4. Results and Discussion 73
3.4.1. Post Layout Simulation Results 73
3.4.2. Experimental Results 77
Chapter 4 86
4.1. Introduction 86
4.2. Pre-biased Inverter for Avalanche Detection 86
4.2.1. Pre-biasing Idea 86
4.2.2. Biasing Circuit 89
4.3. Native 3D Pixel & Body Biasing Issue 94
4.3.1. Body Effect in FD-SOI Technology 94
4.3.2. Body Independent Biasing Circuit 95
4.3.3. Body Independent Current Source 96
4.4. Pre-Biased Active Quenching Active Reset Circuits 99
4.5. Results and Discussion 102
4.5.1. Post Layout Simulation Results 102
4.5.2. Experimental Results 109
Chapter 5 120
5.1. Conclusions 120
5.2. Perspectives 125

List of Acronyms

ACR	Afterpulsing Count Rate
ANR	Agence Nationale de la Recherche Française
APP	Afterpulsing Probability Percentage
AQ	Active Quenching
AQAR	Active Quenching-Active Reset
AR	Active Reset
ATP	Avalanche Triggering Probability
BBI	Body Biased Inverter-Based
BOX	Buried Oxide layer
BSI	Back-Side Illumination
CTC	Current Transfer Characteristic
DAC	Digital to Analog Converters
DCR	Dark Count Rate
DRC	Design Rule Checking
DS	Direct Sensing
FAQ	Fast Active Quenching
FD--SOI	Fully Depleted-Silicon on Insulator
FLIM	Fluorescence-lifetime imaging microscopy
FPGA	Field-Programmable Gate Array
FSI	Front-Side Illumination
FW	Flipped Well
FWHM	Full Width at Half Maximum
InDS	InDirect Sensing
INL	Institut des Nanotechnologies de Lyon
LiDAR	Light Detection and Ranging
MS	Mixed Sensing
OPW	Optical Wireless Communication
PDE	Photon Detection Efficiency

PDP	Photon Detection Probability
PQ	Passive Quenching
PQPR	Passive Quenching-Active Reset
PR	Passive Reset
RW	Regular Well
SCR	Space Charge Region
SPAD	Single Photon Avalanche Diode
STI	Shallow Trench Isolators
TCSPC	Time-correlated Single Photon Counting
VLC	Visible Light Communication
VLQC	Variable Load Quenching Circuit
VTC	Voltage Transfer Characteristic

Introduction

Thanks to their high sensitivity in the optical signal detection and picosecond range temporal resolution, Single Photon Avalanche Diodes (SPADs) play a significant role in a variety of applications that require photon counting or photon timing measurements. Light Detection and Ranging (LiDAR), quantum computing, quantum cryptography, Fluorescence Lifetime Imaging Microscopy (FLIM), and Time-Correlated Single Photon Counting (TCSPC) are some of these applications.

SPADs are PN junctions, reversely biased over the breakdown voltage. This biasing enables them to detect single photons with an incredibly low jitter. SPADs flag a photon by generating a surging current, called the avalanche current. Quenching circuits control the flow of the avalanche current inside the SPAD, and finally extinguish it completely to avoid devices overheating or destruction.

In some emerging applications of the SPAD such as Optical Wireless Communication (OPW), high density arrays and high photon count rates are required. High count rate by a single device is only achievable by shortening the dead time of the SPAD, the time during which SPAD is unresponsive to single photons. Unfortunately shortening the dead time severely increases the afterpulsing effect. Afterpulsing effect is a correlated noise in SPAD that generates spurious avalanches in a short random time after the main photo-generated avalanche. This effect degrades the overall system performance by reducing the dynamic, the signal-to-noise ratio and it also modifies the temporal profile of the signal. Afterpulsing effect can be reduced by utilizing Active Quenching-Active Reset (AQAR) circuits. However, the increased pixel area due to the implementation of these circuits, reduces the sensitivity of the SPAD pixel, which is expressed as Photon Detection Efficiency (PDE). This trade-off between maximum photon count rate, afterpulsing effect, and PDE has raised new research interest on AQAR circuits and innovative pixels structure such as stacked 3D pixels.

The aim pf this Ph.D. thesis is to propose ultra-fast AQAR circuits in a 28 nm Fully Depleted-Silicon on Insulator (FD-SOI) CMOS technology, alongside with a novel avalanche sensing

method and a new SPAD pixel structure, that are able to mitigate the mentioned trade-off without the extra expenses of 3D stacking process at wafer level. This work is a collaboration between ICube laboratory of University of Strasbourg and Institut des Nanotechnologies de Lyon (INL), which provide SPAD devices structures and layouts. The project is funded by the Agence Nationale de la Recherche Française through the grant of the project “SPAD-FDSOI” (ANR-18-CE24-0010).

The manuscript organization is as follows: in chapter 1 a brief introduction on SPAD physics and its operation principles is presented. It follows by introducing some of the SPAD important parameters such as Photon Detection Probability (PDP), jitter, Dark Count Rate (DCR), afterpulsing effect. The rest of this chapter is dedicated to the state of the arts of AQAR circuits. At the end of this chapter, the SPAD devices used in this thesis are briefly introduced.

In chapter 2, an ultra-fast AQAR circuit, featuring a simple and compact avalanche detection circuit is presented. Through post layout simulations and experimental results, the functionality and efficiency of the proposed AQAR circuit in the afterpulsing reduction are investigated.

In chapter 3, a novel avalanche detection circuit with three different variants is presented, which results in three AQAR circuits. The concept of monolithic 3D SPAD pixel is introduced and based on that the first active monolithic 3D SPAD pixel is introduced. A new avalanche sensing method is also proposed that can increase the efficiency of AQAR circuits. Thanks to the post-layout simulation and experimental results, all the proposed circuits are validated. Also, effects and challenges of placing electronics over the SPAD in the active monolithic 3D SPAD pixel is discussed.

In chapter 4, body biasing in the FD-SOI CMOS technology and its effect on the active monolithic 3D SPAD is investigated. The first body voltage independent current source is presented to bias the circuits in the monolithic 3D pixel. Another avalanche detection circuit in three different variants is proposed. Each variant leads to a low power ultra-fast AQAR circuit. Like the last two chapters, post-layout simulations and experimental results are presented.

Finally, chapter 5 concludes the manuscript by comparing the different proposed AQAR circuits. Furthermore, this chapter outlines potential avenues for future development and research, highlighting the prospects of the thesis.

Chapter 1
SPAD and Quenching Circuits

1.1. SPAD Operation Principals

Reversely biasing a diode over its breakdown voltage generates a strong electric field across the diode junction. Under this biasing condition, the impact ionization can occur in the multiplication region [1]–[3]. The multiplication region is part of the depleted area of the junction where has the highest electric field. In this region, a free carrier, either an electron or a hole, has enough energy to break a bond in the semiconductor lattice and thus generates an electron-hole pair. The newly generated carriers also undergo the same scenario, and a positive feedback loop of impact ionization arises in the Space Charge Region (SCR). At breakdown voltage the multiplication factor of the PN junction is infinity [4], [5]. Thus, a fast-rising avalanche of charge carriers flows through the SCR. The initial free carrier can be generated by an incident photon thanks to the internal photoelectric effect [6], [7]. Single Photon Avalanche Diodes (SPADs), based on these principles, are able to convert a single photon to a measurable current, called the avalanche current, within a few tens to hundreds of picoseconds[1], [8]–[11]. The avalanche process in a reversed biased PN junction is illustrated in Fig. 1.1.

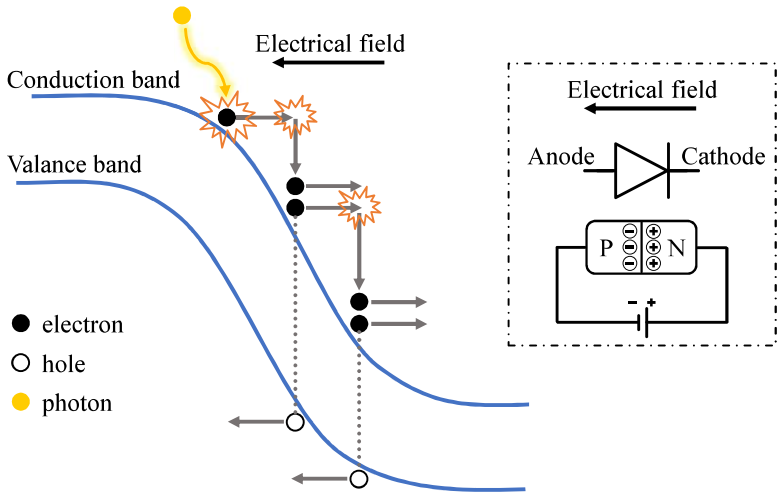


Fig. 1.1 Avalanche process in a reverse biased PN junction due to the photo electric phenomenon. Inset: a diode in the reverse bias with its electrical symbol.

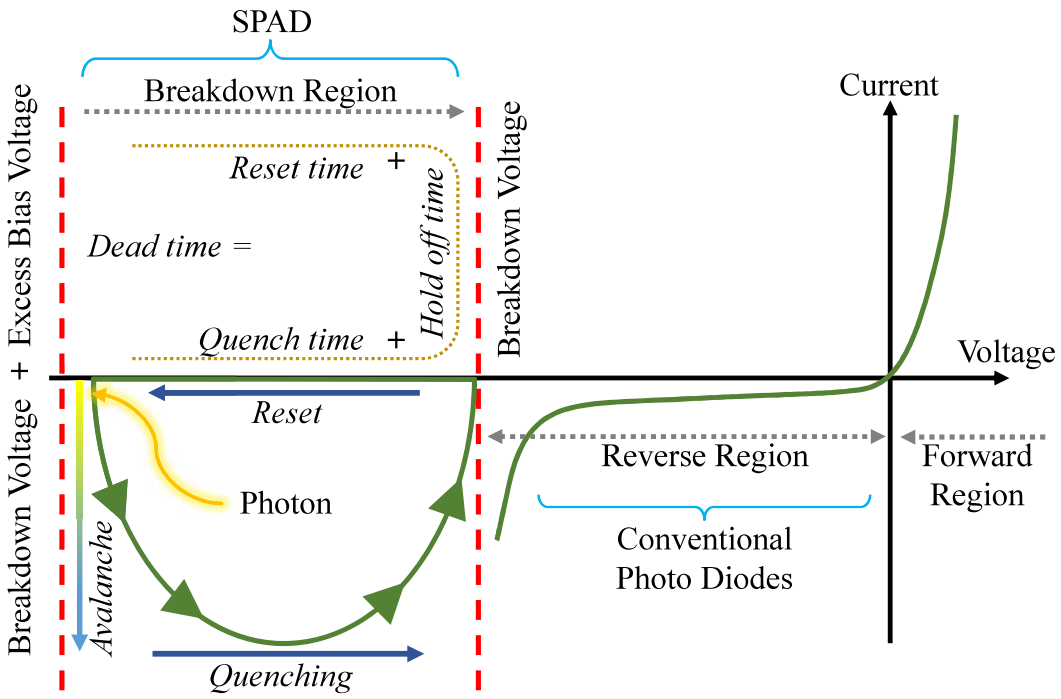


Fig. 1.2 Symbolic I-V characteristic of a SPAD with quench-reset dynamic.

Fig. 1.2 shows the I-V characteristic of a typical diode. To operate as a SPAD, the diode is biased in the breakdown region, over the breakdown voltage. The difference between the SPAD biasing voltage in the breakdown region and the breakdown voltage is called excess bias voltage. In this region, an incident photon can generate a self-sustaining avalanche current [12]. The current can be high enough to destroy the diode if it is not stopped. The process of stopping avalanche current to flow is called quenching [13]. To quench the SPAD, the voltage across its junction must drop below the breakdown voltage. The quenching can be done by means of an active circuit, which is called Active Quenching (AQ), or simply by putting a large resistor in series with the SPAD, in the Passive Quenching (PQ) case [14], [15]. Once the avalanche is quenched and the SPAD is biased out of its breakdown region, the electric field across the junction is reduced and consequently the SPAD cannot generate a self-sustaining avalanche current. To prepare the SPAD for the next photon detection, it must be biased again in the breakdown region. Similar to the quenching process, resetting (recharging) the SPAD can be done by an active circuit, Active Reset (AR), or by a passive resistive path, Passive Reset (PR) [16], [17]. The avalanche dynamic behavior is illustrated in Fig. 1.2. The time interval between the onset of the avalanche and end of the reset phase is known as deadtime, during which the SPAD is unable to detect a new photon. It

is composed of quenching time, reset time, and hold-off time, during which the SPAD is below the breakdown voltage. A short deadtime increases the photon count rate, at the cost of increasing SPAD noise [9]. This trade-off is discussed in the following sections.

The ability to detect a single photon with a high temporal resolution makes SPAD suitable for photon counting and photon timing applications. SPADs play a critical role in a wide range of applications, from bio photonic to quantum computing, including Fluorescence Lifetime Imaging (FLIM), optical tomography, quantum random number generator, Light Detection and Ranging (LiDAR), Visible Light Communication (VLC), and many others [18]–[51].

1.2. SPAD Performance Parameters

There are several parameters that determine the SPAD performance, from the pixel level up to the system level, in the mentioned applications. In this section some of the most important parameters of SPAD are briefly introduced, and the proposed solutions in the literature to improve these parameters are presented.

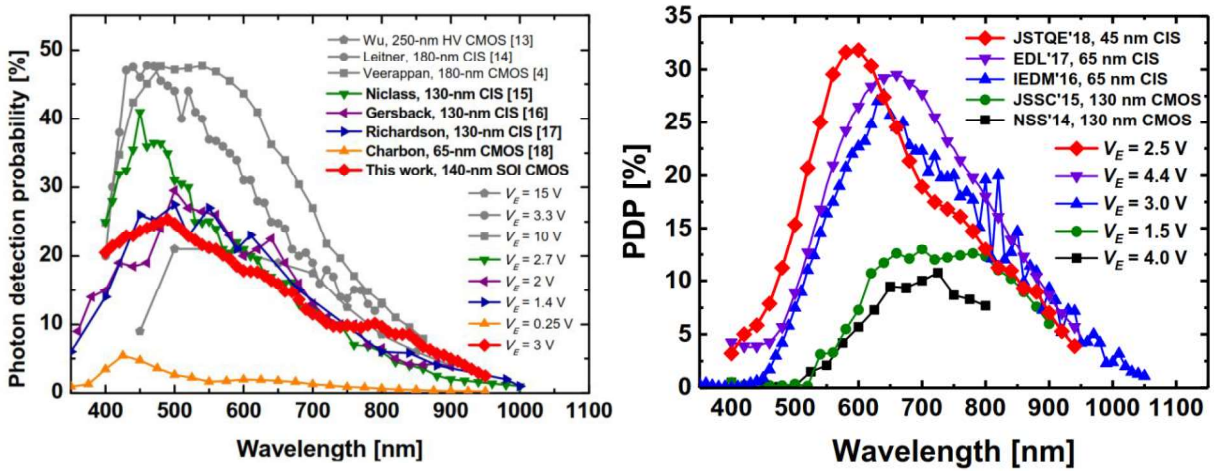
1.2.1. *Photon Detection Probability (PDP)*

SPAD is a single photon detector, however, not all the impinging photons to its light sensitive area can be necessarily detected. The probability that an incident photon generates a self-sustaining measurable avalanche current is defined as Photon Detection Probability (PDP), which indicates the sensitivity of a SPAD. There are several possible scenarios for the incident photons to a SPAD: some of them are reflected by the different material stacks in the technology structure. Some of the incident photons are absorbed in the neutral regions of the junction. Among the resulting photogenerated carriers, the majority carriers (electrons in N and holes in P) are repelled by the electric field across the multiplication region, thus, it is impossible for them to start an impact ionization. The minority carriers have the chance to diffuse into the multiplication region, however, they can also be trapped or recombined before reaching the multiplication region [1], [52], [53]. Even a carrier that reaches the multiplication region or is absorbed in that region, because of the statistical nature of the impact ionization, has a certain possibility to trigger an avalanche, defined as the Avalanche Triggering Probability (ATP) [54], [55]. Therefore, PDP of a SPAD cannot reach 100 %. ATP is increased by the electric field across the junction. Therefore, to increase the PDP, one can increase the excess bias voltage and apply a uniform electric field through the doping

engineering [56]–[64]. Anti-reflection coating is another approach to increase the PDP by decreasing the number of reflected photons [65], [66].

PDP also depends on the wavelength of the incident light and the optical absorption of the SPAD material [1], [58], [67]. Thus, to have an optimum PDP at a specific wavelength range from near ultra-violet to near infra-red (250 nm – 1550 nm), different semiconductor materials (silicon and other compound semiconductors) are used to implement the SPAD [47], [48], [68]–[87].

The PDP peak occurs at different wavelengths regarding the illumination configurations. Fig 1.3 shows that in a Front-Side Illumination (FSI), the peak PDP is around the near ultra-violet and blue range, while in the Back-Side Illumination (BSI), the peak PDP is shifted around the infra-red range. Also, the technology scaling effect on the PDP is also observable in Fig. 1.3 (a). In the lower technology nodes, the PDP peak becomes lower and narrower [88], [89].



(a) (b)

Fig. 1.3 Photon Detection Probability (PDP) versus impinging light wavelength at different technology nodes: (a) Front-Side Illumination (FSI) [88] (b) Back-Side Illumination (BSI) [89].

1.2.2. *Timing Jitter*

Thanks to the high electric field over the multiplication region, SPAD has an ultra-fast time response, i.e., the avalanche process starts almost immediately by an impinging photon [10], [11]. However, the coming photons, as explained previously, are not necessarily absorbed in the multiplication region. Those carriers generated by the photons who land in the neutral regions, pass different paths to reach the multiplication region, which takes different times. Also, each carrier in the multiplication region can experience a different avalanche build-up [90] process due to the

statistical nature of the impact ionization [91]–[95]. These effects introduce fluctuations in the time response of the SPAD, which is called timing jitter. It is expressed as the distribution of the photon arrival time at its Full Width at Half Maximum (FWHM). To measure the jitter, a picosecond laser pulse repeatedly illuminates the SPAD. Each time the interval between the trigger time of the laser pulse and the photon detection time is measured to obtain the photon arrival time histogram [96], [97]. Fig. 1.4 shows an arrival time histogram presented in [98]. The exponential part of the histogram is called the diffusion tail, representing the avalanches generated by the photons absorbed in the neutral regions [99]. The jitter of a SPAD can be as low as a few tens of picoseconds [87], [97], [100]–[103]. Same as in PDP case, a higher electric field improves the SPAD jitter [102].

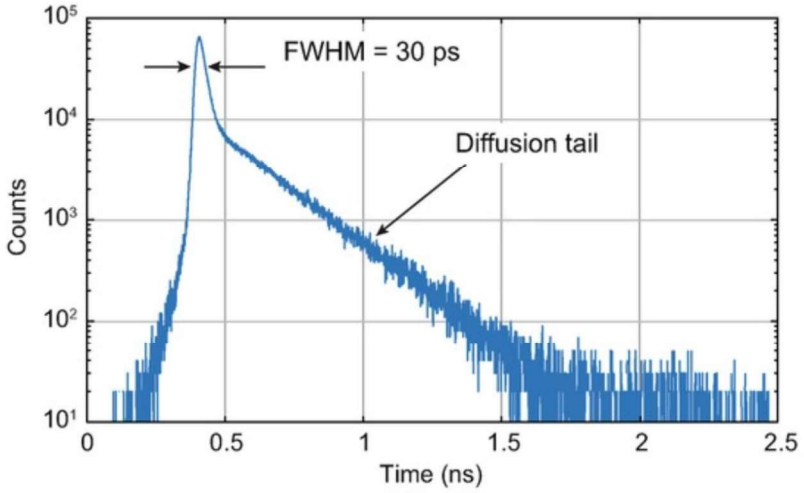


Fig. 1.4 Photon arrival time histogram of a SPAD, composed of a Gaussian peak and a diffusion tail [98].

1.2.3. *Dark Count Rate (DCR)*

Photoelectric phenomenon is not the only way that starts an avalanche. A thermally generated carrier or a free carrier by tunneling effect can also trigger the impact ionization process [1], [104], [105]. Thus, in the absence of light, SPAD may generate non-photogenerated avalanches. These spurious avalanches are one of the main noise sources of SPADs, called dark counts. Thermal generation, trap-assisted generation, band to band tunneling, and trap-assisted tunneling are responsible for dark counts [106], [107], see Fig. 1.5. Dark Count Rate (DCR) is defined as the number of avalanches per second when the SPAD is placed in complete darkness. DCR delimits the minimum photon count rate of a SPAD, thus, for the applications with low intensity input light is crucially important. DCR is dependent on the semiconductor materials, doping profile,

temperature, and the junction electric field [9], [108]. Despite PDP and jitter, a high excess bias voltage increases DCR, therefore, there is a trade-off between these parameters. A common way to decrease the DCR is to cool down the SPAD, since there is a proportional relationship between the DCR and the temperature [9], [56], [87], [97]. Fig. 1.6 shows the dependency of DCR on the excess bias voltage and the temperature.

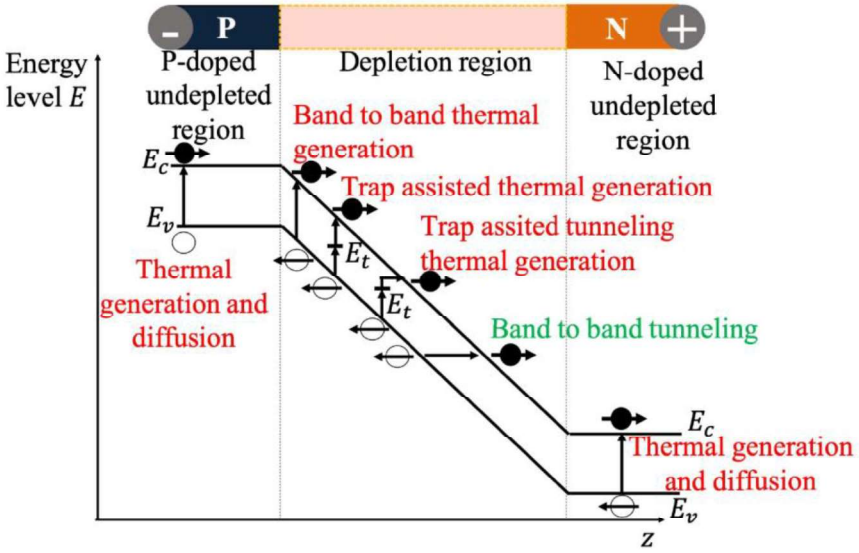


Fig. 1.5 DCR contributions illustration for a reverse biased p-n junction [106].

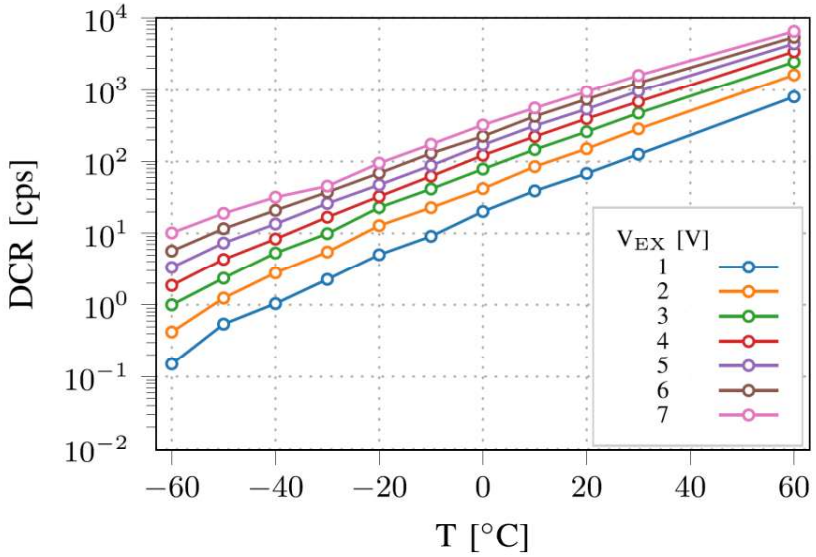


Fig. 1.6 Temperature and excess bias voltage dependence of DCR for a SPAD presented in [56].

1.2.4. *Afterpulsing*

When the avalanche current is passing through the SCR, some of the carriers are captured by the traps inside the semiconductor lattice. These traps, which are the result of defects and impurities in the semiconductor material, hold the carrier for a random time, depending on the temperature and the energy of the trap [109], see Fig. 1.7. The captivity time can reach up to the microsecond range. Two possibilities exist for the released carrier: after the main photo-generated avalanche, if the SPAD is reset sooner than the longest trap lifetime, the released carrier can trigger a new impact ionization, resulting in a non-photo-generated avalanche. This spurious avalanche, which has a correlation with the main photo-generated avalanche, is called afterpulse, and the process of the afterpulse generation is called afterpulsing effect or simply afterpulsing. Afterpulsing effect, alongside with the DCR, is a main noise source in SPADs that delimit their maximum photon count rate [9].

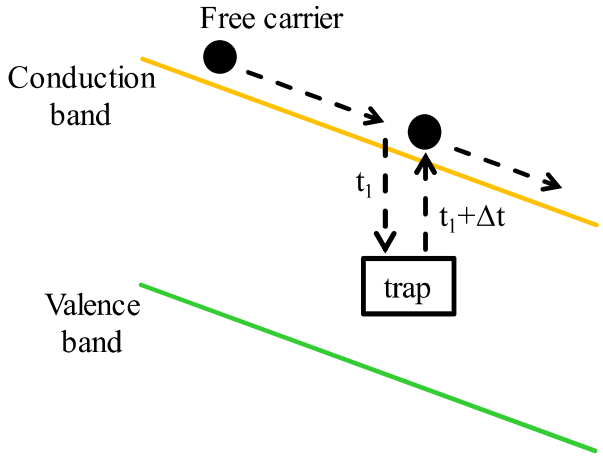


Fig. 1.7 Trapping and release of a free carrier in an avalanche photo diode: the origin of afterpulsing effect.

However, in the second scenario, if the carrier is released while the SPAD is out of the breakdown region, i.e., during the hold-off time, the electric field is not high enough to allow the released carrier to start an avalanche. Therefore, it is a common solution in the afterpulsing reduction to prolong the dead time of the SPAD [110]–[113]. Fig. 1.8 shows the dependency of afterpulsing on the dead time [113]. On the other hand, prolonging the dead time, severely decrease the maximum photon count rate, which is not desired in the photon counting applications such as VLC and LiDAR [114]. Also, cooling down the SPAD to reduce the DCR, increases the afterpulsing effect, because at lower temperatures the traps tend to hold the carriers for a longer

time [109]. A higher electric field, as in the DCR case, increases the afterpulsing effect as shown in Fig. 1.8.

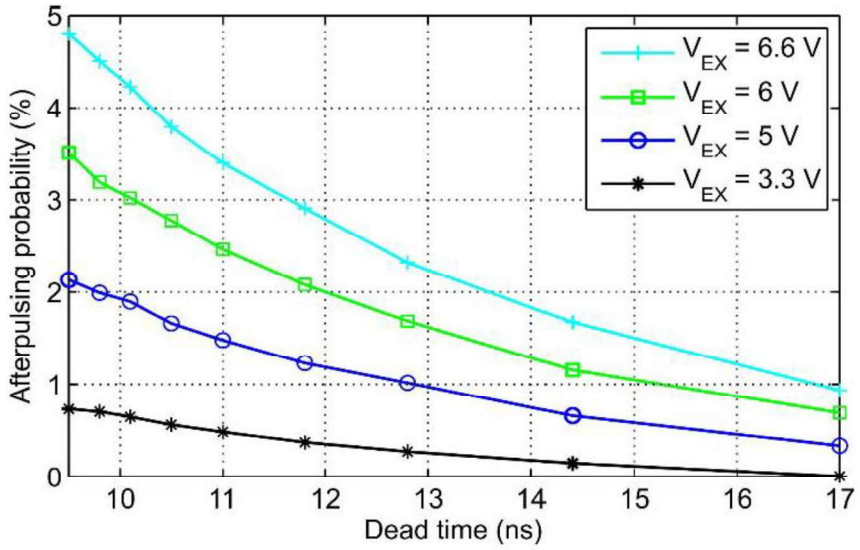


Fig. 1.8 Afterpulsing probability versus dead time at different excess bias voltages (V_{EX}) [113].

Afterpulsing effect depends on the number of traps in the semiconductor lattice and the number of charge carriers during an avalanche. Thus, a way to reduce the afterpulsing effect is decreasing the avalanche charge [115]. This is done by AQ circuits, which provide part of the charge that is needed to bias the SPAD out of the breakdown region and stop the impact ionization process. It allows the SPAD not to generate all the required charge for the quenching, and therefore reduces the charge carriers in the SCR. In this solution, afterpulsing effect is reduced without affecting the maximum photon count rate [116]. The AQ circuits are discussed in detail in this chapter. However, the AQ circuits have some drawbacks that are highlighted in the next section.

1.2.5. *Fill Factor & Photon Detection Efficiency (PDE)*

The ratio between the SPAD active (light sensitive) area and the whole pixel area, including the guard ring, quench/reset circuits, and readout circuit, is defined as fill factor. The fill factor value is presented as a percentage and is always less than 100 %. It measures how efficient the total area of a pixel is used to detect the incident photons. A larger fill factor means more photons can be absorbed in the active area of the SPAD pixel. Thus, there is a proportional relationship between the fill factor and the sensitivity of a pixel [47], [114]. This relationship is defined by another important SPAD parameter, named Photon Detection Efficiency (PDE). PDE is the product of PDP

and fill factor [117]. It is the probability of detecting a photon colliding on the SPAD pixel and measures the sensitivity and efficiency of a pixel in detecting a photon.

To have the maximum sensitivity, the fill factor must be maximized. The electronic interface of the pixel, e.g., Active Quenching-Active Reset (AQAR) circuits, which are necessary for afterpulsing reduction and other functionality of the pixel such as dead time control, severely reduces the fill factor, and therefore the PDE. One can see a trade-off between afterpulsing effect, maximum photon count rate and fill factor or PDE [47], [114]. A few solutions are proposed in the literature to alleviate this trade-off. Using microlenses is one of these solutions [70], [75], [118]–[125]. A microlense is literally a small lens that focuses the coming photons onto the active area of the pixel. Collecting more photons on the SPAD active area and consequently detecting more, results in a higher fill factor and PDE. Fig. 1.9 shows a microlens array integrated with an SPAD array [18]. The main drawback of microlenses is the extra fabrication costs and technological constraints.

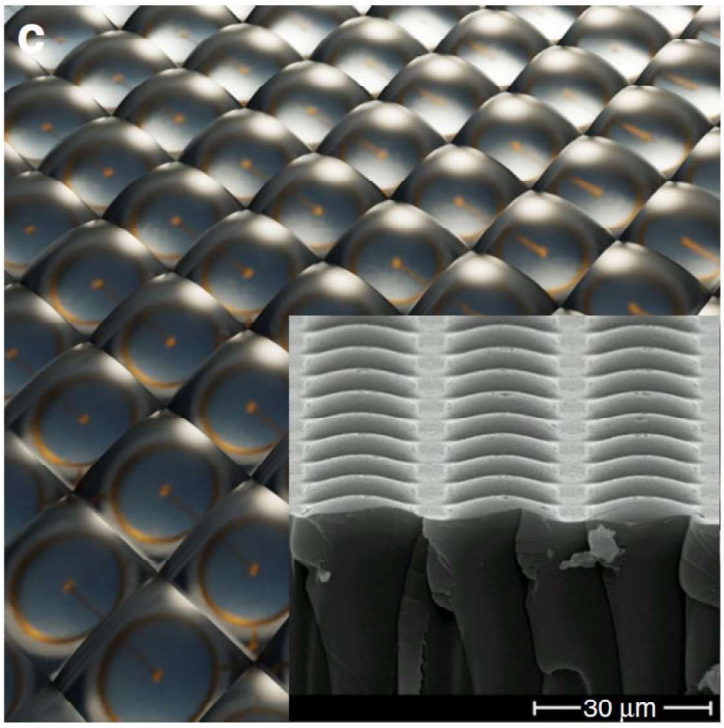


Fig. 1.9 Microlens array integration with a SPAD array [18]. Inset: scanning electron microscope image of the microlens array from ref. [125].

Another solution is the stacked 3D IC technology. In this method, the SPAD array and the peripheric circuitry are fabricated on separate dies and then bonded together through specific

bonding processes [30], [76], [126]–[129], see Fig. 1.10. In the absence of the interface electronics, the pixel area is mostly occupied by the SPAD active area, resulting in a very high fill factor and PDE. Beside this main advantage, stacked 3D IC technology allows utilizing specific technologies for each separate die. Implementing SPADs in customized SPAD technology gives the opportunity to fine tune the SPAD performance parameter such as PDP and DCR. Also, it is possible to use scaled technology nodes for the electronics to benefit from their high area density and low power consumption [64], [130]–[132]. The drawback of this method is the high manufacturing costs due to the two separate dies, in comparison with the conventional 2D IC technology. Also, the fabrication process is complicated by the interconnections of the dies [18], [133].

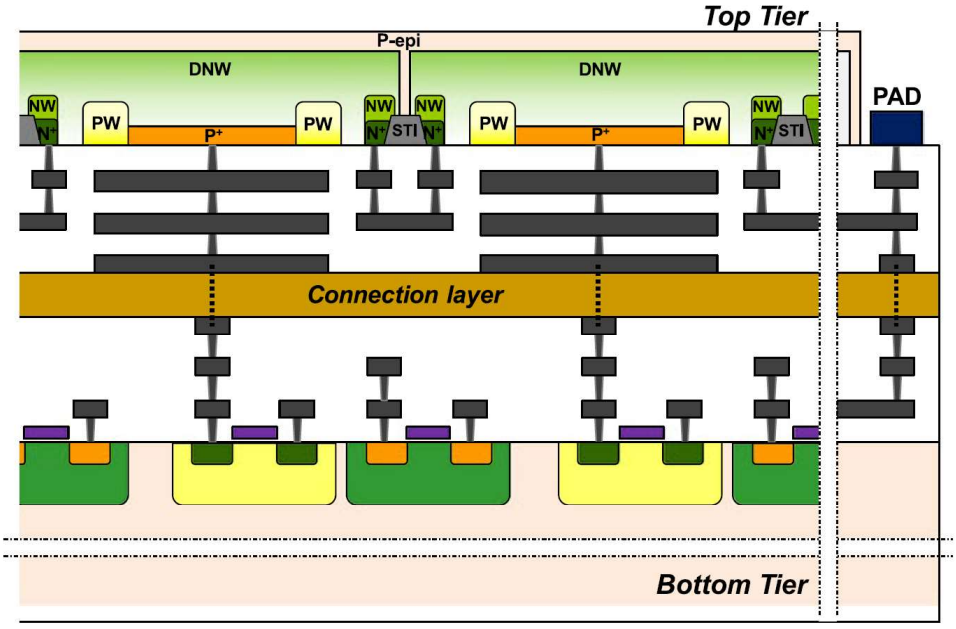


Fig. 1.10 Cross section of a 3D stacked backside illuminated SPAD array [126].

1.2.6. *Crosstalk*

In a SPAD array, avalanche flows through the SCR of a SPAD can trigger other avalanches in the neighboring SPADs. The spurious avalanche that is correlated with the first avalanche is another noise source of the SPAD, called crosstalk. When an avalanche of carriers is flowing into the junction, some of the carriers can diffuse to junction of the neighboring SPADs. There, as a free carrier, it can start a new avalanche. This type of crosstalk is categorized as electrical crosstalk. The avalanche carriers can also undergo an electron-hole recombination process as explained in

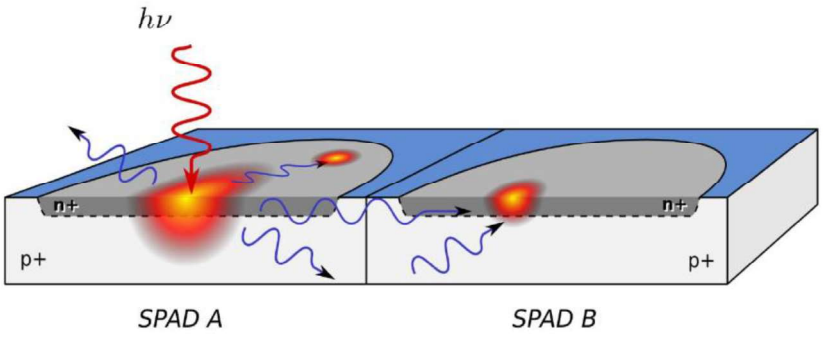


Fig. 1.11 Symbolic representation of optical cross talk between two SPADs [135].

the PDP section. The electron-hole recombination generates a photon that can be absorbed by the neighboring SPADs [52]. This secondary photo-generated avalanche is optical cross talk, see Fig. 1.11. To reduce the crosstalk noise, it is proposed to place insulators between adjacent SPAD or increase the distance between the SPADs in the array [9], [134]–[136]. Also, one can see the source of crosstalk are avalanche charge carriers [137]. Thus, as an ultimate solution, reducing the avalanche charge can reduce crosstalk. Same as in the afterpulsing case, this avalanche charge reduction can be done through the AQ circuits. However, implementing each of these solutions degrades fill factor, pitch, and PDE of the SPAD, which is an obstacle in realizing dense SPAD arrays.

1.3. Quenching Circuits State of the Arts

In the breakdown region, even one single photon can generate an avalanche of current carriers, which can keep flowing till damaging the device. Thus, as it is already explained, once an avalanche has occurred it is mandatory to stop its flow through the SPAD to avoid devices overheating or destruction. This action is done by biasing the SPAD out of its breakdown region through a quenching circuit. When the voltage across the SPAD is below or equal to the breakdown voltage, the SPAD is unable to generate a self-sustaining avalanche current. A quenching circuit is also necessary for determining and limiting the dead time and the count rate of the device, and to make the device reusable for the next coming photons.

1.3.1. *Passive Quenching*

There are two general categories of quenching circuits based on their components: Active Quenching (AQ) and Passive Quenching (PQ) circuits. In PQ circuit, SPAD is connected to a high

impedance path which can be a high value resistor or a transistor that operates as a resistor. Fig. 1.12. shows schematic of a PQ circuit. In this circuit, once an avalanche breakdown occurs, the avalanche current starts to charge the SPAD junction capacitance C_{SPAD}. Thus, the voltage across the quenching resistor R_Q increases till it reaches the excess bias voltage. Now the SPAD is quenched, and the avalanche current has stopped.

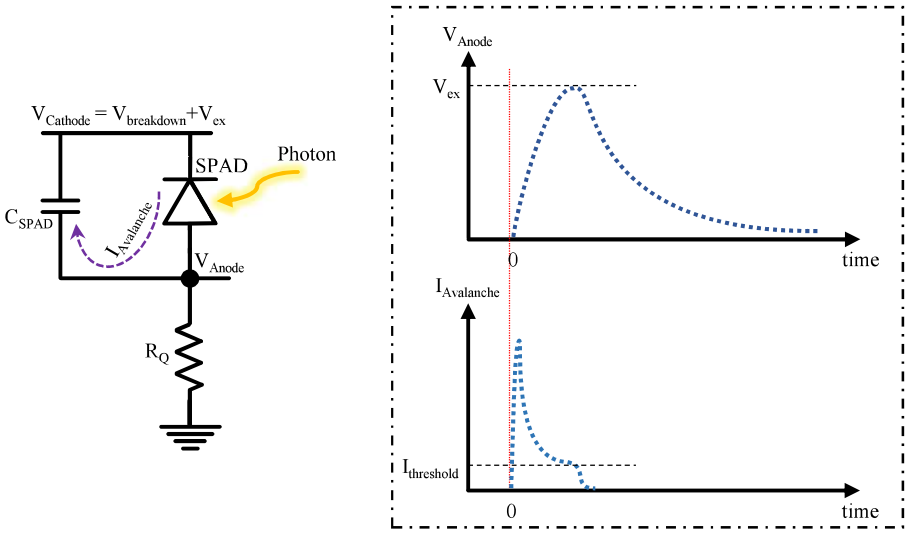


Fig. 1.12 Schematic of the PQ circuit with the anode voltage and the avalanche current transitions during an avalanche.

It should be noted that in PQ circuit, the diode voltage cannot reach below the breakdown voltage, thus, the current can still flow through the SCR. In the literature, however, a threshold is defined for the avalanche current, which below this threshold, the avalanche cannot be a self-sustaining current and the carriers will leave the SCR before triggering a self-sustaining avalanche [13]. It establishes a minimum value for the passive quenching resistor [15]. The maximum current passing through RQ is equal to:

$$I_{max,Q} = \frac{V_{ex}}{R_Q} \tag{1}$$

Where V_{ex} is the excess bias voltage. This current is provided by the SPAD. Thus, to effectively quench the SPAD, this current must be sufficiently lower than the avalanche threshold current that implies a minimum RQ value. A lower R_Q, makes the SPAD to function almost as a Zener diode with a potentially very high current that can destroy the diode.

The quenching time in PQ circuit is mostly determined by C_{SPAD} and other physical phenomena that are involved in the avalanche build up process. In fact, almost all the avalanche current is passing through the SPAD junction capacitance, while only a negligible amount of current is passing through the quenching resistance. Therefore, R_Q has almost no effect on the quenching time. Since the avalanche build up is a statistical process, the quenching time in a PQ circuit is also statistical. Even though, a higher quenching resistor can slightly decrease the quenching time [13].

In the reset phase however, R_Q plays a major role and defines the reset time constant with C_{SPAD}. Since R_Q typically is high, the reset phase in PQ circuit is long. During the reset phase, the SPAD voltage is above the breakdown, thus, the possibility of new avalanche triggering exists before complete reset of the SPAD. A new avalanche in this condition prolongs further the reset time and limits the maximum photon count rate. Moreover, it makes the reset also a statistical process which depends on the number and duration of the weak avalanches during the reset phase.

PQ circuits are easy to implement and occupy a very small area which makes them desirable for realizing SPAD arrays [138], [139]. However, as mentioned already, they suffer from indefinite quenching and reset times. Even though integrating an AR circuit with a PQ circuit can result in a well-defined reset time [17], [64], [97], [140], [141], yet the quenching time is a problem. Besides, by adding electronics for implementing the AR circuit, the resulting circuit cannot be as compact as it is claimed. Another drawback of the PQ circuit is that it cannot reduce the afterpulsing effect, since in this quenching method, all the required quenching current (the current that needs to charge C_{SPAD}) is provided by SPAD itself.

Fig. 1.13 (a) shows a Passive Quenching-Active Reset (PQAR) circuit which is presented in [64] and with slight modifications in [97]. The authors used the cascode technique to enable the SPAD to operate with a higher excess bias voltage than the oxide breakdown of the transistors. By cascoding M_2 with M_3 which is implementing the passive quenching resistor, the anode voltage (and consequently the excess bias voltage) can safely reach a voltage two times higher than the maximum tolerable voltage of the oxide of M_3. Since the MOSFETs M_1 to M_5 are thick oxide MOSFETs, their maximum tolerable voltage is 2.75 V in the used technology. By applying 4.4 V of V_{ex}, thanks to this technique, around 30 % improvement in the peak PDP value of the device and 13 ps reduction of the jitter is achieved in comparison with V_{ex} = 2.4 V. Fig. 1.14 shows the

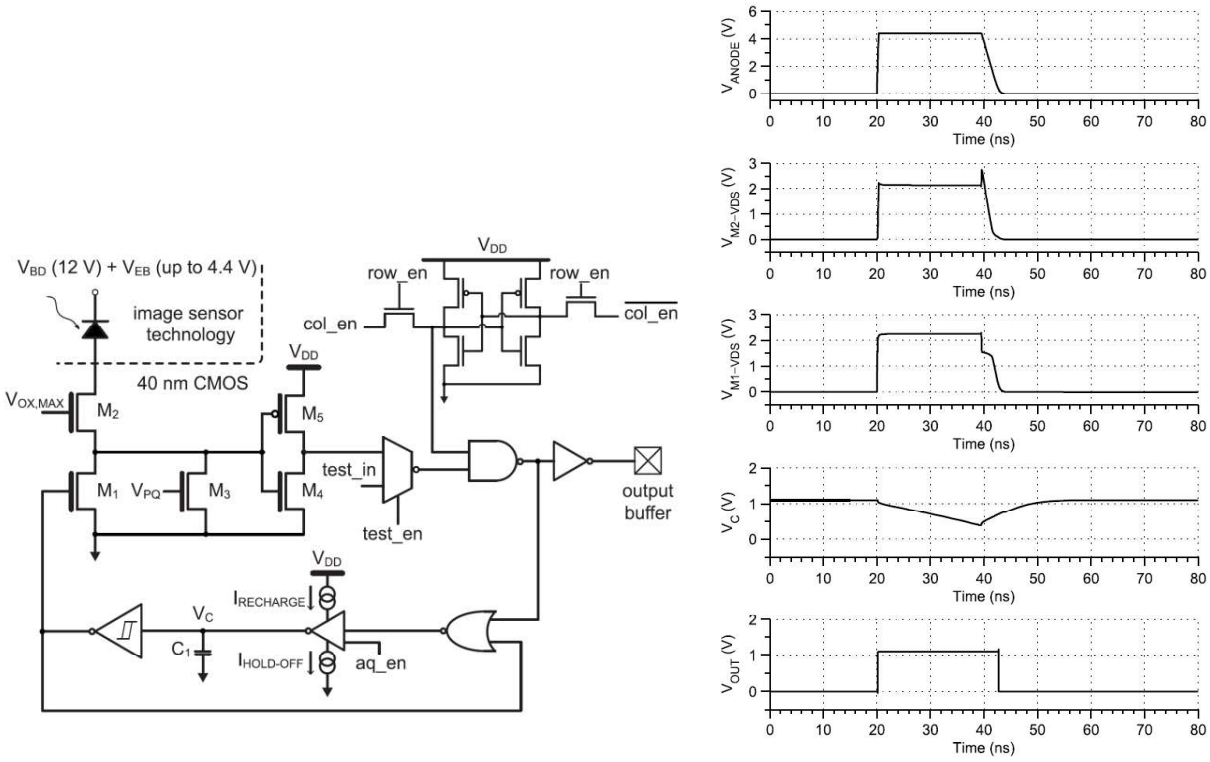


Fig. 1.13 (a) Schematic of the PQAR circuit presented in [64]. (b) Timing diagram of the circuit in (a).

PDP and jitter improvements in this circuit. It should be noted that for a safe operation at 4.4 V of V_{ex}, the gate voltage of M_2 $V_{OX,MAX}$, must be the highest possible value (2.75 V).

The timing diagram of the circuit is presented in Fig. 1.13 (b). Avalanche current flow raises the anode voltage. The anode voltage distributes between the drain-source voltages of M_1 and M_3 as well, till completely quenches the SPAD. During this time M_1 is set to be off. The drain-source voltage rise of M_3 is sensed by the inverter composed of M_4 and M_5. This inverter finally drives a current starved inverter which controls the hold off time and reset time of SPAD through $I_{RECHARGE}$, $I_{HOLD-OFF}$, and C_1. During the quenching phase C_1 is discharged by $I_{HOLD-OFF}$ till it reaches the lower threshold of the Schmitt trigger. Now M_1 turns on and the SPAD is reset through M_1 and M_2. After a determined time, C_1 is charged by $I_{RECHARGE}$ till V_C passes the upper threshold of the Schmitt trigger that turns M_1 off once again to end the reset phase and make the pixel ready for a new event. The total circuit area is 25.2 μm^2.

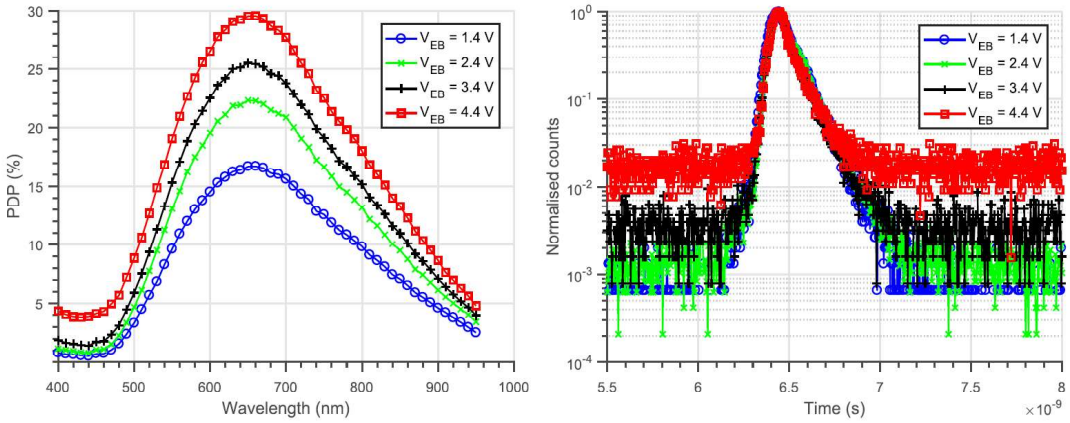


Fig. 1.14 (Left) PDP versus wavelength at different excess bias voltages and (Right) Jitter at different excess bias voltages for the pixel presented in [64].

1.3.2. *Active Quenching*

Nondeterministic quench and reset timing, plus inability in afterpulsing reduction, as two main drawbacks in PQ circuits, has led to another type of quenching circuit, named Active Quenching (AQ) circuit. An AQ circuit is typically composed of two main parts: an avalanche detection circuit and a fast switch or a variable load. The avalanche detection circuit, which can be a comparator, senses the avalanche and reacts to it by generating a driver signal for the switch or the variable load. The switch then turns on and biases the SPAD out of its breakdown region by lowering the voltage across the SPAD. In the variable load case, the driver signal generated by the detection circuit increases the load at the dynamic node of the SPAD to quench the SPAD sooner. Fig. 1.15 shows a symbolic schematic of an AQ circuit.

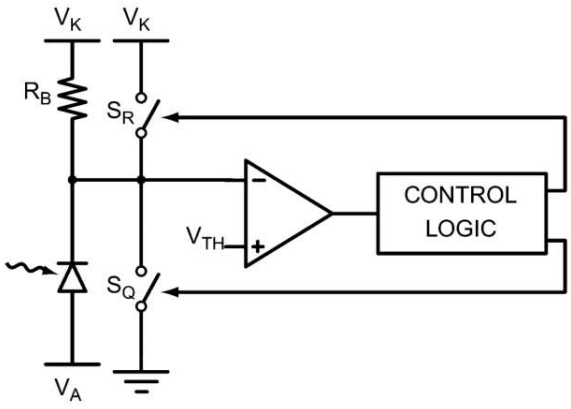


Fig. 1.15 Schematic of an AQAR circuit. S_Q and S_R are AQ and AR switches, respectively.

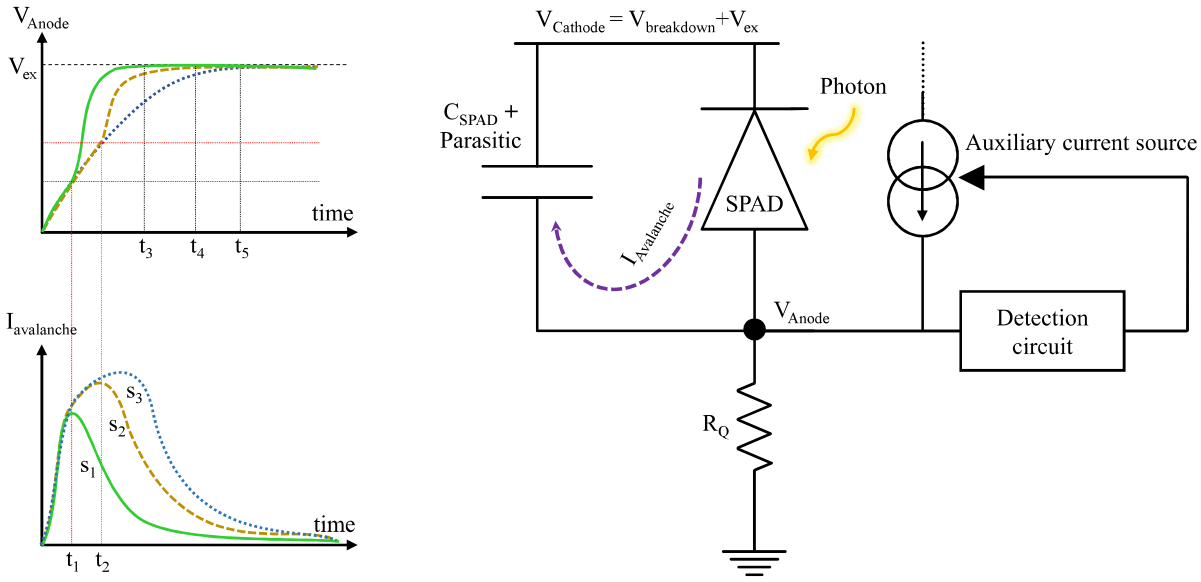


Fig. 1.16 Avalanche current reduction through AQ circuit. Different detection times results in different avalanche currents (S_1, S_2, S_3) and different quenching times (t_3, t_4, t_5).

In an AQ circuit, if the SPAD is quenched sooner than in the pure PQ mode, the avalanche current would be lowered. Lower current in the SCR means less trapping probability, which results in less afterpulsing effect without increasing the dead time. This interesting feature of AQ circuits prompts several research on implementing such a circuit to reduce the noise of the SPAD and mitigate the tradeoff between the afterpulsing effect and the maximum photon count rate of the SPAD. The efficiency of an AQ circuit is determined by its contribution in providing the required charge for charging the SPAD junction capacitance to bias the SPAD out of the breakdown region. In this case one can consider the switch as an auxiliary current source which provides this charge, see Fig. 1.16. Thus, the amount of current that it provides, and the moment that it starts its contribution are two determining factors in the afterpulsing reduction. The more the auxiliary current is, the less the avalanche current and consequently the less the afterpulsing effect. A higher auxiliary current is highly dependent on the time for the switch to be turned on. If the switch starts to conduct at the end of the quenching phase, it cannot speed up the quenching process and help the SPAD in providing the required charge for quenching. Thus, to have an efficient quenching for the afterpulsing reduction, it is of the highest importance to detect the avalanche at its very early stages and react to that as soon as possible. It can be concluded that the efficiency of an AQ circuit in the afterpulsing reduction strongly depends on the speed of its avalanche detection circuit.

Fig. 1.16 shows how a sooner avalanche detection results in a faster quenching and less avalanche charge.

1.3.3. *State-of-the-Art Active Quenching Circuits*

The first integrated AQ circuit, aiming afterpulsing reduction through the avalanche current decrease, has been presented in [14]. Fig. 1.17 shows the schematic of the proposed AQAR circuit in this work. In the standby mode the SPAD is biased in the breakdown region, and a signal from the monostable turns the reset switch (S_{RESET}) and feedback switch ($S_{FEEDBACK}$) off and on, respectively. When an avalanche occurs, the PQ phase starts through R_B which implements the PQ resistor, and thus the cathode voltage (IN) increases. This voltage increase is sensed through the sensing stage realized by MOSFET P_1. Thus, the voltage at node (A) is increased in a positive feedback loop and turns the quenching switch (S_{QUENCH}) on. Now the AQ phase is started till it completely quenches the SPAD, see Fig. 1.18. A delay block controls the hold off time from 50-500 ns. After a certain hold off time, the monostable turns $S_{FEEDBACK}$ off and S_{RESET} on to start the AR phase. After the reset phase, once again the monostable turns the S_{RESET} off and $S_{FEEDBACK}$ on to prepare the SPAD for a new event.

The proposed circuit is fabricated in a high voltage 0.8 µm CMOS technology and occupies an area of $2mm^2$ (pads included). The AQAR circuit detects the avalanche in less than 12 ns and quenches it in 25 ns for a 20 V of V_{ex}. The reset phase takes about 20 ns. The maximum count rate of the pixel is 20 MC/s, and the circuit dissipates 20 mW of power in the quiescent mode.

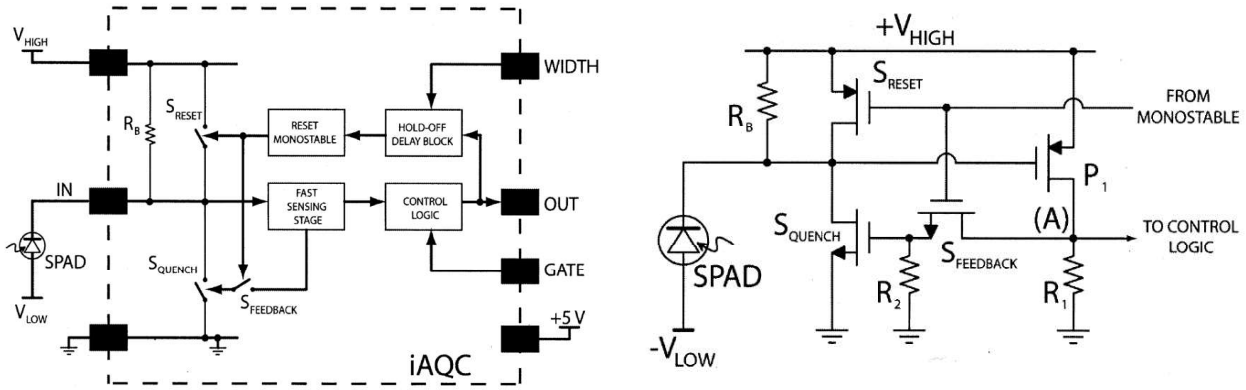


Fig. 1.17 (Left) Symbolic view of the proposed AQAR circuit in [14], (Right) Schematic diagram of the input sensing stage.

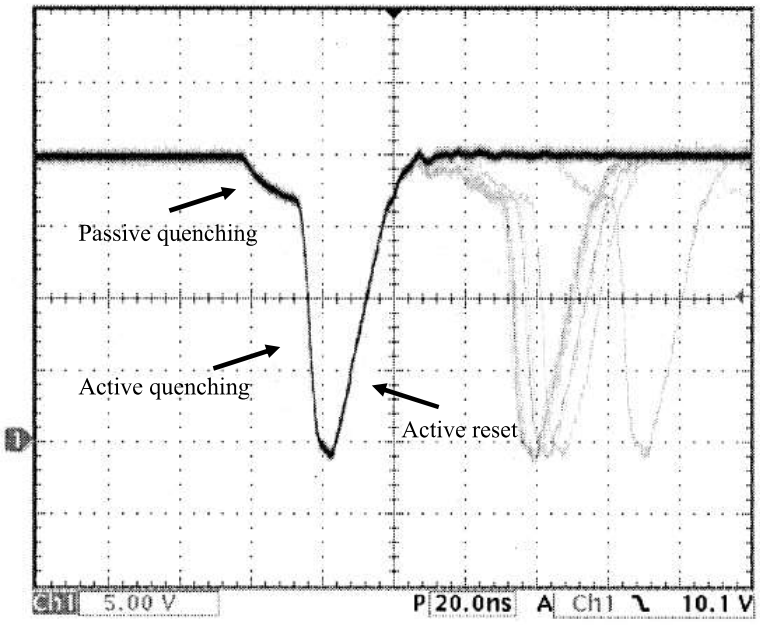


Fig. 1.18 Cathode voltage (IN) transitions during an avalanche. Voltage scale: 5 V/div, time scale: 20 ns/div [14].

One of the solutions to reduce the AQAR circuit area is to realize all the circuit by NMOS transistors to avoid implementing isolated N-wells for PMOS transistors.

Fig. 1.19 shows an AQAR circuit presented in [142]. In this circuit three different paths for quenching exist: R_{POL} which realizes the passive quenching resistor, a FAST PULSER unit which gates on and off the SPAD, and a wideband pnp transistor connected to anode which actively quenches the SPAD. The authors call the gating in this paper "AQ," and the AQ realized by the pnp transistor Fast AQ "FAQ." Occurrence of an avalanche increases the anode voltage thanks to the PQ resistor R_{POL}. This voltage increase is sensed by a comparator which drives the Gating circuit ("AQ") and FAQ pnp. Now the AQ phase is started till the end of the quenching phase. After a certain hold-off time the RESET npn transistor reset the SPAD for another avalanche detection. The right side of the circuit is almost identical to the left side, which is the main branch of the circuit. On the right side a capacitor imitates the SPAD junction capacitor, and the rest of the circuit components exist to reproduce the parasitic at the SPAD anode to generate a precise reference voltage for the sensing comparator. Bipolar technology is used to speed up the quenching process. The avalanche charge reduction for different quenching methods in comparison with the PQ mode is presented in Fig. 1.20 (a). In the best case (FAQ with 3 V V_{ex}), the avalanche charge

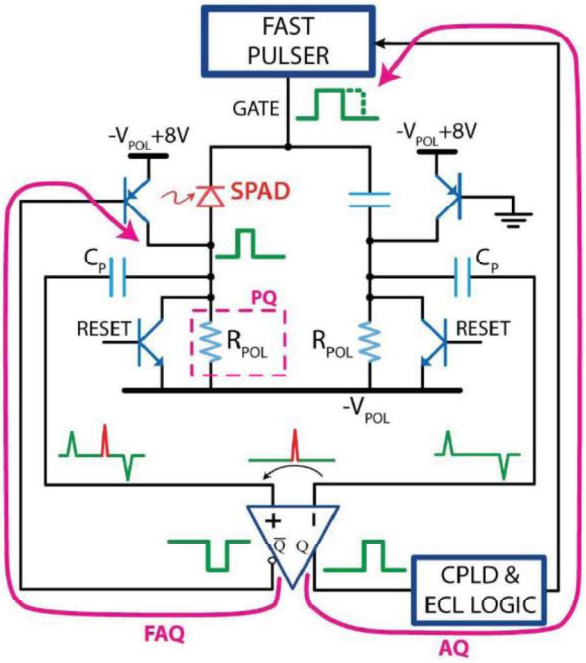
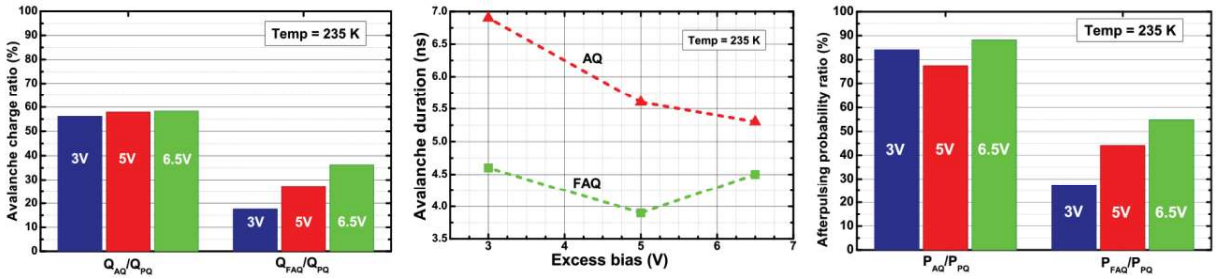


Fig. 1.19 Quenching and reset front end presented in [142] composed of passive quenching (R_{POL}), active quenching (FAQ), Gating (AQ), and reset circuits.



(a) (b) (c)

Fig. 1.20 (a) Estimated avalanche charge with active quenching (Q_{AQ}) and fast active quenching (Q_{FAQ}) normalized to the passive quenching case (Q_{PQ}). (b) Dependence of quenching time on the excess bias voltage in different quenching circuits. (c) Estimated afterpulsing probability with active quenching (AQ) and fast active quenching (FAQ) circuits, normalized to the passive quenching case [142].

is about 20 % of the avalanche charge in PQ mode, which means 80 % charge reduction. This is due to the shortened quenching time in the AQ modes in comparison with the PQ mode. Fig. 1.20 (b) shows the quenching time for the AQ modes. The FAQ mode presents the fastest quenching time of about 4 ns. Fig. 1.20 (c) concludes that a shortened quenching time results in a reduced avalanche charge which ends in a less afterpulsing probability. In the best case (FAQ with 3 V V_{ex}) the afterpulsing probability is reduced by 75 % in comparison with the pure PQ mode.

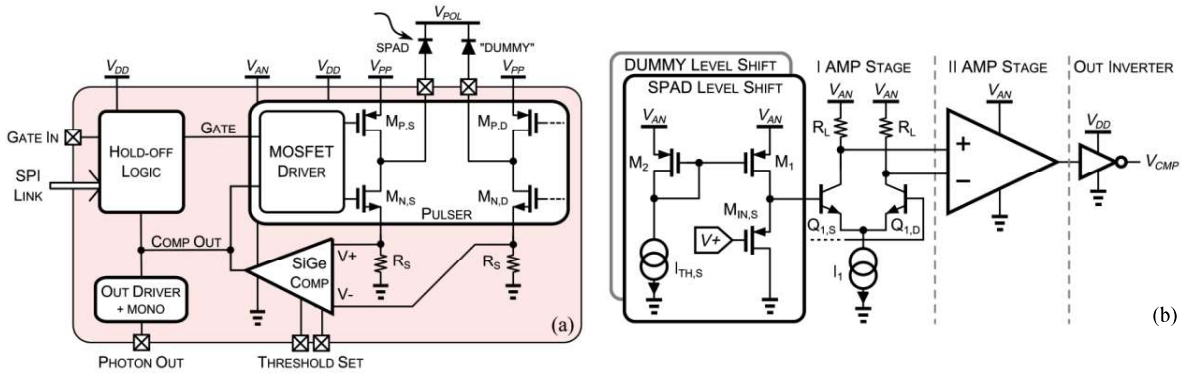


Fig. 1.21 (a) Block diagram of the quenching and gating circuit presented in [143]. (b) Schematic of the comparator in (a).

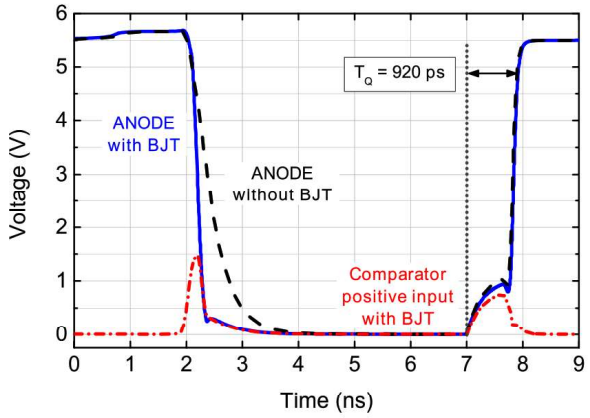


Fig. 1.22 Simulated waveforms at SPAD anode and at comparator positive input at 230 K. Anode voltage is reported with and without Q_S and Q_D BJT transistors that speed up the reset transition [143].

Almost the same circuit is presented by the same team in [143] and [144]. The proposed circuit is fabricated in a 0.35 µm SiGe BiCMOS technology with bipolar transistors with unity gain frequency up to 40 GHz. Fig. 1.21 (a) shows the block diagram of the circuit. Here the quenching is done by the MOSFET $M_{P,S}$ and the avalanche is sensed through a SiGe comparator with 1 GHz bandwidth. The comparator drives the quenching MOSFET and a delay block (HOLD-OFF LOGIC) to control the hold-off and reset time. Fig. 1.21 (b) shows the schematic of the comparator composed by four stages: the first stage is a CMOS level shifter to set the level of the input signals in the common mode range of the following stage. The second stage is a bipolar differential amplifier like the third stage. The last stage is an inverter which generates digital signal to drive the logic parts of the circuit and the quenching MOSFET. According to the simulation results, the avalanche is detected in about 800 ps and quenched in less than 1 ns, see Fig. 1.22. However, in

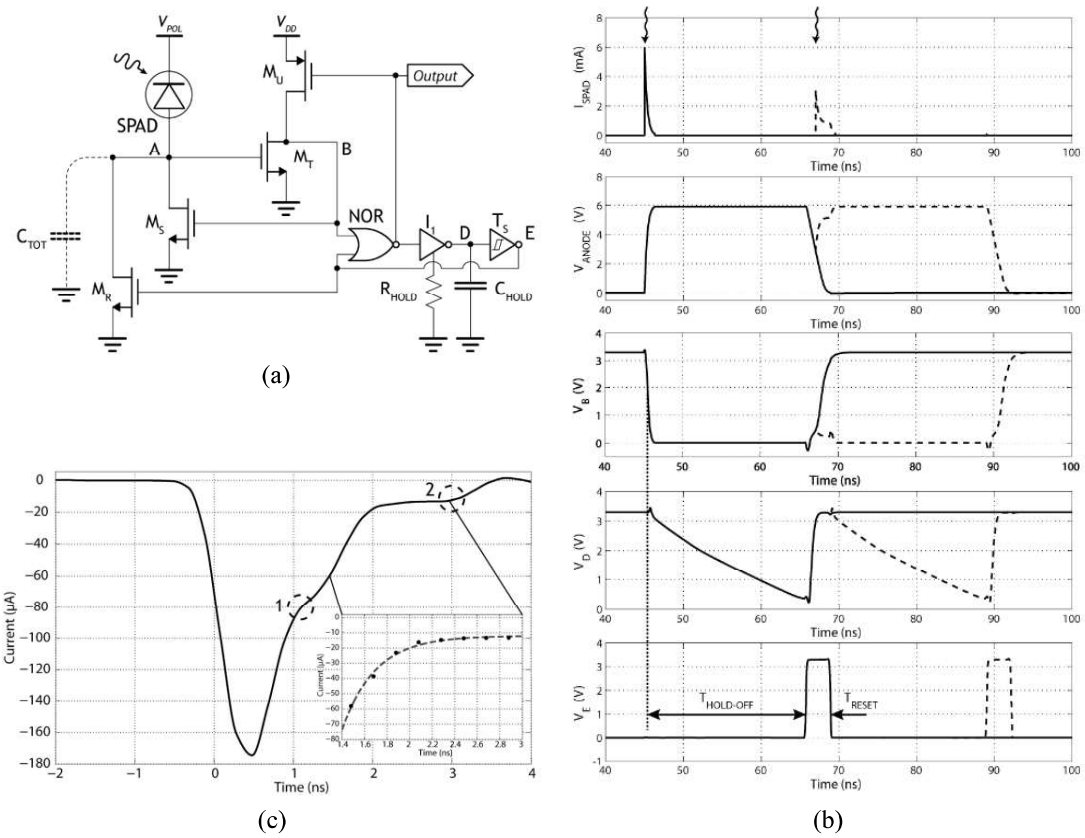


Fig. 1.23 (a) Schematic of the Variable Load Quenching Circuit (VLQC) presented in [145]. (b) timing diagram of the VLQC. (c) Cathode current of the circuit in (a) during an avalanche.

the measurements, the quenching time is reported to be less than 2 ns. The circuit area is 100 µm × 370 µm, and the power consumption is around 30 mW at 1 Mc/s.

In [145] a novel quenching circuit is introduced named Variable Load Quenching Circuit (VLQC). Fig. 1.23 (a) shows the schematic of the proposed VLQC. In this circuit the anode is connected to a transistor (MS) which functions as a PQ resistor, as well as an AQ element. Ms has the minimum size to reduce the parasitic capacitance at anode, and more importantly, to realize a high impedance path for PQ. In the standby mode this MOSFET is on and in the triode region. Thus, at the onset of an avalanche, the anode voltage rises due to the PQ implemented by Ms. This voltage increase is sensed through MT and falls down the voltage at B which is the gate of MS. Therefore, the resistivity of MS is increased until it completely turns off and finally quenches the SPAD. After a hold-off time set by RHOLD and CHOLD, the Schmitt trigger TS turns on the reset MOSFET MR and starts the reset phase. Anode voltage falls (node “A”) and MT turns off. NOR

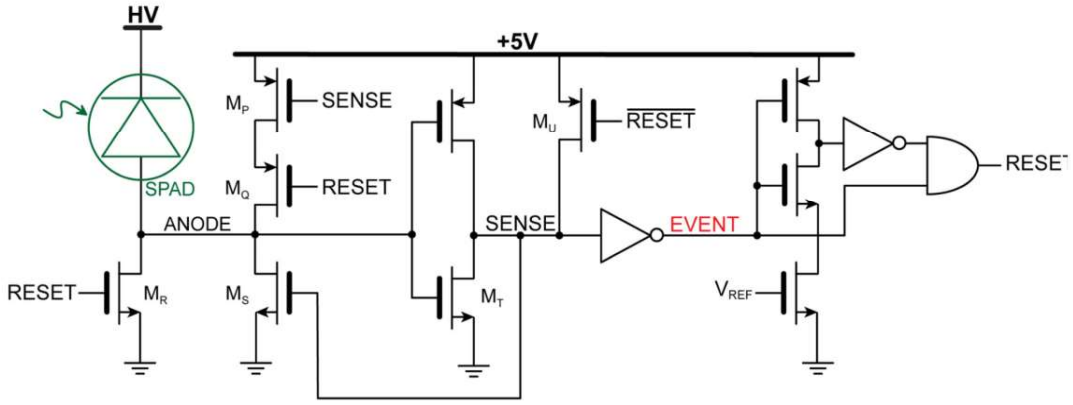


Fig. 1.24 Improved version of VLQC with an AQ switch (M_P) presented in [147].

gate forces MU to increase the voltage at node "B" up to VDD. Now the Schmitt trigger turns MR off to end the reset phase and make the SPAD ready for a new event. The timing diagram of the circuit is presented in Fig. 1.23 (b).

The proposed circuit is fabricated in a 0.35 µm CMOS technology and occupies an area of 28 µm × 24 µm. The maximum count rate can reach 50 Mc/s and the current consumption is 83 µA from V_{DD}. The total quenching time is about 3 ns, of which the first 1 ns is the AQ duration, see Fig. 1.23 (c). Unfortunately, no comparison between the proposed VLQC and a PQ circuit has been presented to evaluate the efficiency of VLQC in avalanche current and afterpulsing reduction.

Even though, a higher quenching resistor results in a faster quenching, and the avalanche is quenched through a positive feedback loop between M_T and M_S, however, calling VLQC an AQ circuit cannot be completely true, since, same as in PQ circuit, this is finally the high impedance path that quenches the SPAD.

Two developed versions of the same circuit as [145] are presented in [146] and [147]. Fig. 1.24 presents the AQAR circuit presented in [147] by the same team as in [145]. The base of the circuit is the VLQC which is already presented. Here, the MOSFET M_P is added to assist the quenching process and realize the AQ switch. Also, M_T is used in the inverter composition. The hold-off time is set by V_{REF} which controls the current of a current starved inverter. When an avalanche starts, the anode voltage rises thanks to the M_S which is in the ohmic region and implements the PQ resistor. The voltage rise is sensed by the inverter which includes M_T. Therefore, the voltage at node "SENSE" increases and turns M_P on. Now the real AQ process starts till completely quenches

the SPAD. Meanwhile, the EVENT signal rises and after a delay set by V_{REF}, turns on the reset transistor M_R and M_U, and turns off M_Q. The reset phase starts, however, there is no control over the reset time. As soon as the anode voltage passes the threshold of the first inverter realized by M_T, before reaching the ground, the RESET signal falls and turns the reset MOSFET M_R off. Thus, the reset phase ends while the SPAD is not fully reset. The maximum reset time is equal to the propagation delays of the inverters and the AND gate. An uncomplete reset makes the SPAD operate in a lower excess bias voltage and thus increases jitter and lowers PDP. Unfortunately, no timing diagram or a simulated waveform is presented by the authors to evaluate the reset functionality of the circuit. The authors have claimed that the quenching time is about 100 ps for 5 V of V_{ex}. However, no evidence, neither a simulation result nor a measured waveform is presented. The circuit is fabricated in a 160 nm BiCMOS DMOS (BCD) technology, and its area is 37 × 35 μm^2. The used transistors are thick oxide MOSFETs. The dead time is around 930 ps which can result in a maximum count rate of more than 1 Gc/s. An extremely low APP of 0.14 % is reported. However, no comparison is made between the AQ circuit and pure PQ mode of operation to see how much of this low APP is due to the AQ circuit contribution in the quenching process, and how much is related to the low intrinsic noise of the SPAD.

Fig. 1.25 shows the AQAR circuit presented in [146]. The circuit is very similar to the previous circuit: MOSFETs M_S and M_R have the same functionality as in the previous circuit. M_T is a diode connected MOSFETs which increases the impedance of M_S for a faster PQ and M_Q realizes the AQ switch. The quenching is done through positive feedback between INV1, M_S, and M_Q. The reset has a separate path which is controlled by a Schmitt trigger, thus better control over the reset time is achieved. The circuit is fabricated in 0.18 µm CMOS technology with an area of 306 μm^2. The quenching time is 700 ps (simulation results) and the maximum count rate is 200 Mc/s. At 4 ns hold-off time 0.75 % APP is reported. Here also no comparison between the AQ circuit performance and the PQ mode is made which remains the effectiveness of the proposed AQAR in afterpulsing reduction ambiguous.

In [113] and [148]–[150] an almost same team has presented the same quenching strategy with slight changes at circuit level. Fig. 1.26 (a) shows the circuit presented in [113]. The main part of the circuit is a fast comparator (“AMP”) which quenches the SPAD in a positive feedback loop.

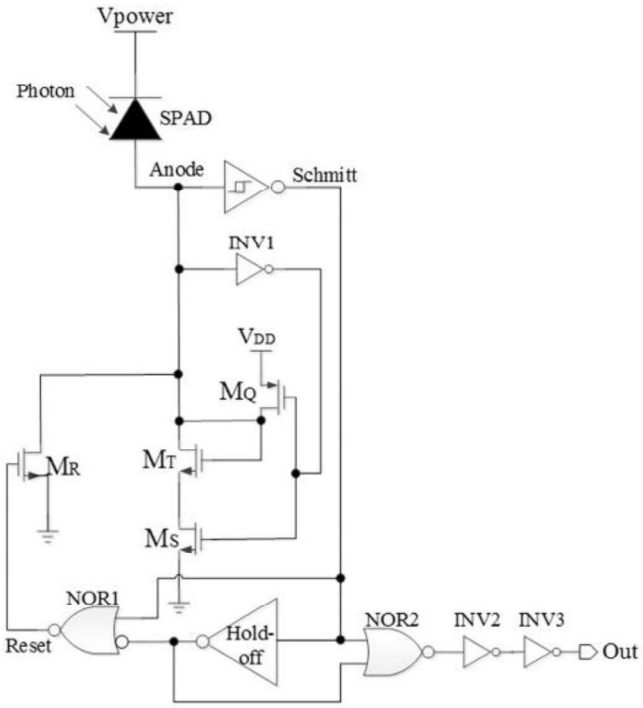


Fig. 1.25 Schematic of the proposed quenching circuit in [146].

Fig. 1.26 (c) shows the timing diagram of the circuit. In the quiescent mode, the cathode voltage is almost equal to V_{SUP+} (3.3 V), neglecting the voltage drop over M_2 and M_3. The output of the Schmitt trigger ST_1 is high, thus M_6 has charged C_{TQ} (node "B"). The output of the second Schmitt trigger ST_2 falls (node "C"), which turns on "AMP" to detect an avalanche. When an avalanche occurs (t_0), PQ through M_3 and M_2 starts and decreases the cathode voltage and so the non-inverting input of the comparator. The voltage at this input reaches the reference voltage of the comparator V_{REF} at t_1. From t_1 to t_2 it takes the comparator to start the AQ phase and finally at t_3 the quenching phase is finished by lowering the cathode voltage down to V_{SUP-}. Now the output of ST_1 rises and turns off M_6. Thus, C_{TQ} is discharged through M_5. Node "B" falls, and node "C" rises. After a certain delay, the comparator turns off and M_7 turns on (t_4). Now the reset phase starts until the fully recharge of the SPAD at t_5. By reaching to the threshold of ST_1, at t_6 once again the comparator is ready to detect an avalanche.

The comparator is a 4-stage amplifier, see Fig. 1.26 (b). Two differential pairs (M_8-M_{13}) followed by a PMOS common source stage (M_{14}). The Last stage is a cascode stage realized by M_{15} and M_{C5}. M_{I1}-M_{I8} are biasing and pre-biasing current sources. M_{C1}-M_{C9} are cascode transistors.

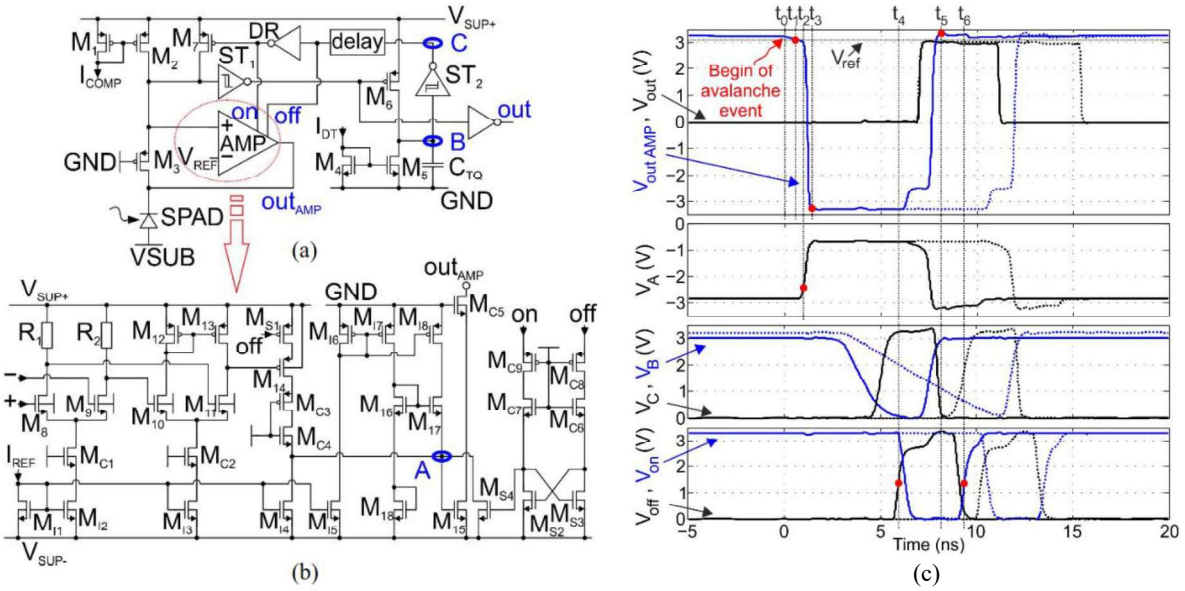
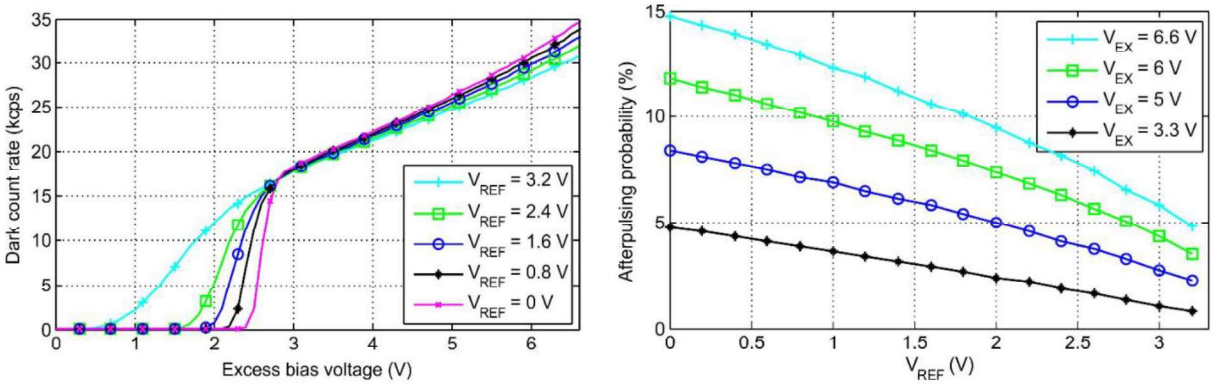


Fig. 1.26 (a) Proposed AQAR circuit in [113]. (b) Schematic of the AMP comparator used in (a). (c) timing diagram of the AQAR circuit.

The last stage (M_{15}) is pre-biased M_{16} and M_{17} to have a faster response. M_{S2} and M_{S3} form a latch that controls M_{S4}, which along with M_{S1}, partially turn off the amplifier. However, the two first stages and the pre-biasing current mirrors still consume current. It results in 4.8 mW of power consumption in the idle mode and 10.03 mW at 100 Mc/s.

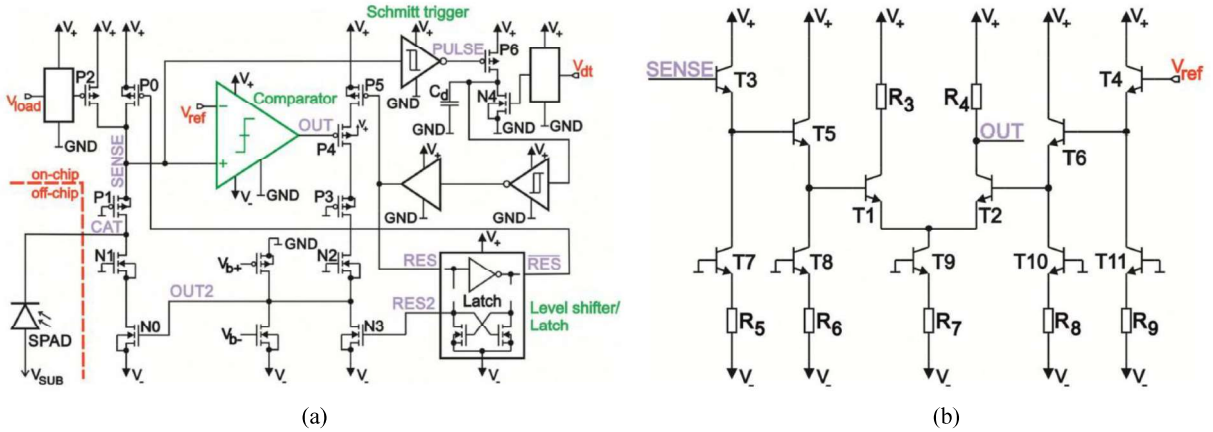
The circuit is fabricated in a 0.35 µm CMOS technology and occupies an area of 130 × 134 µm^2. With V_{ex} = 6.6 V, in the fastest and most sensitive case (V_{REF} = 3.2 V and so a threshold of 100 mV for the comparator), the total quenching time is 1.04 ns (simulation results). Fig 1.27 (a) presents the DCR versus excess bias voltage of the used SPAD in this work, measured by the proposed AQAR circuit at different V_{REF}. A higher V_{REF} means a higher detection sensitivity and a lower threshold for the comparator. It shows that a less sensitive circuit (low V_{REF}) cannot detect avalanches with lower V_{ex}, while the same SPAD shows a higher DCR value with a more sensitive circuit (higher V_{REF}). This result is an important proof of the underestimation of the SPAD noise in the measurement with a low sensitivity avalanche detection circuit. In other words, the measured noise of a SPAD can strongly depend on the sensitivity of the avalanche detection circuit. Fig. 1.27 (b) shows the APP at different excess bias voltages for different sensitivity levels. At 6.6 V of V_{ex}, in the highest sensitivity (V_{REF} = 3.2 V) the APP is 4.8 %, and in the lowest sensitivity (V_{REF} = 0 V) this value is 14.7 %. It means around 67 % relative improvement in the APP.



(a) (b)

Fig. 1.27 (a) DCR versus excess bias voltage at different sensitivity level. (b) APP versus sensitivity level at different excess bias voltage. Dead time is 9.5 ns [113].

In [150] the same circuit is redesigned by employing bipolar transistors. Fig. 1.28 (a) shows the AQAR circuit diagram. It is similar to the circuit in Fig. 1.26 (a). The first two stages of the "AMP" in the previous circuit are shown by "Comparator" in the schematic. The third stage is single ended common source amplifier realized by P_4, and the last stage which acts as a quenching switch is implemented by N_0. The "comparator" is realized by bipolar transistors to speed up the quenching process. Fig. 1.28 (b) shows the bipolar comparator schematic: two bipolar level shifters placed in series at each input of a bipolar differential pair. In comparison with the CMOS version of the circuit in [113], the power consumption is increased by a factor of 4. At 40 Mc/s the bipolar AQAR circuit consumes 58 mW (simulation results). The circuit is fabricated in 3.3 V/ 0.35 µm BiCMOS technology and its area is twice the area of the CMOS version. At a V_{REF} = 3 V the active



(a) (b)

Fig. 1.28 (a) Schematic of the AQAR circuit presented in [150]. (b) Schematic of the bipolar comparator with level shifters.

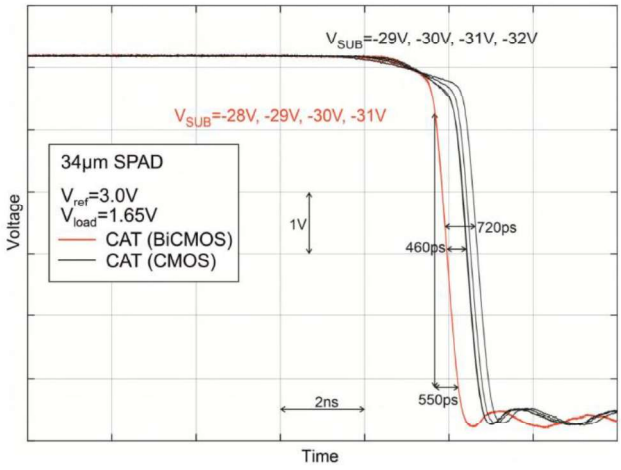


Fig. 1.29 Cathode voltage variations during an avalanche at different excess bias voltages for CMOS and bipolar comparators [150].

quenching phase starts sooner than the CMOS version, from 460 ps to 720 ps depends on the excess bias voltage. A higher V_{ex} results in a shorter reaction time of the comparator. However, the AQ phase slope is the same for both bipolar and CMOS versions (550 ps fall time at the cathode voltage). The total quenching time is slightly higher than 2ns. Fig. 1.29 shows the cathode voltage variations for bipolar and CMOS AQAR circuits at different V_{ex} and V_{REF} = 3 V.

In [148] the same circuit as in [113] is presented with the modifications in the timing control circuits. Fig. 1.30 shows the schematic of the proposed AQAR circuit, named triple voltage quenching circuit. The reason for this naming is 9.9 V of V_{ex} that the pixel can tolerate which is 3 times of the technology voltage headroom. To achieve this ability, many high voltage (5 V) MOSFETs (drawn in bold) are used in a cascode structure that increases the circuit area. The proposed circuit is marked by five parts from A to E. Part A is the quenching and reset switches (M_{Q1} and M_{R1}) with their cascode MOSFETs. Part B is almost the same comparator as in [113]. The third stage of the comparator is realized by an inverter instead of a common source stage. In [113] the active quenching switch is included in the four stages of the used comparator as the fourth stage, however, in this circuit a separate quenching switch is driven by the 4-stage comparator. A current conveyor, controlled by V_{PB} (“CC”), pre-biases the quenching switch M_{Q1} for a faster quenching. Since the quenching switch must be a very wide transistor, the pre-biasing can increase its leakage current. Part C turns off the quenching switch by pulling its gate voltage to the ground. Part D generates the reset signal and controls the reset time through V_{RT} and by driving the reset

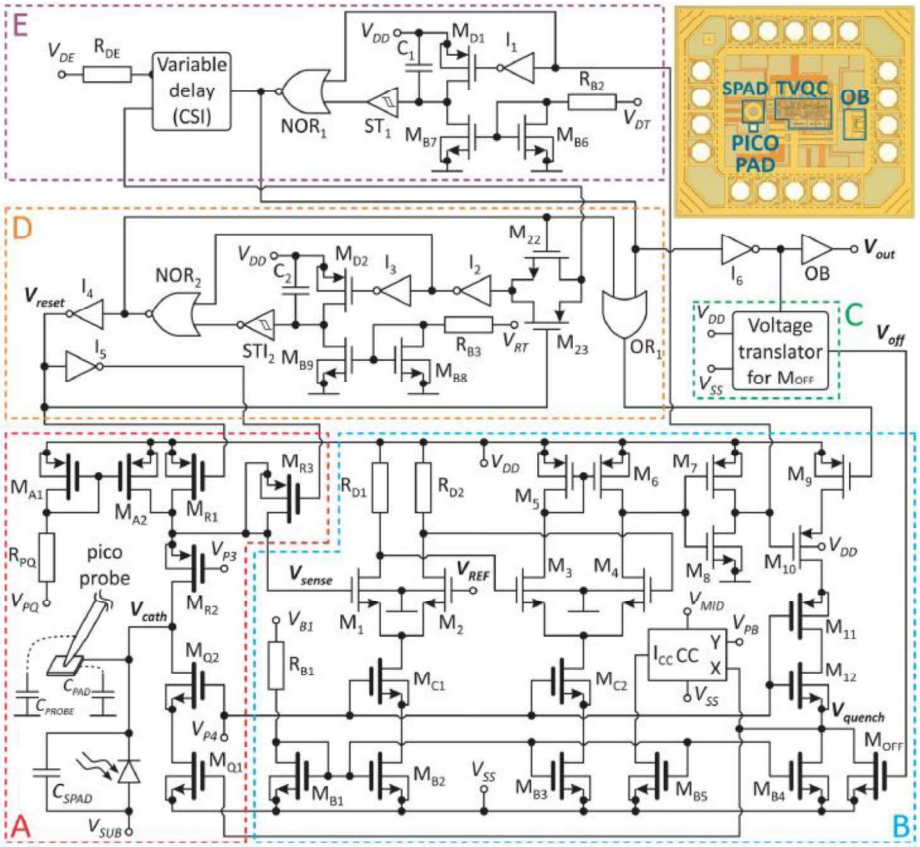
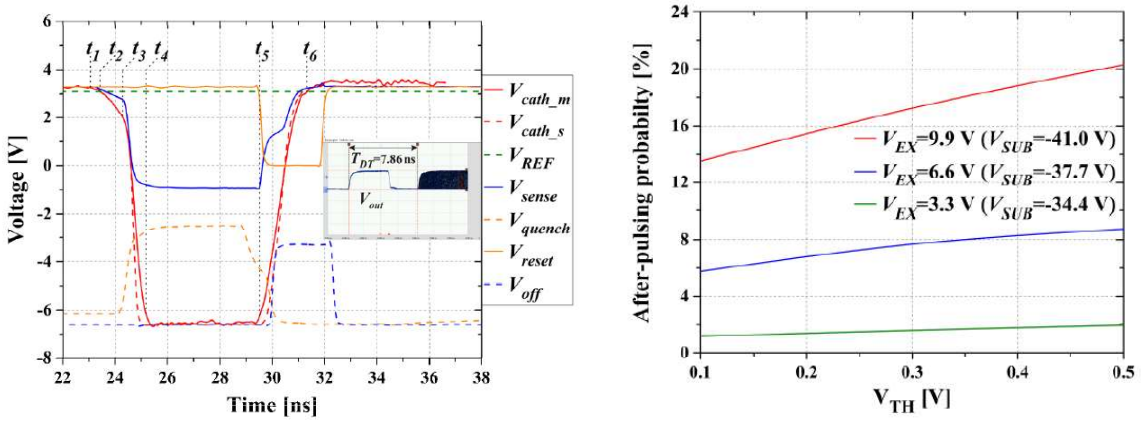


Fig. 1.30 Schematic of the proposed triple voltage AQAR circuit in [148] with micrograph of the chip.



(a) (b)

Fig. 1.31 (a) Operation principal of the presented circuit in [148] at its highest sensitivity (V_{th} = 100 mV). V_{cath_m} and V_{cath_s} are measured and simulated cathode voltages. (b) APP versus sensitivity level at different excess bias voltage. Dead time is 8 ns.

switch M_{R1}. Part E controls the hold-off time through V_{DT} by triggering the reset process through

a tunable delay ("CSI") controlled by V_{DE}. It also turns off the last stage of the comparator (M_{10}) by switching off M_9 for the reset phase. The circuit is fabricated in a 0.35 µm CMOS technology and occupies an area of 236 × 108 µm^2. The power consumption of the circuit with 8 ns dead time is 14.4 mW. The measured quenching time is claimed to be 1.7 ns, however due to the measured waveform presented in Fig. 1.31 (a), this time is almost 2 ns ($t_1 - t_4$). Fig. 1.31 (b) shows the APP versus the sensitivity (V_{TH} = 3.3 V - V_{REF}) of the comparator controlled by V_{REF}. In the best case, at V_{ex} = 9.9 V and 8 ns dead time, the highest sensitivity (V_{TH} = 100 mV) results in 13.4 % APP, while at the lower sensitivity of V_{TH} = 500 mV, the APP is 20.3 %. It means about 34 % relative improvement in APP.

1.4. Thesis Objective

At the beginning of this chapter, we discussed the noise sources of the SPAD and the solutions to reduce these noises. It is already explained how avalanche charge reduction can reduce the afterpulsing effect as a main noise source. However, as it is addressed previously, there are severe trade-offs between fill factor, PDE, maximum photons count rates, and the afterpulsing effect. A very few solutions to mitigate these trade-offs are costly, such as the 3D stacked IC technology. Besides, according to the state-of-the-art, the AQ circuits still suffer from several drawbacks in afterpulsing reduction, such as low sensitivity in the avalanche detection, long quenching time, high static power consumption, and large circuit area.

In this thesis, we aim to introduce the innovative AQAR circuits, which are sensitive enough to strongly reduce the afterpulsing effect, while consuming low power and small areas to mitigate the aforementioned trade-offs. Also, we investigate the monolithic 3D integration of SPAD with AQ circuits for the first time, which can remove the dependence of the fill factor and thus PDE on the area of the SPAD front end circuitry in a Back-Side Illumination (BSI) configuration. This can be an ultimate solution for the trade-offs between PDE, noise, and the count rate. To better understand this concept, we explain this technology in the following.

1.4.1. Monolithic 3D SPAD Pixel in FD-SOI CMOS Technology

The technology that is used in this thesis is a 28 nm Fully Depleted Silicon On Insulator (FD-SOI) CMOS technology. In Fig. 1.32 (a) a cross section of this FD-SOI technology is presented. The transistors are fabricated over a 25 nm thick Buried Oxide layer (BOX). This structure yields

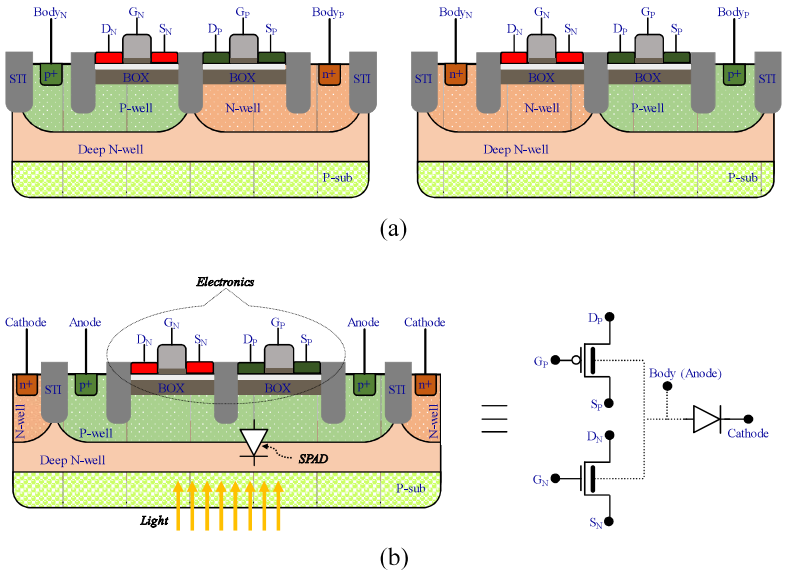


Fig. 1.32 Symbolic cross section of: (a) FD-SOI CMOS technology, regular well (left) and flip well (right), (b) mix of the flip well and the regular well in the FD-SOI CMOS technology and forming of the SPAD at the junction of P-well and the deep N-well which results in the intrinsic 3D pixel (left) and its schematic counterpart (right). The pixel is expected to be backside illuminated.

electrical isolation for the active devices over the BOX from their corresponding wells. In this technology, two different MOSFET families exist; Flipped well (FW) and Regular Well (RW). In FW MOSFETs, NMOS are fabricated in an N-well and PMOS in a P-well, while in the RW family it is the opposite [151]–[153]. It is possible to mix these two MOSFET categories and have both NMOS and PMOS in the same well as it is shown in Fig. 1.32 (b). One can note that the junction between P-well and deep N-well can be used to form a SPAD.

Based on this, the first SPAD in a 28 nm FD-SOI CMOS technology is presented in [154]. The SPAD is implemented in a commercial FD-SOI CMOS technology without any process modifications i.e., with the default standard junction doping profile. The SPAD is laid out without any Design Rule Checking (DRC) violations. It has a quasi-octagonal shape to respect DRC, and its diameter is 25 µm. These all have resulted in a promising albeit non-optimized SPAD. However, for the purposes of this paper, a non-optimized SPAD can demonstrate more clearly the effectiveness of the proposed AQAR circuits in the afterpulsing reduction. In this regard, this SPAD is chosen to evaluate the proposed AQAR circuit. Nonetheless, the optimizations to improve the FD-SOI SPAD parameters are presented in [155]–[159].

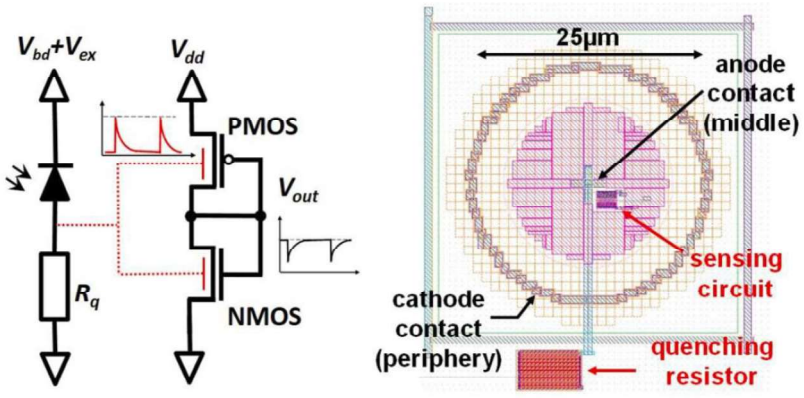


Fig. 1.33 (right) Schematic of a voltage divider as a sensing circuit for SPAD-FDSOI (the dashed red line represents the body-biasing, i.e., the indirect coupling between the SPAD anode and transistors). (left) layout view of the SPAD-FDSOI with the divider on top of the SPAD and also the quenching resistor.

Thanks to the insulating BOX layer, electronic devices can be placed over the SPAD while being electrically isolated from it. This forms a monolithic 3D SPAD pixel, see Fig. 1.32 (b). Despite the conventional 3D pixels, this pixel occupies one die and reduces the fabrication costs. The first attempts to realize this pixel are presented in [160], [161], in which two MOSFETs, realizing a divider as an avalanche detection circuit, are placed over the SPAD, see Fig. 1.33. Yet, the structure has more potential: integrating an ultra-fast AQAR circuit in the pixel, alongside a Back-Side Illumination (BSI), can drastically reduce or remove the dependence of the FF and PDE on the electronics.

Although a die thinning is mandatory to allow the BSI in the monolithic 3D SPAD pixel, the manufacturing process cost of the used FD-SOI CMOS technology is 10% less than the classical bulk CMOS technology [151].

In this thesis the focus is on the design and characterization of the first ultra-fast AQAR circuits which are compatible with the monolithic 3D SPAD pixel as well as conventional SPAD pixels.

Chapter 2
Body Biased Inverter Based Active Quenching Active Reset Circuit

2.1. Introduction

It has already been discussed that to decrease the afterpulsing effect without degrading the maximum count rate, an AQ circuit is mandatory. Such a circuit is composed of two main parts: avalanche detection or sensing circuit, and quenching circuit. The quenching circuit can be as simple as a switch or a current source that discharges the SPAD junction capacitance. The effectiveness of an AQ circuit in the afterpulsing reduction is determined by these two parts. A sensitive detection circuit can signal early stages of an avalanche to the quenching part, and the quenching part should start its contribution in the quenching process as soon as possible, before the SPAD itself completely quenches the avalanche. The quenching part should provide a portion of the required charge carriers that discharge the SPAD capacitance. The greater the contribution of the AQ circuit, the less the flow of avalanche current in the SCR, and thus less afterpulsing effect. Also, the sooner the AQ part starts to intervene, the more it can contribute. The time that the AQ part starts its auxiliary role is determined by the detection circuit. Therefore, it is crucially important to design a very fast and sensitive avalanche detection circuit to have an effective AQ circuit.

Another important consideration in designing an AQAR circuit is its area. The circuit added area increases the pixel area while the light sensitive area (active area) remains constant. This reduces the pixel fill factor and thus its PDE. Therefore, an AQAR circuit should be compact and area efficient.

In this chapter we introduce a new AQAR circuit in a 28 nm FD-SOI CMOS technology. Utilizing specific characteristics of FD-SOI CMOS technology, a very fast detection circuit is proposed, which detects the avalanche quickly and reduces the afterpulsing effect effectively. The circuit can be very compact and does not add considerably parasitic to SPAD.

2.2. Low Threshold Inverter as an Avalanche Detector

Despite their fame in digital electronics, inverters are also used in analog circuits, especially in more scaled technology nodes. They are used as comparators, output buffer stages, and even analog amplifiers [162]. Thanks to their compact area and low power consumption, inverters are also used for the avalanche detection as it is reported in the literature [163], [164]. Fig 2.1 shows the schematic of a SPAD pixel utilizing an inverter as a detection circuit. When an avalanche occurs, the anode voltage rises, till it reaches switching threshold of the inverter. At this point inverter output changes and flags the photon arrival. The time that it takes to detect an avalanche through an inverter can be characterized by the high to low propagation delay (t_{pHL}) or switching delay.

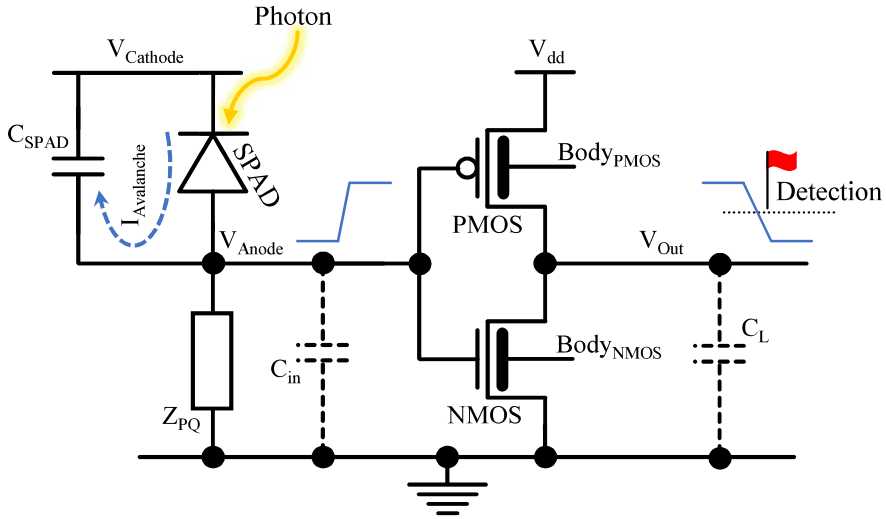


Fig. 2.1 Schematic of a SPAD pixel featuring an inverter as the avalanche sensing circuit.

The propagation delay is the interval between the time that input signal of an inverter is at its 50 % and the time that its output signal reaches to its 50% [165]. Even though, in the avalanche detection case, 50 % of input signal swing translates to a slow detection, yet the propagation delay can give an insight to the efficiency of an inverter in detecting an avalanche. Based on this definition, for the very fast transient signals where the rise and fall time of the input signal are within the range of those of the output signal, finding a precise equation for the propagation delay is not trivial. However, it is possible to form a proportional relationship between the switching delay and circuit parameters:

$$t_{pHL} \propto \frac{C_L \cdot V_{dd}}{(\frac{W}{L})_n \cdot (V_{dd} - V_{th_n})} \tag{2.1}$$

where C_L is all the parasitic capacitances at the output node of the inverter, V_{dd} is the power supply, $(\frac{W}{L})_n$ is the width over length ratio of NMOS, and V_{th_n} is the threshold voltage of NMOS. The shorter this time, the sooner the avalanche is detected, and thus the more efficient is the AQ circuit. According to this relationship, to have a short delay, it is required to properly size the inverter while keeping C_L minimum.

2.2.1. *Delay Reduction Through Inverter Sizing*

Relationship 2.1 shows that a wider NMOS can result in a shorter delay. Increasing NMOS width, while keeping the one of PMOS constants, decreases the switching threshold of the inverter by the following equation:

$$V_M = \frac{V_{dd} - |V_{th_P}| + V_{th_N} \cdot \sqrt{\frac{\mu_N \cdot (\frac{W}{L})_N}{\mu_P \cdot (\frac{W}{L})_P}}}{1 + \sqrt{\frac{\mu_N \cdot (\frac{W}{L})_N}{\mu_P \cdot (\frac{W}{L})_P}}} \tag{2.2}$$

where V_{th_N} and V_{th_P} are the threshold voltages of NMOS and PMOS. μ_N and μ_P are the mobilities of electrons and holes, respectively, and $(\frac{W}{L})_N$ and $(\frac{W}{L})_P$ are the width to length ratios of NMOS and PMOS. This is shown in Fig. 2.2 (a), where Voltage Transfer Characteristic (VTC) of an inverter is pushed back by increasing NMOS width. Therefore, at a lower input voltage (V_{in}) the output state changes, which in the time domain means faster detection.

However, increasing the size of its transistor also increases the load capacitance (C_L) of the inverter. Regarding relationship 2.1, this increases the delay. Thus, there is an optimum width that results in the minimum delay and increasing the width further than that, increases the delay. This is observable in Fig. 2.2 (b), where a 100 ps ramp is applied to an inverter for different NMOS widths. Here we define another delay indicator that is more intuitive in the avalanche detection

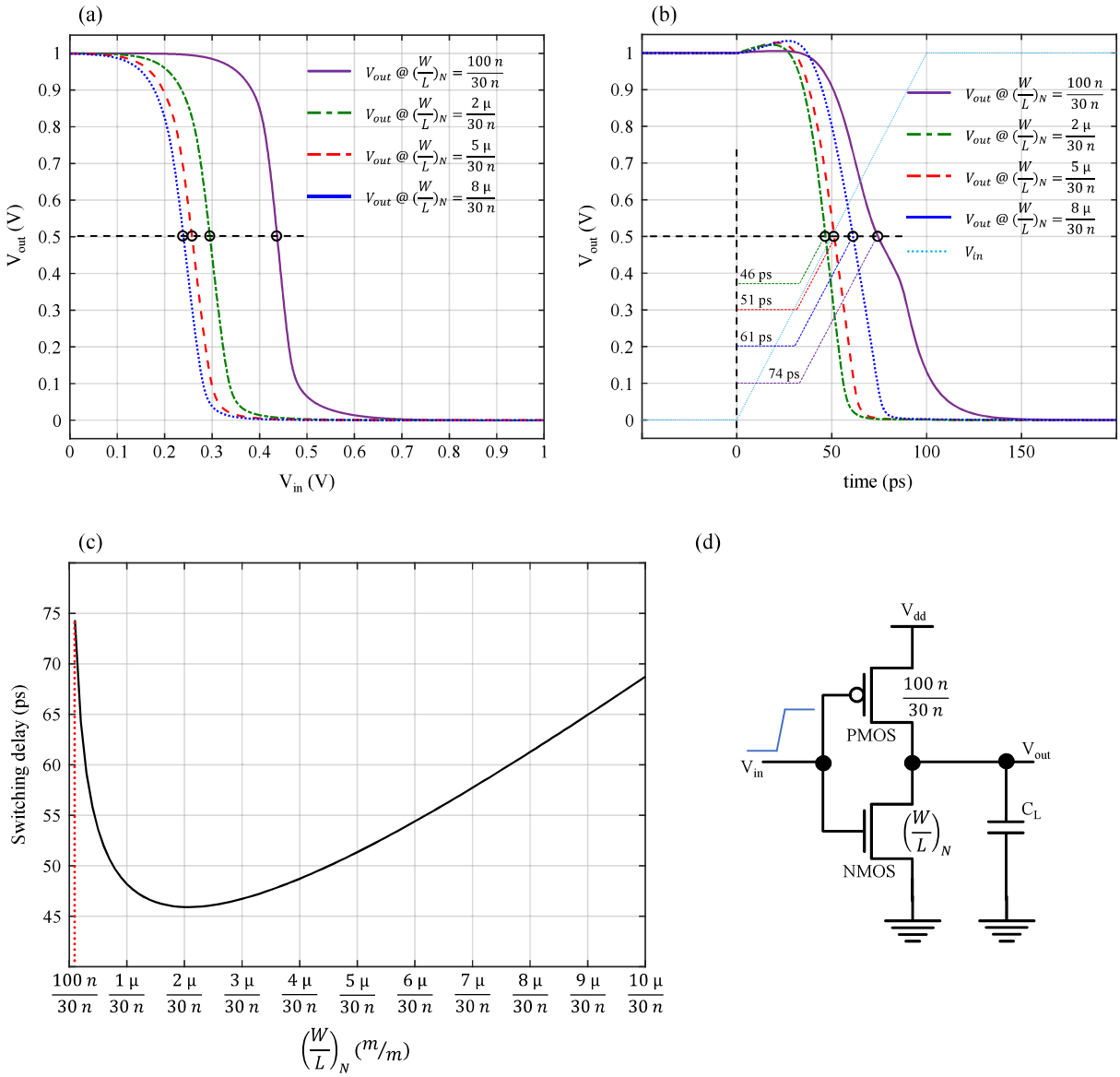


Fig. 2.2 Reducing switching delay of an inverter through sizing of its transistors. (a) VTC of the inverter at different NMOS sizes. (b) Transient response of the inverter for a 100 ps input ramp at different NMOS sizes. (c) Switching delay of the inverter versus NMOS size. (d) schematic of the tested inverter.

case. The switching delay is defined as the time interval between the beginning of the rise of the input signal and the time that the output signal reaches the switching threshold of a (following) standard inverter gate. Fig. 2.2 (c) shows the switching delay variations versus $(\frac{W}{L})_N$. A maximum 40 % (~ 30 ps) delay reduction is obtained by 20 times increasing the size of NMOS.

Besides the parasitic capacitances, the area consumption is another limiting factor in this method. The added area for the purpose of delay reduction, increases the inactive pixel area and therefore degrades the fill factor and PDE, and increases the pitch.

2.2.2. *Delay Reduction Through Body Biasing*

Another solution to reduce the delay, regarding the relationship 2.1, is decreasing the NMOS threshold voltage V_{th_N}. Equation 2.2 also propose to decrease V_{th_N} and increase the PMOS threshold voltage V_{th_P} to lower the switching threshold of the inverter. However, in standard CMOS technology, this is not a very practical solution. As it is shown in Fig. 2.3 (a), in the bulk CMOS technology, there are parasitic diodes between P-well and N-well, and between each well and Drain and Source of inside wells. These diodes limit the maximum applicable body voltage to their forward bias voltage, which is typically about 0.7 V.

In the FD-SOI CMOS technology an insulator (BOX layer) separates the MOSFETS and other devices from their corresponding wells and creates a capacitive link between the body and the channel of MOSFETs, see Fig. 2.3 (b). This structure yields two advantageous features: 1) the parasitic diodes cannot be formed anymore between Drain/Source and wells 2) the BOX layer is in fact a second gate or a back gate. These features allow to apply body biasing to the MOSFETs up to the maximum tolerable voltage of the buried oxide (BOX). In the used FD-SOI CMOS technology, the thickness of the back oxide is three times of the front gate oxide. Therefore, the back gate can tolerate three times of the technology voltage headroom (3×1 V). Despite the square root relationship in the bulk CMOS technology, the threshold voltage depends almost linearly on the body voltage in the FD-SOI CMOS technology [166]. This will be further discussed in chapter 4. The rate of this dependence, called body biasing factor (α), is about 80 mV/V, i.e., 1 V body voltage variation results in 80 mV threshold voltage variation. Therefore, a maximum threshold voltage modulation of about 240 mV is achievable in this technology.

To reduce the switching threshold of the inverter a positive voltage should be applied to the bodies of NMOS and PMOS. This requires a well separation through a deep N-well layer. Consequently, the circuit area is increased which is not desirable for a quenching circuit. However, in the FD-SOI CMOS technology there is a solution for this problem. As it is depicted in

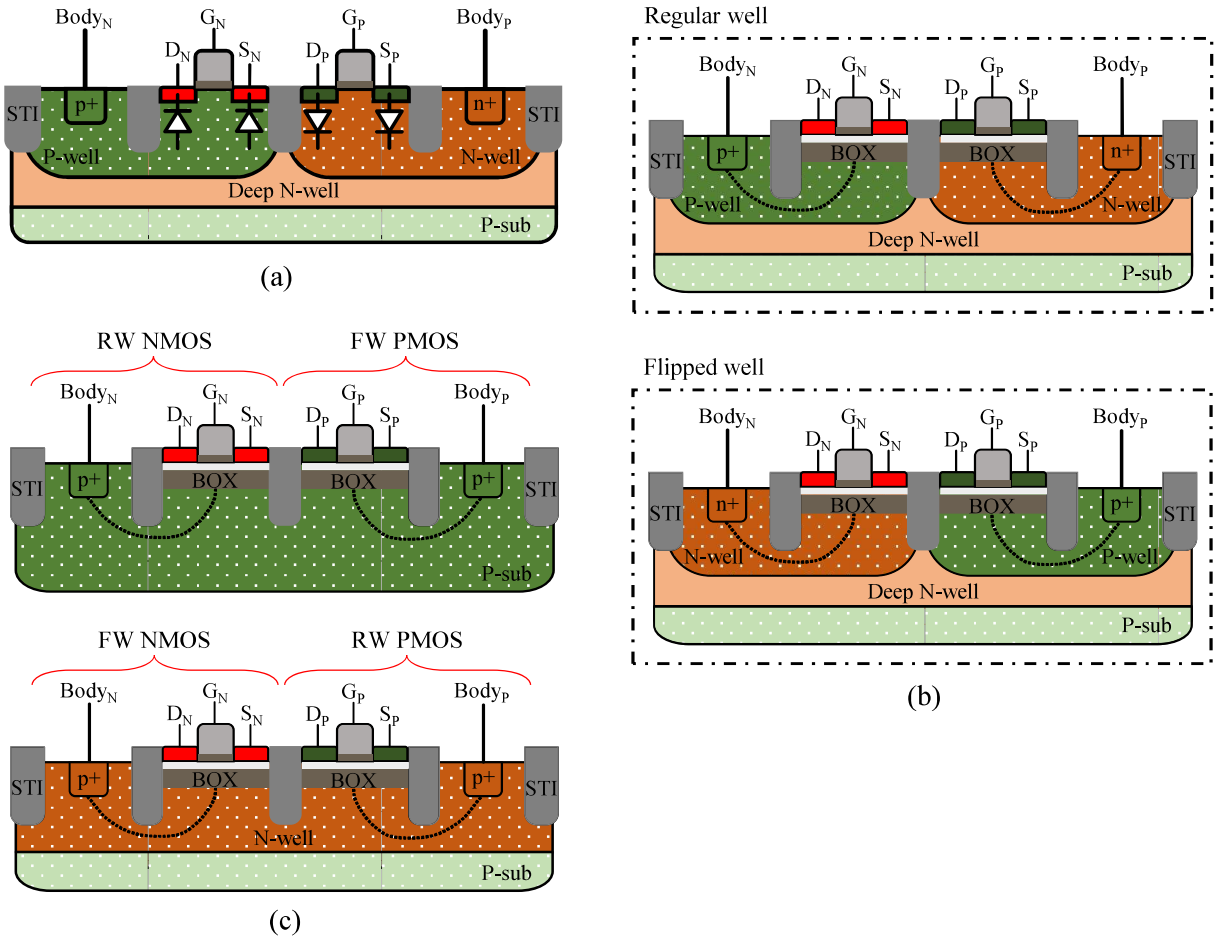


Fig. 2.3 (a) Symbolic cross section of a standard CMOS technology featuring a deep N-well layer. (b) Symbolic cross sections of the used FD-SOI CMOS technology. Top: regular well family, bottom: flipped well family. (c) Combining the two families of FD-SOI technology in a same well. Top: regular well NMOS and flipped well PMOS in a shared P-well, bottom: flipped well NMOS and regular well PMOS in a shared N-well.

Fig. 2.3 (b), there are two families of MOSFET in this technology: Regular Well (RW) and Flipped Well (FW). In the flipped well family, NMOS is implemented in an N-well and PMOS in a P-well, while in the regular well this is the reverse. By respecting the electrostatic considerations, and since the two transistors require the same body voltage, it is possible to mix the two families and have both NMOS and PMOS in the same well. As a result, two structures are possible that are shown in Fig. 2.3 (c): RW NMOS and FW PMOS in a P-well or RW PMOS and FW NMOS in an N-well. Both implementations are valid and yield a small circuit area. Even though in our design a conventional well separation is used for measurement purposes.

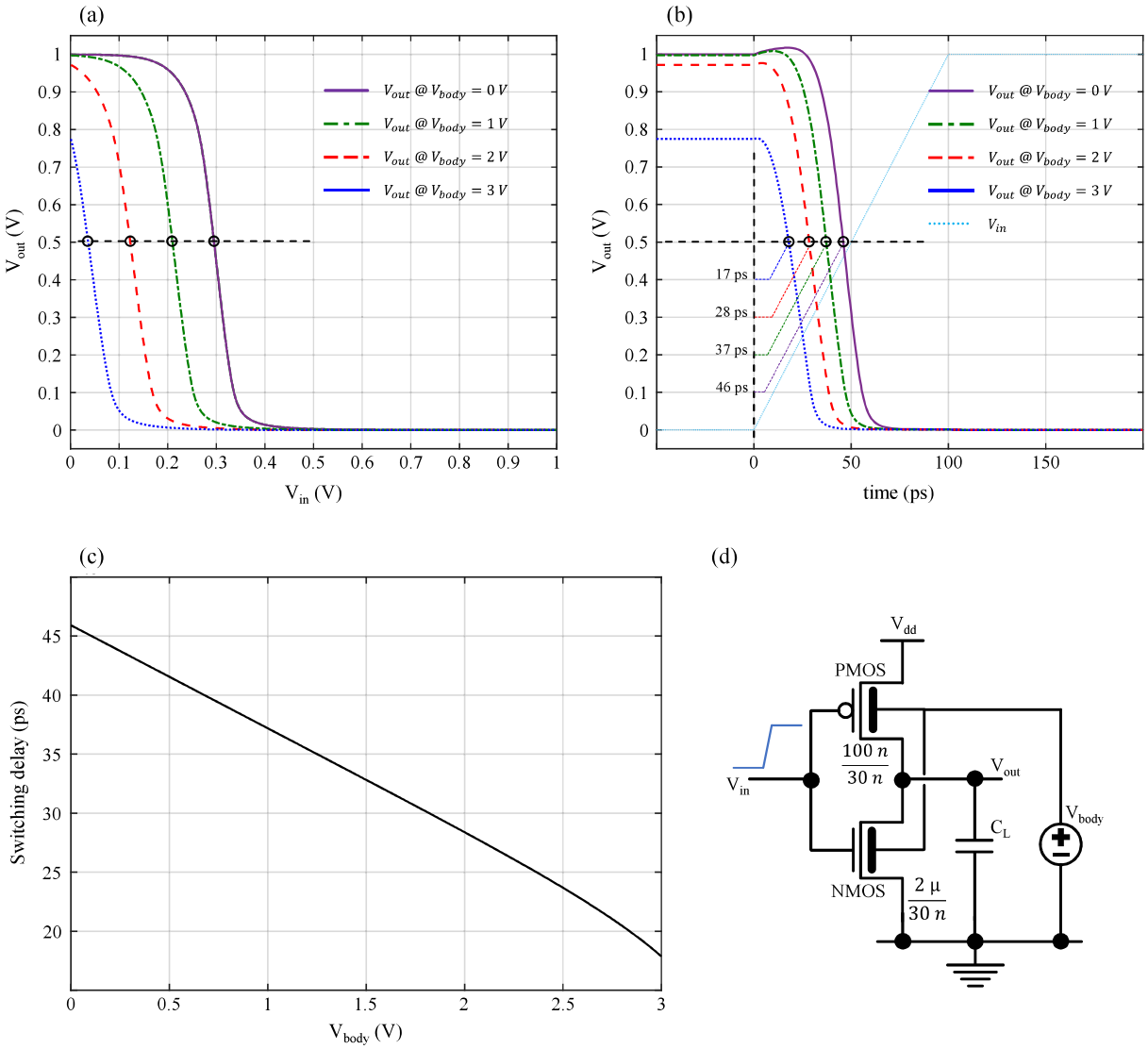


Fig. 2.4 Reducing switching delay of an inverter through body biasing of its transistors. (a) VTC of the inverter at different body voltages. (b) Transient response of the inverter for a 100 ps input ramp at different body voltages. (c) Switching delay of the inverter versus body voltage of NMOS and PMOS. (d) schematic of the tested inverter.

Fig. 2.4 shows the simulation setup and results for body biasing of an inverter. By applying the same voltage to NMOS and PMOS and sweep it from 0 V to 3 V, the switching threshold of the inverter is reduced by about 240 mV, as can be seen in Fig. 2.4 (a). In this simulation the optimum sizing which has been obtained in the previous section is used. Applying a 100 ps ramp to the inverter (driving the same C_L as in the sizing simulations), while sweeping the body voltage is presented in Fig. 2.4 (b). Despite the sizing method, here at each sweep step the switching delay is reduced. The full range of variation for switching delay versus body voltage is presented in

Fig. 2.4 (c). There is an almost linear relationship between the switching delay and body voltage. This is due to the linear relationship between the body voltage and the threshold voltage in the FD-SOI CMOS technology. A maximum 66 % (~ 30 ps) switching delay reduction is obtained by 3 V increase in the body voltage.

One can notice that in Fig. 2.4 (a) and (b), for a body voltage of 3 V, the maximum output is less than 800 mV. This is because at this body voltage the threshold voltage of PMOS is so high that it cannot be fully conducting. This is not a problem in the avalanche detection and in the quenching phase, though, for the SPAD reset phase, the impact of a weak 1 signal should be noticed.

The two methods, sizing, and body biasing, together reduce the switching delay of about 80 % (~ 60 ps). Therefore, combining these two methods can result in an avalanche detection in the range of a few tens of picoseconds. Based on these observations, we propose to use the body biased inverter at its optimum sizing, for avalanche detection and active quenching. In the following section, an ultra-fast AQAR circuit, featuring the body biased inverter is presented.

2.3. Body Biased Inverter Based AQAR Circuit

Fig. 2.5 presents the schematic of the proposed Body Biased Inverter-Based (BBI) AQAR circuit. The body biased inverter is composed of N_1 and P_1, which bodies are connected to a(n) (external) voltage source. Through this voltage source, the sensitivity and delay time of the AQAR circuit can be tuned. The body biased inverter drives a wide PMOS M_{AQ}, which is the AQ transistor. M_{AQ} is directly connected to the SPAD anode. The detection inverter also drives a second standard inverter (Inv) that generates a digital signal (PixOut) to flag the avalanche detection. M_{PQ} is the active implementation of the passive quenching transistor, and its value can be tuned externally through its gate (PQ). There are also two ultra-wide switches, Quenching switch P_2 and Reset switch N_2., that enable switching between AQ mode and PQ mode, or between AR and PR. These two switches are only there for measurement purposes and can be excluded from the final pixel implementation. The AR transistor M_{AR} is driven by a delay line which controls the hold off time and reset time of the SPAD.

Fig. 2.6 shows the schematic of the delay line. It is composed of three inverters, two delay elements, one D flip flop, and one NAND gate. The PixOut is delayed by the delay element for a

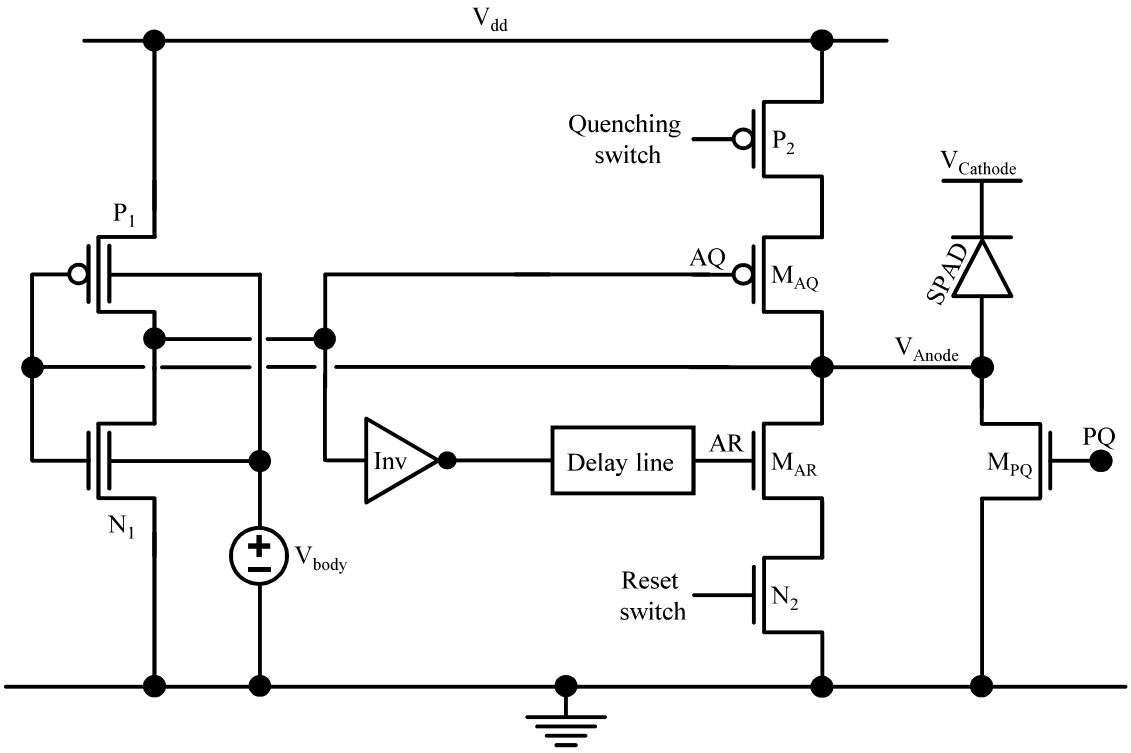


Fig. 2.5 Schematic of the proposed Body Biased Inverter-based Active Quenching Active Reset (BBI AQAR) circuit.

certain hold of time, the output is conditioned by an inverter to drive the clock of the D flip flop and consequently the AR signal is applied to M_{AR} that starts the reset phase. The NAND gate and External reset signal are there to ensure that in any unexpected case the SPAD can be reset externally, thus in the normal functioning of the circuit they can be excluded from the analysis. The delay element is a current starved inverter which applies the delay only on the falling edge of its output. The output is followed by a capacitor C_{delay}. The input current is controlled by an external trimmer to tune the desired delay. There is an internal feedback loop that reset the D flip flop through a delay element and determine the reset pulse width.

The operation chronogram of the circuit is shown in Fig. 2.7. At the photon arrival, the avalanche process starts. The generated avalanche current discharges the junction capacitance of the SPAD and thus raises the anode voltage (t_0). Quickly after that, this rise in anode voltage is detected by the body biased inverter. Meanwhile, the AQ signal falls and thus the AQ transistor M_{AQ} starts to conduct and generate a high and fast rising current to contribute to the quenching process. After the quenching, the anode voltage stays high till the end of the determined hold off time (t_1). At this moment, the AR signal rise and thus M_{AR} start the reset phase by charging the

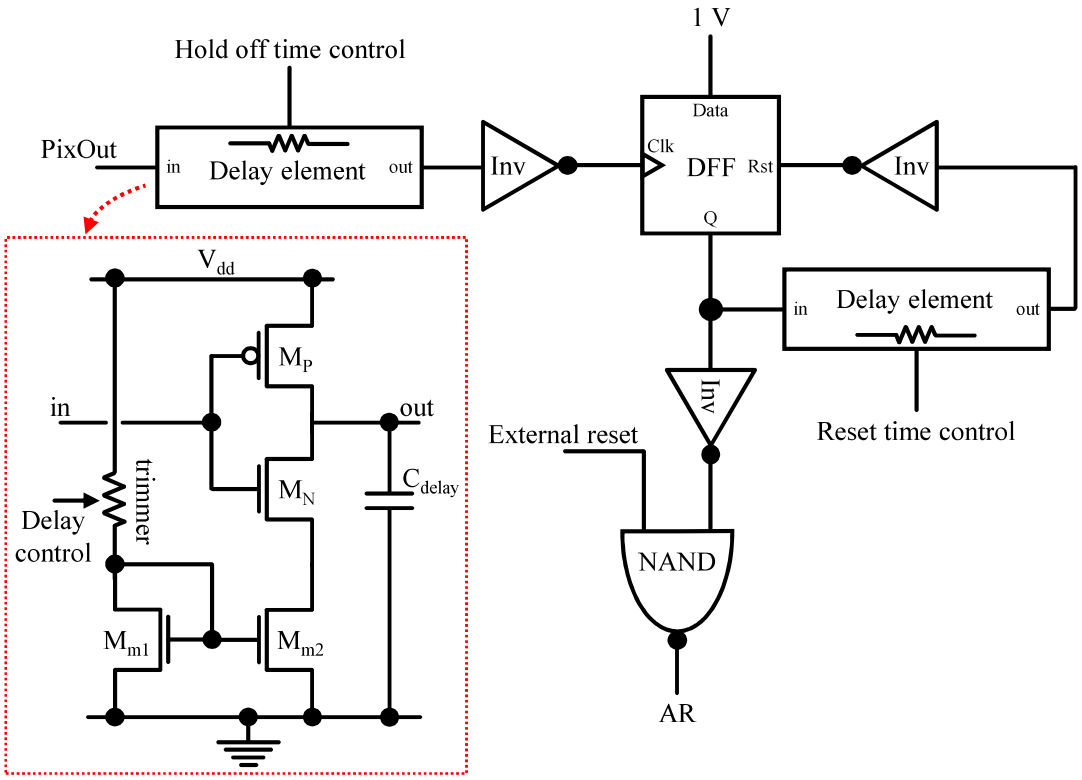


Fig. 2.6 Schematic of the delay line of the BBI AQAR circuit. Inset: schematic of the delay element in the delay line.

junction capacitance of the SPAD. Therefore, the anode voltage falls. The AR signal remains high until the determined reset time (t_2). The AR signal falls to zero to turn off M_{AR}. From now on the pixel is ready for a new photon and avalanche detection.

It should be noted that to have a fast AQ phase, the AQ transistor M_{AQ} should be sized properly. At the beginning of the AQ phase M_{AQ} acts as a current source which charges the SPAD junction capacitance. At the end of this phase M_{AQ} is almost a resistor. In both cases to charge the SPAD capacitance faster, it needs to increase the width of the PMOS. However, increasing the size of this transistor can add parasitic capacitances to the SPAD junction capacitor which slows down the quenching phase. Therefore, there is an optimum size for M_{AQ}.

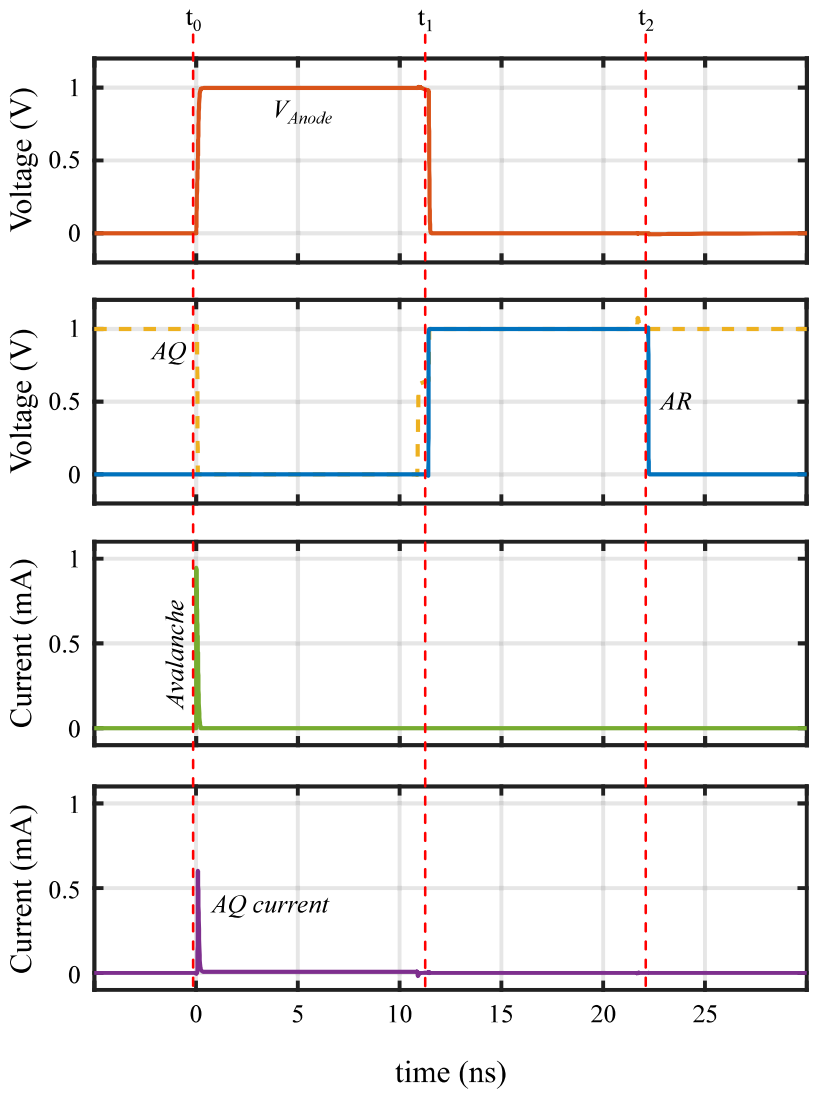


Fig. 2.7 Timing diagram of the BBI AQAR circuit operation. From top to bottom: anode voltage, active quenching and active reset signals, avalanche current, and active quenching current. At t_0 the avalanche starts. At t_1 hold off time is finished and reset phase is started. At t_2, the reset phase is finished, and the pixel is ready for a new event.

2.4. Results and Discussions

In this section, the measurement process, test bench, and the chip, which is exclusively designed to characterize the BBI AQAR pixel, are discussed. The post layout simulation results are presented to justify the functionality of the circuit and are verified by the experimental results.

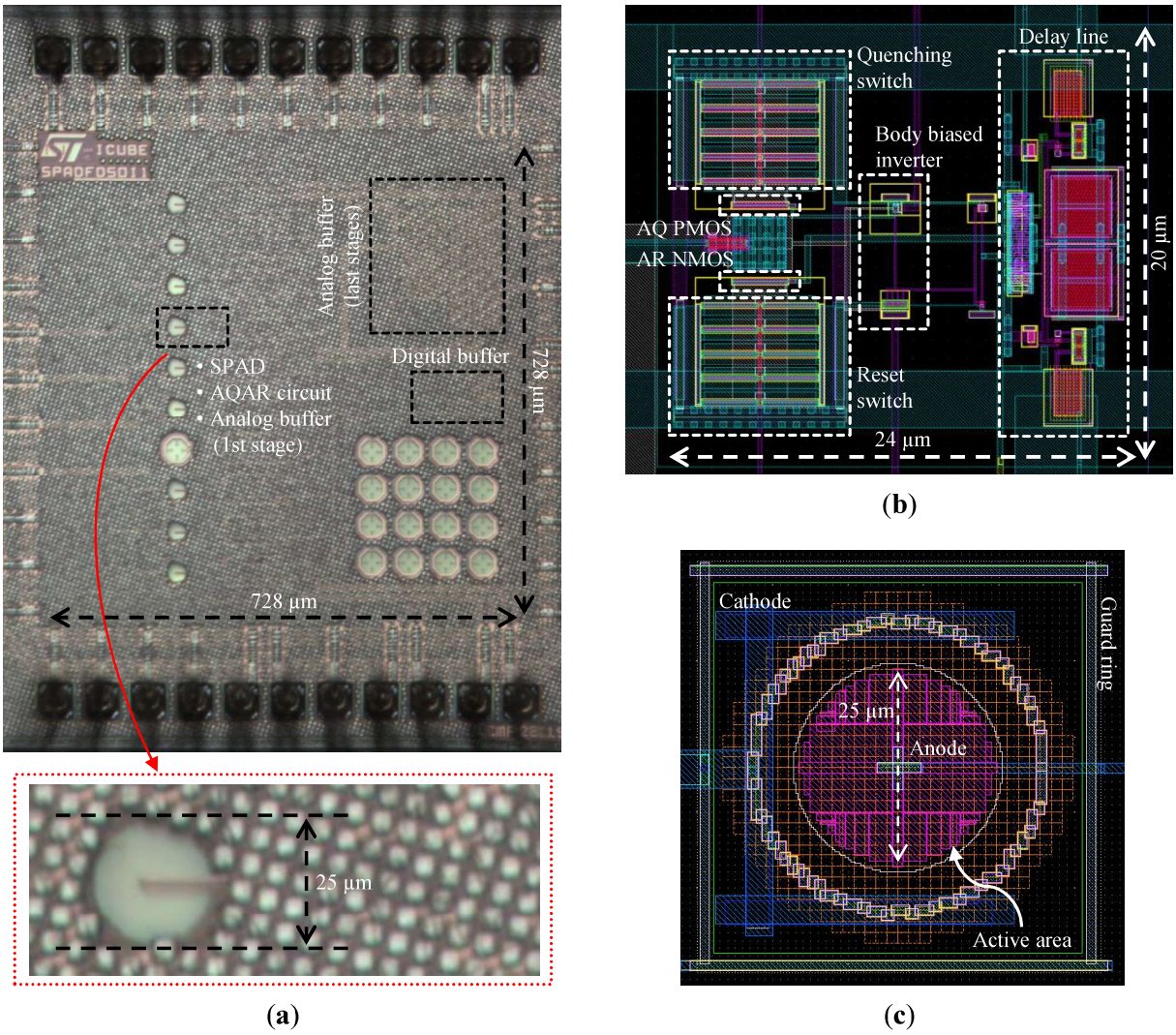


Fig. 2.8 (a) Micrograph of the designed chip. Zoomed area of the device under the test. (b) Layout of the BBI AQAR circuit. (c) Layout of the quasi-circular SPAD.

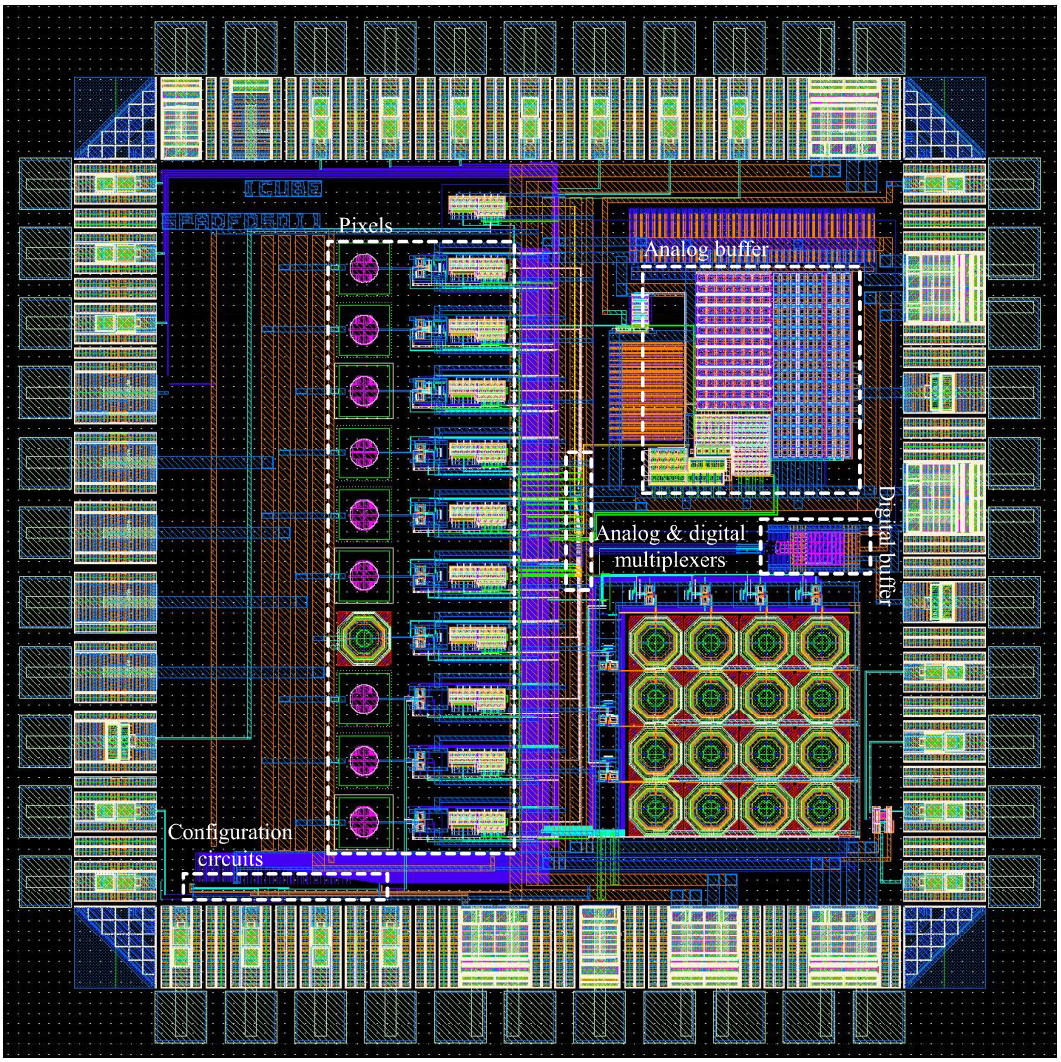
2.4.1. *Chip and Test Bench*

Fig. 2.8 (a) shows a 1×1 mm^2 (728×728 µm^2 internal area) chip, designed and fabricated in a 28 nm FD-SOI CMOS technology. The chip is composed of different SPADs and their peripheric circuitry. The SPAD under test in this chapter is presented in Fig. 2.8 (b). It is a quasi-circular SPAD with a 25 µm diameter, presented in [154]. The SPAD is designed in the standard process of the used 28 nm FD-SOI technology, without any process modifications. The layout of the SPAD has no Design Rule Check (DRC) violation according to the DRC of the used technology. This is the reason it is not perfectly circular. The total area of the SPAD, including the guard rings, is 53× 53 µm^2, while the active area is about 490 µm^2. The BBI AQAR circuit is placed beside the SPAD,

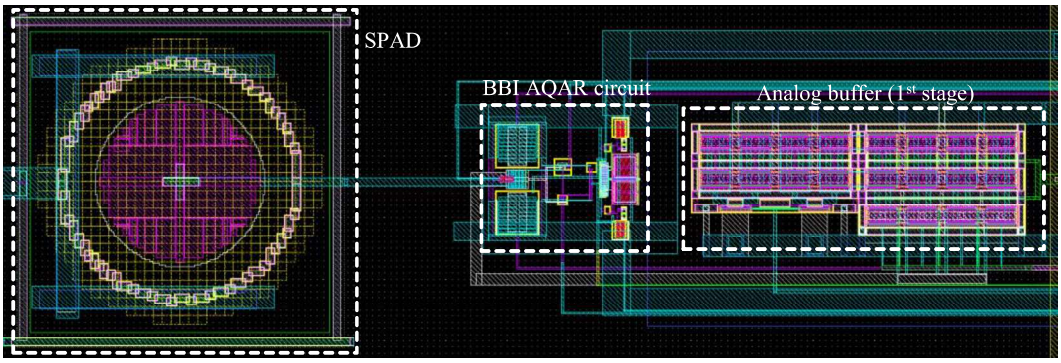
see Fig. 2.8 (c). The two wide switches are at the left side of the layout view and since they are there only for measurement purposes, can be excluded from the circuit area. The delay line is on the right and the detection circuit is just the two inverters at the middle of the layout. It is mentioned that the well separation is for the sake of measurements, otherwise it is possible to implement the detection inverter in one single well. Nevertheless, the total area of the BBI AQAR circuit (switches included) is less than 24×20 μm^2.

Except the SPADs, all the chip surface is covered by metal filling, making it impossible to see the electronics. Thus, the layout view of the full chip is presented in Fig. 2.9. In the vertical array of SPADs, each SPAD is connected to a quenching and detection circuit. There is a wide band analog voltage follower that is connected to the anode to observe its voltage variations during an avalanche. The first stage of the buffer is repeated for each SPAD in the array and to save the silicon area and reduce the power consumption, the outputs of these first stages are connected to the last three stages of the buffer through an analog multiplexer. The avalanche detection signals are also fed to a digital buffer by a digital multiplexer. Thanks to the addressing and configuration circuits, we can observe the digital and analog output of each pixel. A simplified symbolic schematic view of the chip is depicted in Fig. 2.10. The pixel under test, composed of the SPAD, BBI AQAR circuit, and first stage of the analog voltage buffer is shown in Fig. 2.9 (b).

In order to characterize the chip, a testing platform based on a Field-Programmable Gate Array (FPGA) board is developed, as shown in Fig. 2.11. The test chip was mounted on a daughter board which is plugged into a motherboard. The daughter board has an SMA connector that is connected to the output of the analog buffer to observe the anode voltage transient behavior. The motherboard integrates the power supplies of the chip, circuits for the configuration and addressing inside the chip (i.e., a bitstream sent by the FPGA), Digital to Analog Converters (DACs) to generate the body biasing voltages, digital trimmers to control the hold off and reset time of the quenching circuit, and a USB2 controller for communication with a LabVIEW interface. The LabVIEW interface controls the system, reads out the data from the FPGA, and allows saving them in a text file. The FPGA is configured from the LabVIEW interface to send configuration bits to the chip, to set the DAC and the digital trimmer on the motherboard. The bitstream allows switching between PQ and AQ through the switches in the circuit. The FPGA implantation also includes the afterpulsing measurement unit, which consists of a 24 bits gray counter (operating at 200 MHz),



(**a**)



(**b**)

Fig. 2.9 (a) Layout view of the full chip, including pixels, analog and digital buffers and multiplexers, and configuration circuit. (b) Layout view of the pixel under the test, including the SPAD, BBI AQAR circuit and the first stage of the analog buffer.

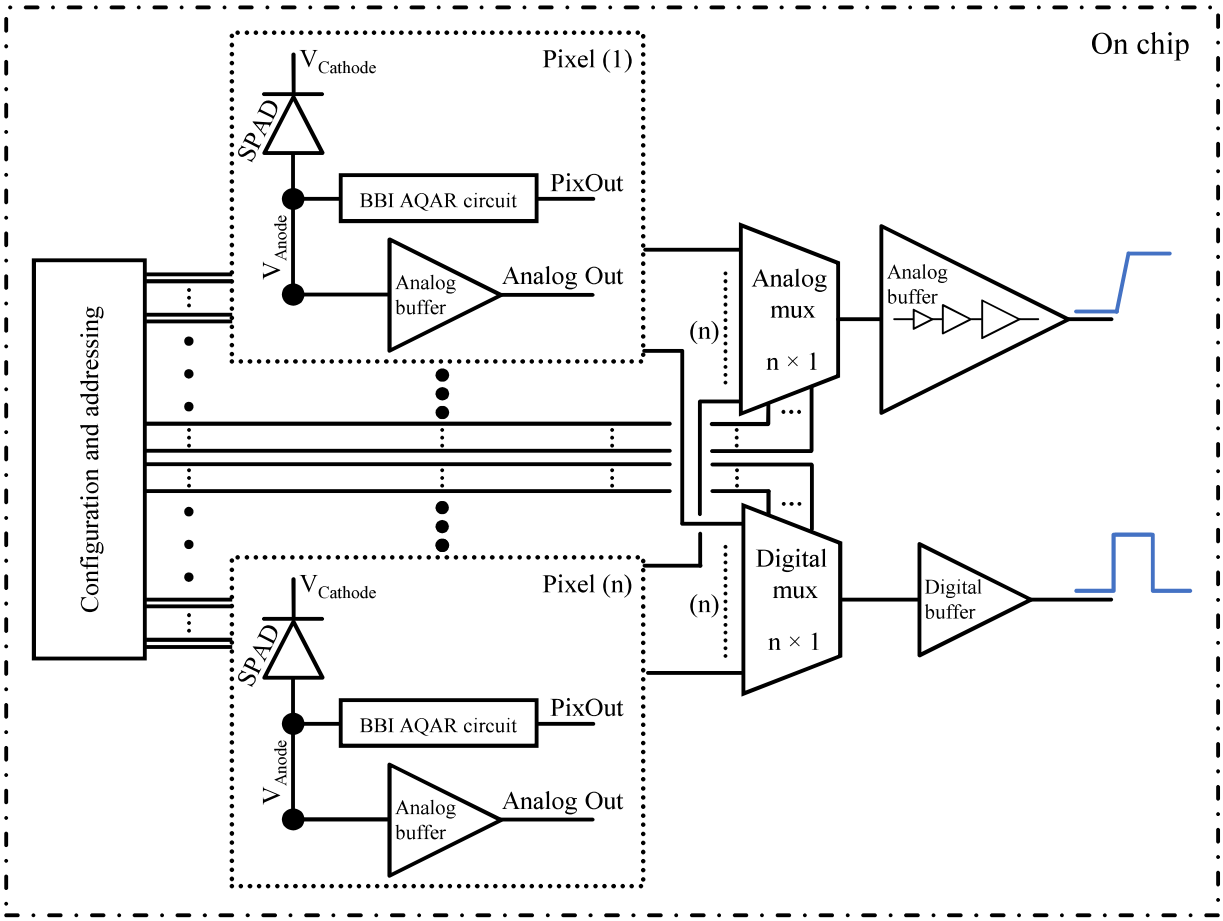


Fig. 2.10 A simplified symbolic view of the chip: different blocks and their interconnections.

written in a FIFO when a rising event occurs (a detected avalanche) on the digital output of the AQAR circuit. This allows measuring the inter arrival time of the SPAD detected events with 5 ns resolution. All the test bench boards are powered by a laboratory power supply. The analog output is measured by a 13 GHz wide-band oscilloscope. And, finally, the testing platform (FPGA board, daughter board, and motherboard) is placed in a thermal chamber for the constant temperature operating condition. All the results are based on the DCR measurements at 40 °C, while the SPAD is in the free-running mode under the dark condition.

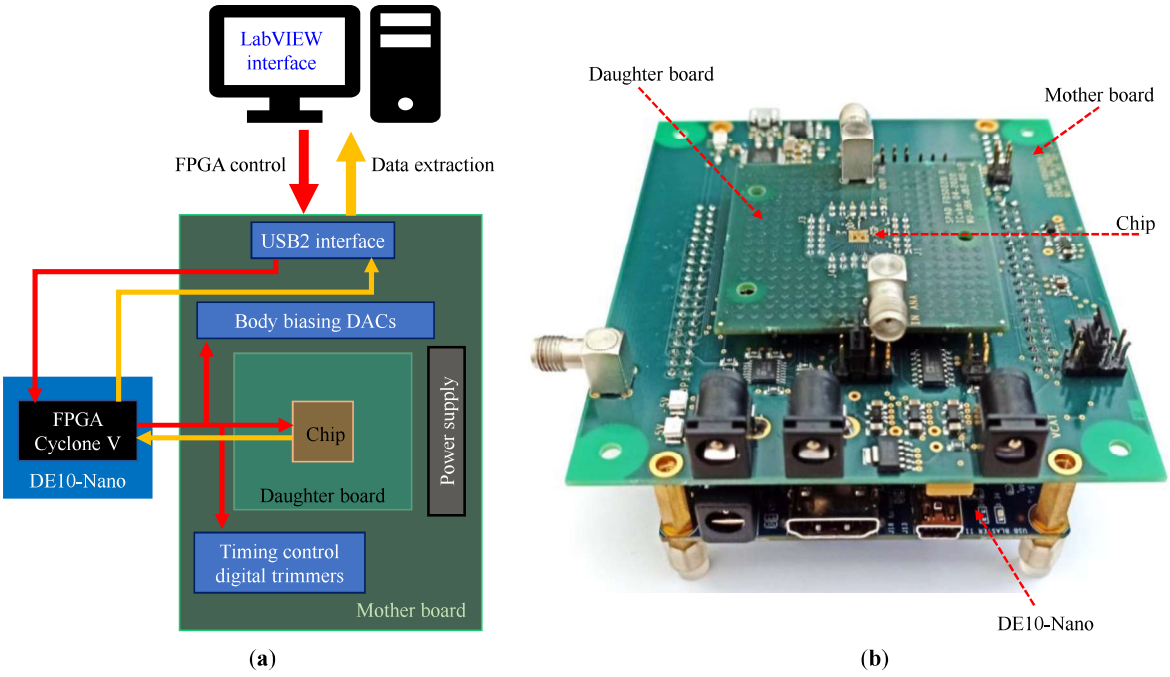


Fig. 2.11 (a) Block diagram of the experimental setup. (b) photo of the experimental test bench.

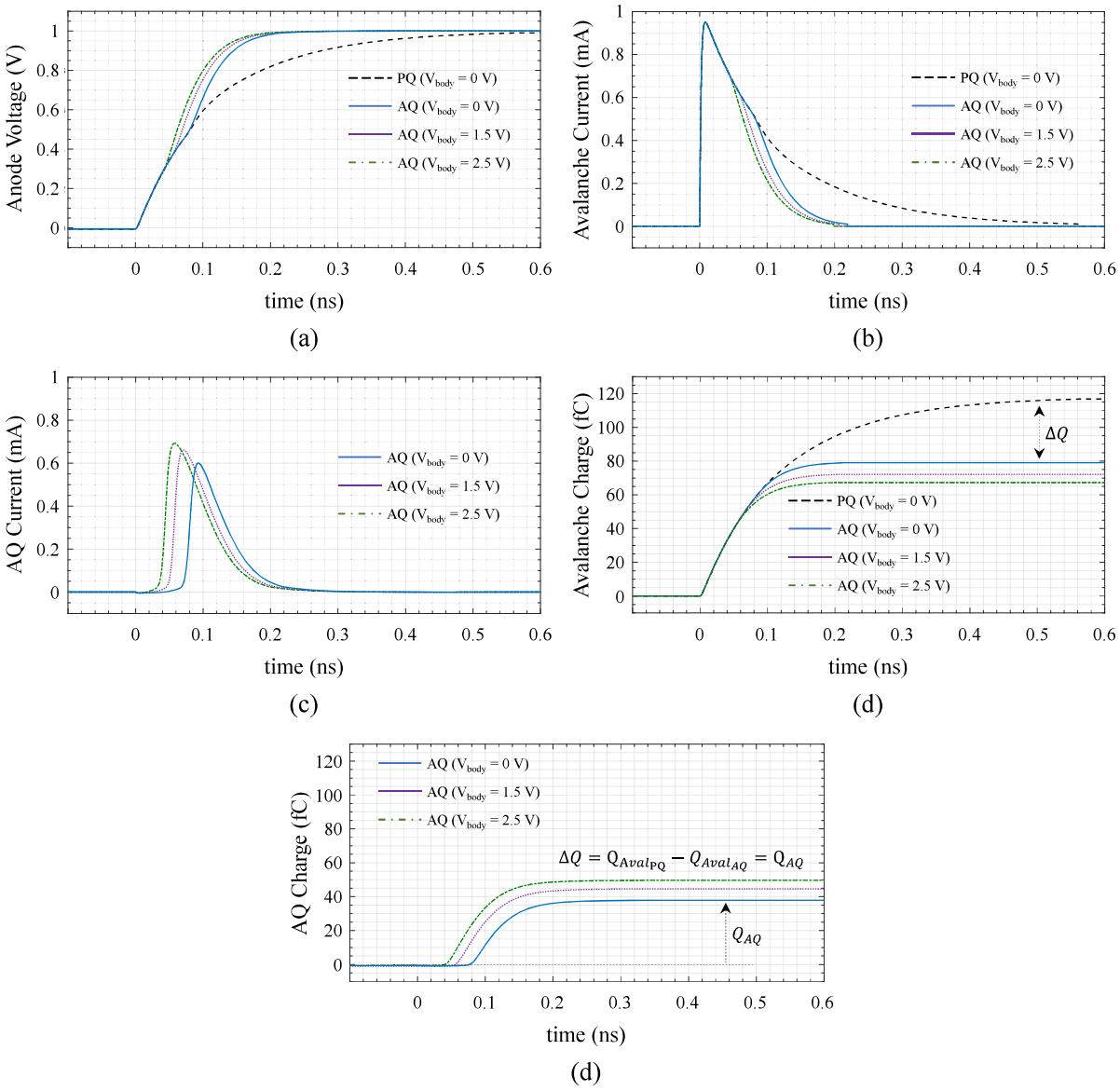


Fig. 2.12 Simulation results during an avalanche in PQ and AQ modes for different body voltages. (a) anode voltage. (b) avalanche current. (c) active quenching current. (d) avalanche charge. (d) active quenching contributed charge.

2.4.2. *Simulation Results*

To simulate the SPAD behavior, a SPICE macro model is used which is reported in [167]. The junction capacitance of the SPAD is set to 100 fF and 1 V of excess bias voltage is applied to the SPAD. To investigate the effectiveness of the proposed BBI AQAR circuit in afterpulsing reduction, the simulations are performed in both AQ mode and PQ mode thanks to Quenching switch. Fig. 2.12, shows the transient simulation results during the occurrence of an avalanche. In

the anode voltage transitions (Fig. 2.12 (a)), the anode voltage curves deviate from the PQ curve and their slopes are increased, and finally they reach the excess bias voltage sooner. This means, at the deviation point, the quenching transistor M_{AQ} is injecting current to anode. That demands the avalanche to be detected sooner than the deviation starts. Therefore, one can see, in the slowest case, an avalanche is detected in less than 80 ps. By increasing the body voltage (V_{body}), the sensitivity level of the detection inverter is increased. As can be seen, the AQ curves deviation starts sooner by increasing V_{body} which means the avalanche is detected sooner. The fastest detection time is when V_{body} = 2.5 V, which is less than 40 ps.

The same behavior is observable in the avalanche current (Fig. 2.12 (b)); the difference between the PQ mode avalanche current and those of the AQ mode. The quenching time in the AQ mode is distinctively shorter than the PQ mode. It is evident in the figure that the sooner the avalanche is detected, the shorter the quenching time. Taking 50 µA as the current at which the avalanche is considered as quenched, the total quenching time in the AQ mode can be as short as 150 ps, 2.5 times faster than in the PQ mode.

These differences between the PQ mode and AQ mode in the avalanche current and anode voltage curves are due to the contribution of the AQ transistor M_{AQ} in the quenching process. These differences are proportional to the amount of injected current by M_{AQ}, and also the time that this current injection starts, which is mainly determined by the sensitivity of the detection circuit. Fig. 2.12 (c) shows the current of M_{AQ} (I_{AQ}). Faster detection (a higher sensitivity) causes M_{QA} to contribute earlier and more. The sooner the M_{AQ} conducts, it experiences a higher drain-source voltage (V_{dd} - V_{Anode}) for a longer time, thus its current can be higher.

Fig. 2.12 (d) presents the avalanche charge in the SCR. As was expected faster quenching reduces the avalanche charge. Thanks to the proposed BBI AQAR circuit, the avalanche charge is reduced up to 42 % (50 fC) in comparison with the PQ mode. This charge difference (ΔQ) is equal to the injected charge by the AQ transistor M_{AQ}, see Fig. 2.12 (e). This avalanche charge reduction should decrease the afterpulsing effect. This claim is justified in the next section. Table 2.1 presents all the simulation results of comparison between PQ and AQ modes at different sensitivity levels.

It should be mentioned that we could apply 3 V of body voltage, though, the maximum is kept to 2.5 V to avoid excessive subthreshold leakage current. The static power consumption of the circuit, including the delay line, is around 6.7 µW at 2.5 V of the body voltage.

TABLE 2.1
SIMULATION RESULTS FOR BBI AQAR CIRCUIT IN PQ AND AQ MODES AT DIFFERENT BODY VOLTAGES.

	V_{body} = 0 V	V_{body} = 1.5 V	V_{body} = 2.5 V
Avalanche detection time (ps)	< 80	< 60	< 40
Quenching time in PQ (ps)	380	380	380
Quenching time in AQ (ps)	170	160	150
Avalanche charge in PQ (fC)	117	117	117
Avalanche charge in AQ (fC)	79	72	67
AQ charge contribution (fC)	38	45	50
Avalanche charge reduction (%)	32	38	42

2.4.3. *Experimental Results*

Thanks to the integrated wideband analog buffer, the experimental replica of simulation result in Fig. 2.12 (a), is presented in Fig. 2.13: anode voltage variations during the avalanche occurrence. By changing the body voltage through the associated DACs, different anode voltage curves are extracted for V_{body} = 0 V, 1 V, and 2 V. The applied excess bias voltage is about 500 mV, as it can be seen in the PQ mode anode voltage curve. At more than 500 mV of excess bias voltage, the used SPAD is saturated by noise. Also, for a higher body voltage, some spurious avalanches are detected by the AQ circuit, indicating that the threshold is too low, and the circuit can be triggered by the noise. Therefore, to have fairer and more realistic results we keep the maximum body voltage to 2 V in our measurements to be sure that these results are applicable for an array of pixels. It is not easy to find the avalanche detection time from Fig. 2.13, however, the quenching time is observable. The quenching time in the PQ mode is about 1 ns, while the proposed BBI AQAR circuit reduces this time by 500 ps (total quenching time of about 480 ps). It means the detection time is much less than 480 ps. One can also see the sweep of the quenching time through changing the body voltage V_{body}. These experimental results prove the efficiency of the proposed circuit in the quenching time reduction and are in accordance with the reported simulation results. The following measurements demonstrate how this quenching time reduction reduces the avalanche charge and therefore the afterpulsing effect.

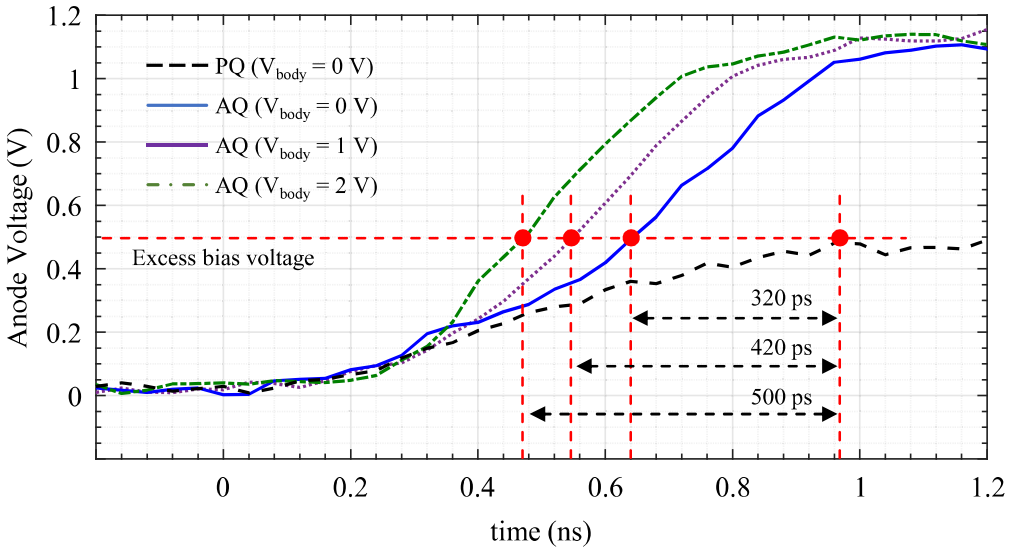


Fig. 2.13 Measured anode voltage variations during an avalanche in PQ and AQ modes at 40 °C with 500 mV of excess bias voltage at different body voltages.

Fig. 2.14 shows the avalanche interarrival time histogram in PQ mode and AQ mode, with different body voltages at V_{ex} = 400 mV. It is worth mentioning that, in order to have a fair comparison between the PQ and the AQ circuits, the reset is always set to the active mode after a hold off time set to only 5 ns. Thanks to the gray counter, the intervals between two consecutive avalanches are calculated to obtain the histogram of the interarrival time. Since the probability of the afterpulsing is very low after a few microseconds, this part of the histogram (in our case, after 6 µs) is fitted by an exponential function to obtain the afterpulsing-free DCR curve, modeled as a Poisson process. Then, for an interarrival time of less than 10 µs (to guarantee that no afterpulse is missed), the afterpulsing effect is extracted as the area between the measured interarrival time histogram and the fitted afterpulsing-free DCR curve. By subtracting the fitted DCR curve from the interarrival time histogram, the number of afterpulses during the measurement is extracted. It is clearly observable in Fig. 2.14 that the higher the body voltage, the lower the detection threshold, and so the lower the afterpulsing effect.

The same measurement procedure is performed for different dead times and body voltages at V_{ex} = 400 mV. The results are announced in Afterpulsing Probability Percentage (APP): the ratio between the counted number of afterpulses and the total counts of the histogram. Fig. 2.15 shows

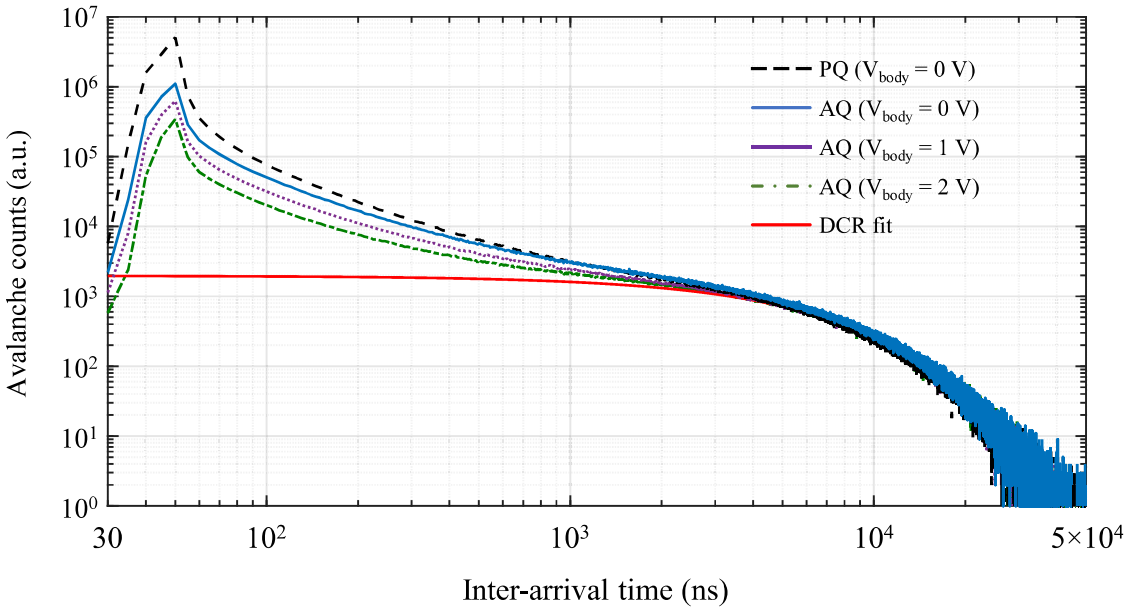


Fig. 2.14 Avalanche interarrival time histogram in PQ and AQ modes at different body voltages for a hold off time of 5 ns. The excess bias voltage is 400 mV, and the temperature is 40 °C.

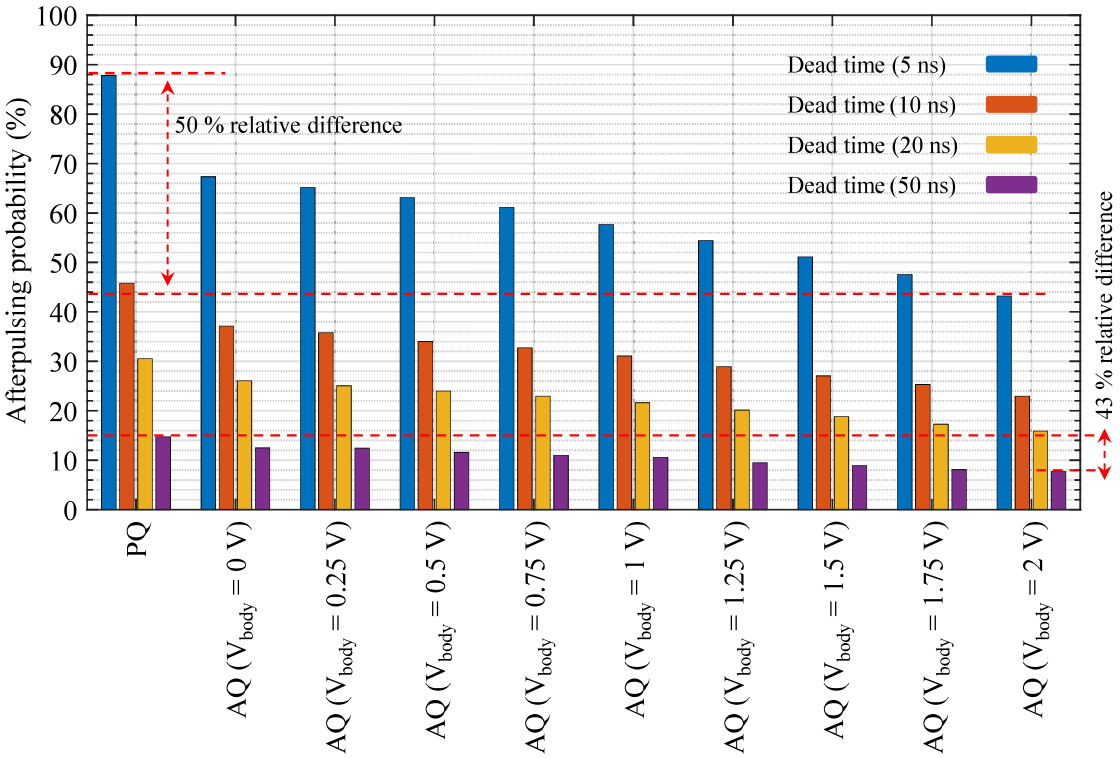


Fig. 2.15 Afterpulsing probability in PQ and AQ modes at different body voltages for hold-off times of 5, 10, 20, and 50 ns. The excess bias voltage is 400 mV, and the temperature is 40°C.

the afterpulsing probability for hold off times of 5 ns, 10 ns, 20 ns, and 50 ns, while the body voltage is swept from 0 V to 2 V, to investigate the effectiveness of the proposed circuit in the afterpulsing reduction in comparison with the PQ. Table 2.2 presents the APP values for both PQ and AQ modes at different body voltages and different dead times, with 400 mV of excess bias voltage. The APP relative reduction is the ratio between the difference of APP value in PQ and AQ modes at a certain body voltage, and the APP value in PQ mode:

$$APP\ relative\ reduction_x = \frac{APP_{PQ} - APP_{AQ(V_{body\,x})}}{APP_{PQ}} \tag{2.3}$$

Even though the absolute value of APP is high (88 % at 5 ns of dead time) for the used SPAD, the proposed BBI AQAR circuit is able to reduce this value from 24 % up to 51 %. It is expected that this APP reduction ability remains the same with a less noisy SPAD. These results justify that the proposed AQAR circuit alleviate the common tradeoff between the afterpulsing and the dynamic range of the SPAD by enabling the increase of the maximum photon count rate, thanks to the afterpulsing reduction by a factor of two at a very short hold off time.

TABLE 2.2
MEASURED APP AND APP REDUCTION OF BBI AQAR CIRCUIT AT DIFFERENT BODY VOLTAGES AND HOLD OFF TIMES. THE EXCESS BIAS VOLTAGE IS 400 MV, AND THE TEMPERATURE IS 40°C.

	APP (%) (5 ns, 10 ns, 20 ns, 50 ns)	APP relative reduction (%) (5 ns, 10 ns, 20 ns, 50 ns)
PQ	88, 46, 30, 15	_
AQ (V_{body} = 0 V)	67, 37, 26, 12	23.8, 19.5, 13.3, 20
AQ (V_{body} = 0.25 V)	65, 36, 25, 12	26, 21.7, 16.7, 20
AQ (V_{body} = 0.5 V)	63, 34, 24, 11.5	28.4, 26, 20, 23.3
AQ (V_{body} = 0.75 V)	61, 33, 23, 10.9	30.6, 28.2, 23.3, 27,3
AQ (V_{body} = 1 V)	57, 31, 21, 10.5	35.2, 32.6, 30, 30
AQ (V_{body} = 1.25 V)	54, 29, 20, 9.4	38.6, 37, 33.3, 37.3
AQ (V_{body} = 1.5 V)	51, 27, 18.7, 8.8	42, 41.3, 37.6, 41.3
AQ (V_{body} = 1.75 V)	47, 25, 17, 8	46.5, 45.6, 43.3, 46.6
AQ (V_{body} = 2 V)	43, 23, 16, 7.7	51.1, 50, 46.6, 48.6

TABLE 2.3
MEASURED APP AND APP REDUCTION OF BBI AQAR CIRCUIT IN PQ MODE AT 0 V OF BODY VOLTAGE AND DIFFERENT HOLD OFF TIMES. THE EXCESS BIAS VOLTAGE IS 400 MV, AND THE TEMPERATURE IS 40°C.

	APP (%)	APP relative reduction (%)
PQ (5 ns)	88	_
PQ (10 ns)	46	47
PQ (20 ns)	30	66
PQ (50 ns)	15	83

Table 2.3 shows the APP reduction versus dead time. As it is expected by increasing the dead time, APP strongly reduces: 83 % relative reduction with 50 ns dead time in regard with the APP value in 5 ns. However, the maximum count rate is also severely reduced by 10 times.

A comparison between APP in PQ and AQ modes is made, while the excess bias voltage (cathode voltage) is increasing. The body voltage is 1.5 V, and the hold off time is set to 20 ns; see Fig. 2.16. The true breakdown voltage of the used SPAD is around 9.6 V, in which we start to observe some very weak avalanches (low amplitude anode voltage signal), thanks to the analog buffer. In fact, the amplitude of the avalanches is lower than the threshold of the readout electronics; thus, they cannot be detected. This is the reason in Fig. 2.16 the first value of APP is recorded in about 9.75 V. Another interesting point in this figure is that the APP of AQ mode is higher than the one of PQ mode until 9.84 V. Yet, the reason is that the anode voltage amplitude is not very high; therefore, some of the events are missed in the PQ mode. On the contrary, in the AQ mode, when these very weak avalanches are detected by the body biased inverter, the AQ PMOS M_{AQ} raises their amplitude, so they can be detected by the following digital gate. By increasing the excess bias voltage, the AQAR circuit has more room to contribute to the quenching process, so, as it is shown in Fig. 2.16, the afterpulsing reduction is increasing when the excess bias voltage is increased, except for a cathode voltage above 10.15 V. Indeed, as the SPAD in PQ mode is saturated for a cathode voltage above 10.15 V, where the afterpulsing effect is also saturating at 100%, the APP relative reduction appears to be lowered.

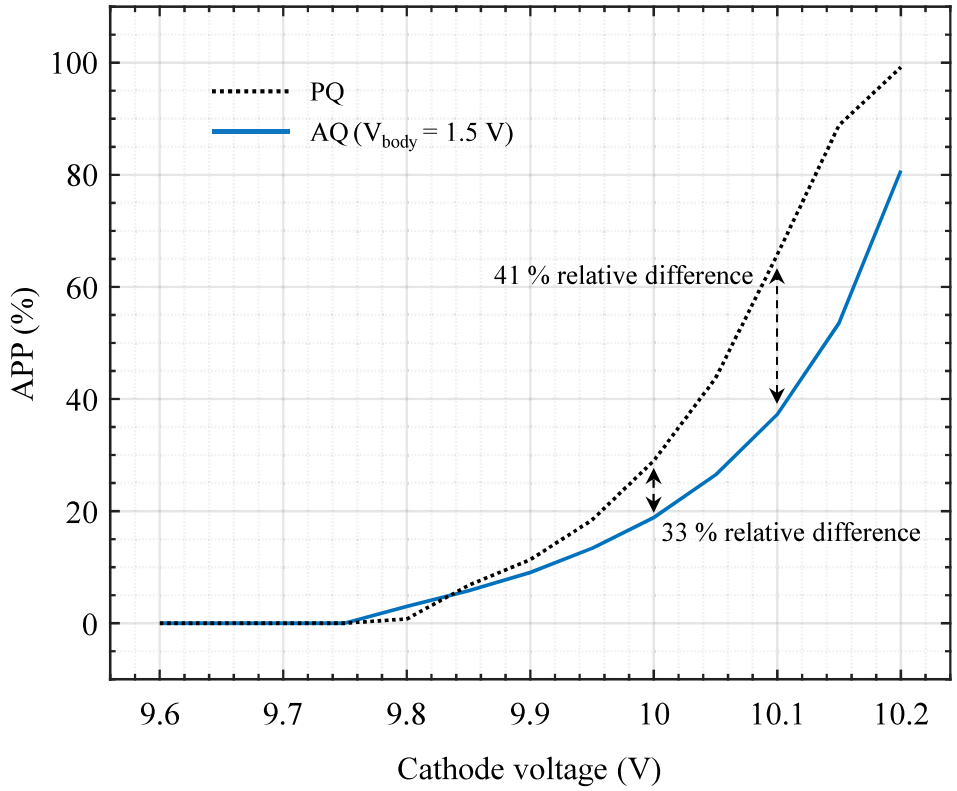


Fig. 2.16 Afterpulsing probability versus excess bias voltage for PQ and AQ modes at 1.5 V of body voltage and 20 ns of hold off time. The temperature is 40°C.

Chapter 3
Ticklish Inverter and Monolithic 3D Pixels

3.1. Introduction

It is shown that, properly sizing the transistors of an inverter and biasing their bodies can reduce the switching threshold of the inverter and make it a good choice for fast avalanche detection. Increasing W/L of the NMOS, while keeping the one of the PMOS constant (or decreasing it), pushes back Voltage Transfer Characteristic (VTC) of the inverter. Recalling the propagation delay equation (in our case high to low propagation delay t_{pHL}), one can see a reversely proportional relationship between the t_{pHL} and the W/L of NMOS, and a proportional relationship with the output capacitance (load + parasitic: C_L):

$$t_{pHL} \propto \frac{C_L \cdot V_{dd}}{(\frac{W}{L})_n \cdot (V_{dd} - V_{th_n})} \tag{3.1}$$

A wider NMOS means a higher parasitic capacitance at the output of the inverter. Therefore, there is an optimum W/L, in regard with the delay. It means, after a certain increase of the width, the switching delay is increased too. In the body biasing case, shortening the delay is achieved by reducing the threshold voltage of NMOS, V_{th_n}. Even though the design rules of the technology restrict the applicable body voltage. In the FD-SOI CMOS technology used in this thesis, maximum body voltage can be 3 V and the body factor is around 80 mV/V. Thus, a maximum threshold reduction of around 240 mV is achievable. In the previous chapter, the two aforementioned methods were combined to obtain a fast and low threshold inverter as an avalanche detection circuit. In this chapter, a new technique is presented, which makes the inverter an ultra-fast switching device without suffering from the drawbacks of the previous methods.

In the second part of this chapter, for the first time, the integration of an AQAR circuit in the native 3D structure is discussed. The challenges of placing the electronics over the SPAD are addressed. A new avalanche sensing method is introduced, which is not only robust against these

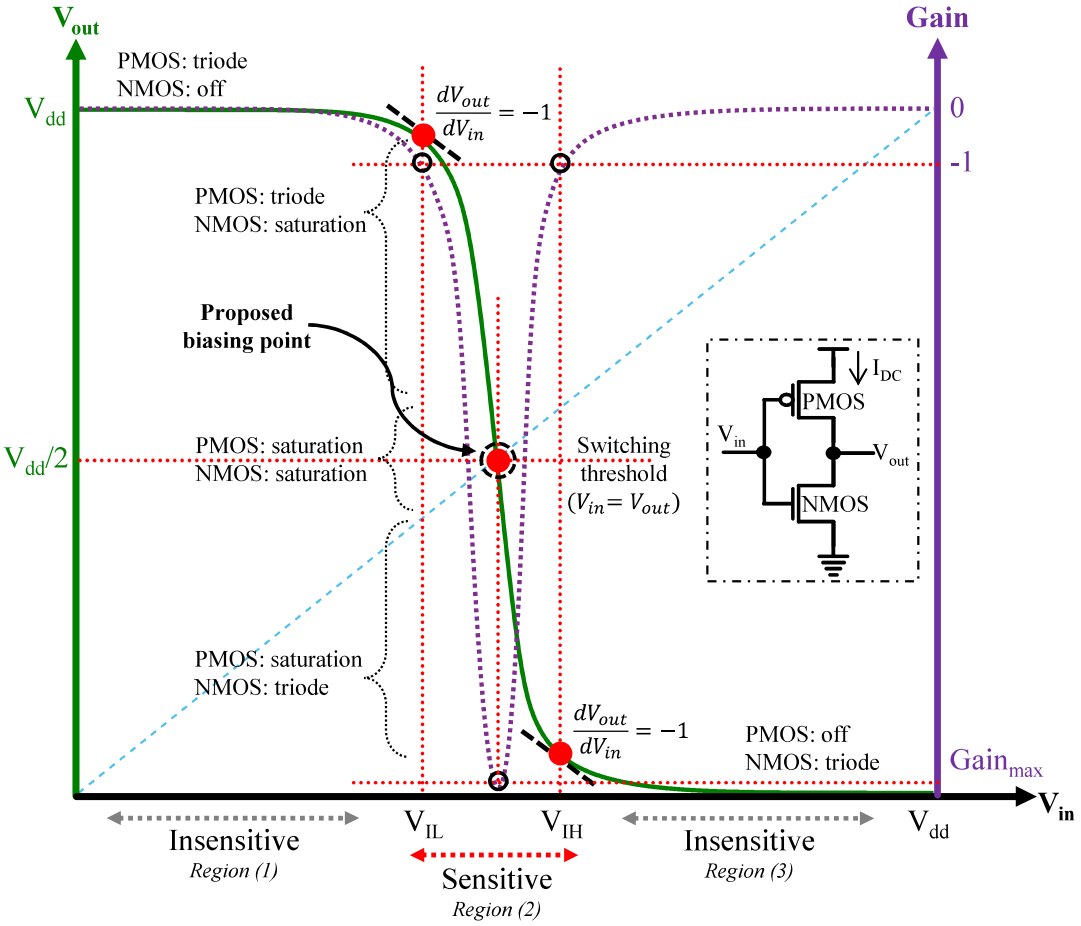


Fig. 3.1 Voltage Transfer Characteristic (VTC) – green curve - and large signal gain – purple dotted curve - of a typical inverter. The blue dashed line is V_{in} = V_{out}. Operating regions of the transistors at each part of the VTC are shown. Maximum gain is at the switching threshold. Inset: schematic of a CMOS inverter.

challenges, but also takes advantage of them. Based on that, the first AQAR circuit, compatible with the native 3D structure, is proposed.

3.2. Ticklish Inverter for Avalanche Detection

Fig. 3.1 shows the VTC of a typical inverter and the biasing condition of its NMOS and PMOS. It can be divided in two parts; One in which the output voltage of the inverter remains almost unchanged regardless of its input variations, region (1 and 3), and the second in which the output voltage is very sensitive to the input, region (2). In digital electronics, it is preferred to operate in region (1) and (3), where at each, one MOSFET is in the cut-off region while the other is in the triode. In consequence, a very low static power dissipation (almost zero) and highly distinctive “0” and “1” are achieved. However, as it was discussed, for an efficient avalanche detection, the

detector needs to be sensitive to the very small voltage variations. This is another explanation for the efforts to push back the VTC of the inverter in order to reduce the time that it operates in the region (1).

According to the large signal gain of the inverter $\frac{dV_{out}}{dV_{in}}$, as depicted in Fig. 3.1, region (2) is limited by V_{IL} (input low voltage) and V_{IH} (input high voltage), where the absolute value of the gain is 1. By moving from the region (1) and (3), where the gain is almost zero there, to the region (2), the gain is increasing until it reaches its maximum value (highest sensitivity) at the middle point, where $V_{in} = V_{out}$. Therefore, for efficient avalanche detection, we propose to bias the inverter in this sensitive region (2), particularly at the middle point, where both NMOS and PMOS are on and in saturation region.

3.2.1. *Ticklish Avalanche Detector*

To bias the inverter at its switching threshold, we propose to put a T-gate in parallel with the inverter. When the T-gate is closed, the input and the output voltages of the inverter are forcedly equal (middle point). Then when it turns open, the input node becomes floating while holding the previously set voltage. Now, any small variation at the input can be amplified by the maximum large signal gain of the inverter at its output. We called it a ticklish inverter. To adapt this biasing scheme with the avalanche sensing, the circuit in Fig. 3.2 (c) is proposed. In this circuit, T-gate feedback biases the detection inverter (inv1) at its middle point. Inv1 is followed by a second inverter (inv2) with a switching threshold slightly lower than the first. A coupling capacitor Cc transfers the anode voltage variations to the input of the detection inverter and decouples their DC levels.

The operation principles of the ticklish detection circuit are shown in Fig. 3.2 (a), (b), and (c). To start, the T-gate must be switched off after it biased the inverter at the middle point. This is generating a voltage across Cc (V_C), equal to the switching threshold of the detection inverter. When an avalanche begins, the anode voltage starts to rise from 0 V (Fig. 3.2 (c)). Only a few millivolts of variation are needed at V_{Anode} to change V_{out} from almost zero to almost V_{dd} (Fig. 3.2 (b)). These few millivolts are the difference between the switching threshold voltages of the detection inverter and the following inverter (inv2). A conventional inverter as a detection circuit needs a higher input swing to detect an event. This high swing is because the input voltage must reach the middle point of the inverter, whereas the ticklish inverter is already biased at the

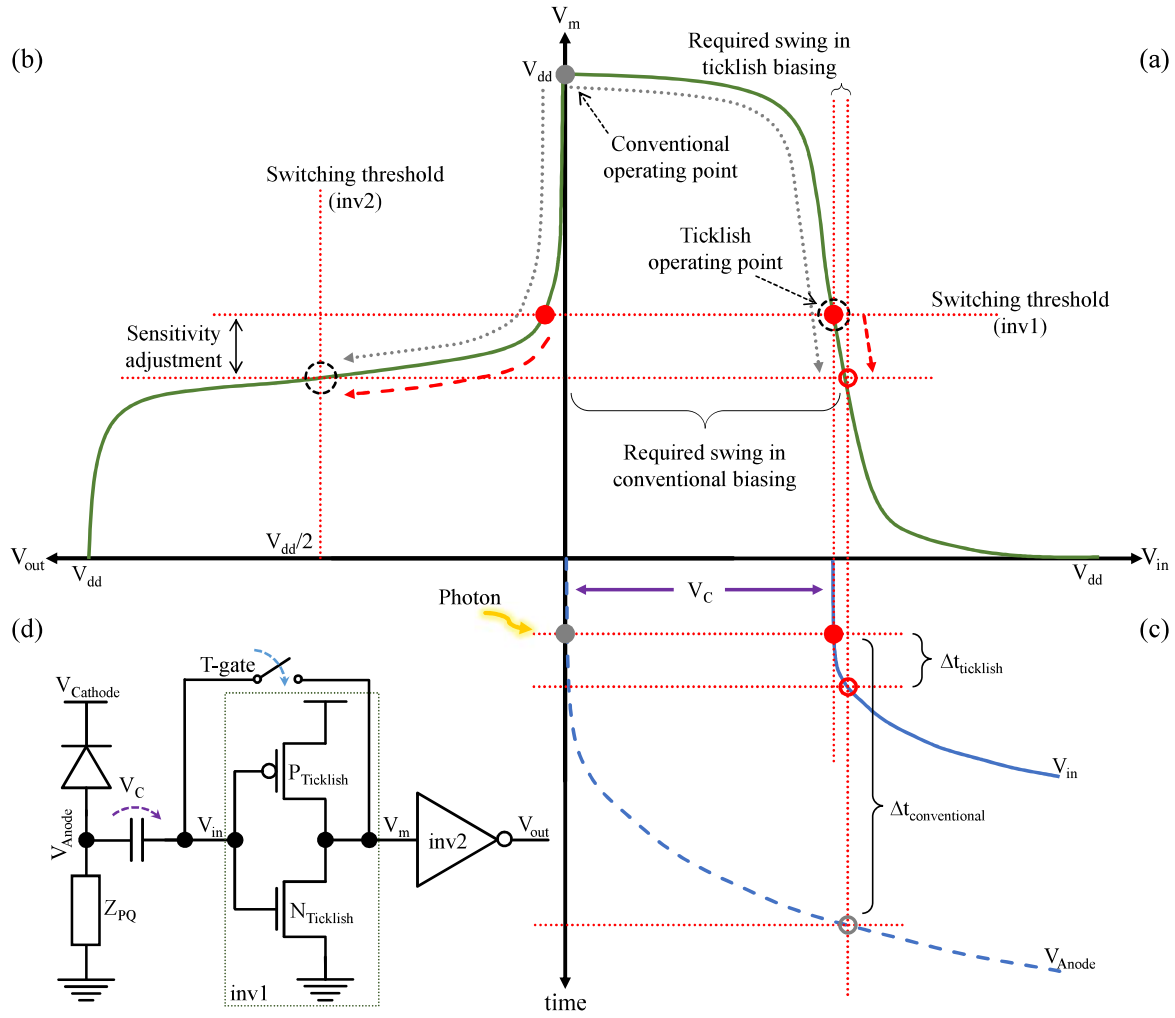


Fig. 3.2 Illustration of how the ticklish detector works. (a) VTC of a typical inverter that can be a ticklish detector by being biased at the middle point (b) VTC of the following inverter with a lower switching threshold than its precedent (c) Anode voltage and input voltage of the ticklish inverter (V_{in}) versus time during an avalanche (d) schematic of the proposed ticklish detector.

middle point (Fig. 3.2 (a)). This voltage swing is translated to the time domain as presented in Fig. 3.2 (c). Thanks to the low input voltage swing required in the ticklish inverter, the avalanche detection time should be shorter than the time needed for a conventional inverter. Consequently, since a faster avalanche detection means a higher afterpulsing reduction, this approach should allow us to obtain even better metrological performances. There are a few remarks that are crucial for designing the ticklish detection circuit. First, precisely choose the switching threshold of the following inverter (inv2). From Fig. 3.2 (b) one can see if the middle point of inv2 be higher than of the ticklish inverter, V_{out} would be around V_{dd}, and by increasing the anode voltage at each

avalanche, it becomes higher till reaches V_{dd}. The circuit would be locked in a non-functioning loop that cannot detect the avalanche. Thus, this is mandatory to design inv2 to have a lower middle point than inv1. The following equation shows how to tune the switching threshold of the inverter by sizing its transistors:

$$V_M = \frac{V_{dd} - |V_{th_P}| + V_{th_N} \cdot \sqrt{\frac{\mu_N \cdot (\frac{W}{L})_N}{\mu_P \cdot (\frac{W}{L})_P}}}{1 + \sqrt{\frac{\mu_N \cdot (\frac{W}{L})_N}{\mu_P \cdot (\frac{W}{L})_P}}} \quad (3.2)$$

Where V_{th_N} and V_{th_P} are threshold voltages of NMOS and PMOS. μ_N and μ_P are mobilities of electrons and holes, respectively, and $(\frac{W}{L})_N$ and $(\frac{W}{L})_P$ are width and length ratios of NMOS and PMOS.

The closer the two middle points the more sensitive the ticklish detector. However, this sensitivity is limited by the maximum tolerable noise at V_{in}, and the mismatch and process variations between the two inverters. Therefore, a Monte Carlo simulation can define the sensitivity level due to the random mismatch and process variations. Besides, an optimum layout can decrease the sensitivity limit due to the systematic mismatch. Otherwise, the noise or any random threshold voltage fluctuations of the MOSFETs of the inverters can result in a false avalanche detection and decrease the reliability of the circuit. In this regard, we performed two Monte Carlo simulations; one for V_m (output of the ticklish inverter), when the T-gate is switched off, and the other for the switching threshold of the second inverter (inv2). Fig. 3.3 shows the statistical distributions of the voltages of these two nodes. In the current design, 60 mV margin (sensitivity level) is chosen between the two middle points. This results in 20 mV margin between these two points in Monte Carlo simulations. No overlap between the histograms ensures the reliability of the design, while it would be possible to increase the sensitivity of the circuit by another 20 mV.

Another crucial design remark is the charge injection generated by the T-gate. A T-gate is composed of an NMOS and a PMOS in parallel. When these transistors are switched on and off, they inject some charges at V_{in} and V_m. This is important especially when the T-gate is switched off and the detection inverter is waiting for a new avalanche. At this moment, any increase at V_{in} due to the charge injection would result in a fake avalanche detection. When turning off the T-gate,

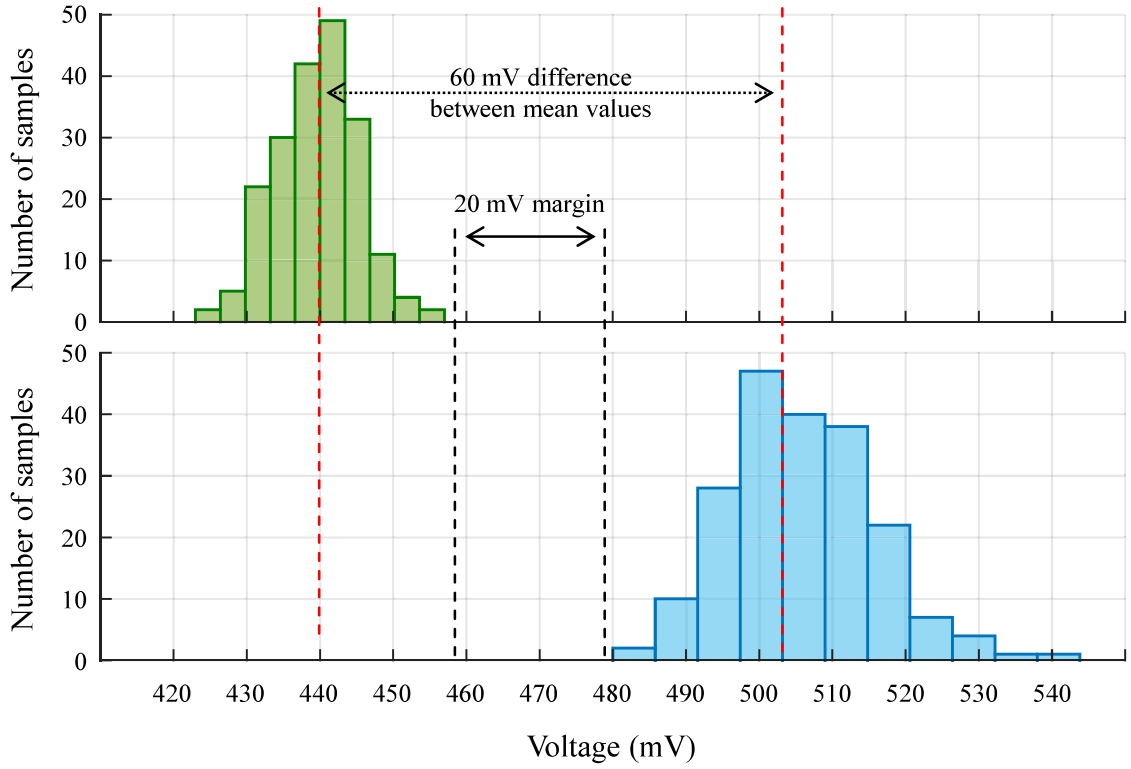


Fig. 3.3 Monte Carlo simulation results for the difference between middle points of inverters in ticklish detector. Green bars show the switching threshold distribution of the switching threshold of the second inverter (inv2) for different samples. Blue bars show the distribution of the output voltage of the ticklish inverter (V_m) after switching off the T-gate, for the same number of samples.

charges are injected through the PMOS gate while others are tapped out through the NMOS gate. Therefore, it is recommended to choose a stronger NMOS than PMOS, to not only compensate the injected charge, but also to insure to obtain a small voltage decrease at the input of the inverter when opening the T-gate to prevent false detections.

3.2.2. *Ticklish Active Quenching Active Reset Circuit*

Based on the proposed detection circuit in the previous section, an AQAR circuit is presented in Fig. 3.4, named ticklish AQAR circuit. The ticklish inverter is composed by P_1, N_1 and a T-gate. Inverter 2 and 3 (inv2 and inv3) are conditioning the detection signal to drive the wide AQ PMOS, P_2. A tunable delay line controls the hold off and reset times of the SPAD by driving the T-gate and the Active Reset (AR) NMOS, N_2. The delay line in this circuit is composed of two of the delay lines described in chapter two, connected together, to be able to control the T-gate too. A tunable Passive Quenching (PQ) resistor is implemented by N_3 for measurement purposes and C_c is the

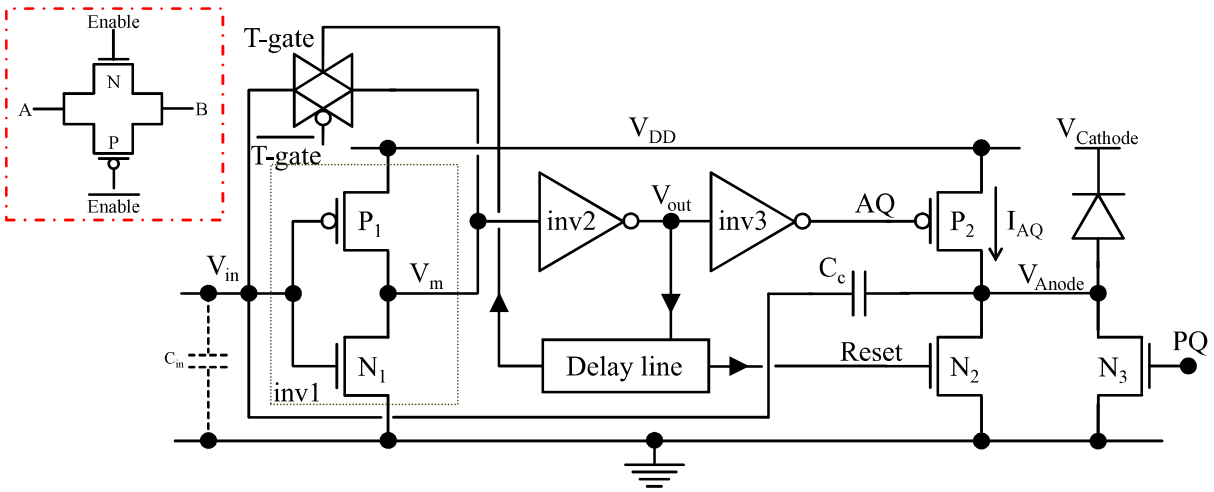


Fig. 3.4 Schematic of the proposed AQAR circuit, based on the ticklish detector. C_{in} is all the capacitances seen at V_{in}. All the bodies are connected to their nominal values. Delay line and passive quenching transistor can be controlled externally. Inset: schematic of a T-gate.

coupling capacitor that transfers anode voltage variations to the input of the detection inverter, forming a capacitive divider with C_{in}:

$$\Delta V_{in} = \frac{C_c}{C_c + C_{in}} \cdot \Delta V_{Anode} \tag{4.4}$$

Fig. 3.5 shows the timing diagram of the ticklish AQAR circuit. To operate the circuit, firstly the T-gate is switched on and off. This is biasing the detection inverter at its middle point. When an avalanche occurs (t_0), the anode voltage starts to rise. The ticklish detector senses this through the capacitive divider. A voltage equal to the difference of the middle points of the two inverters (inv1 and inv2) is enough to detect the avalanche. This detection at very early stages of the avalanche allows to drive P_3 quickly and raise the anode voltage to bias the SPAD out of the breakdown region (t_1). Thus, the avalanche is quenched faster than in PQ mode. After a certain delay (t_2), T-gate is switched on to prepare the ticklish detector for the next event by biasing the inverter at the middle point. At the end of the desired hold off time (t_3), N_2 resets the SPAD to its breakdown region. After the reset phase (t_4), SPAD is ready for photon detection once again.

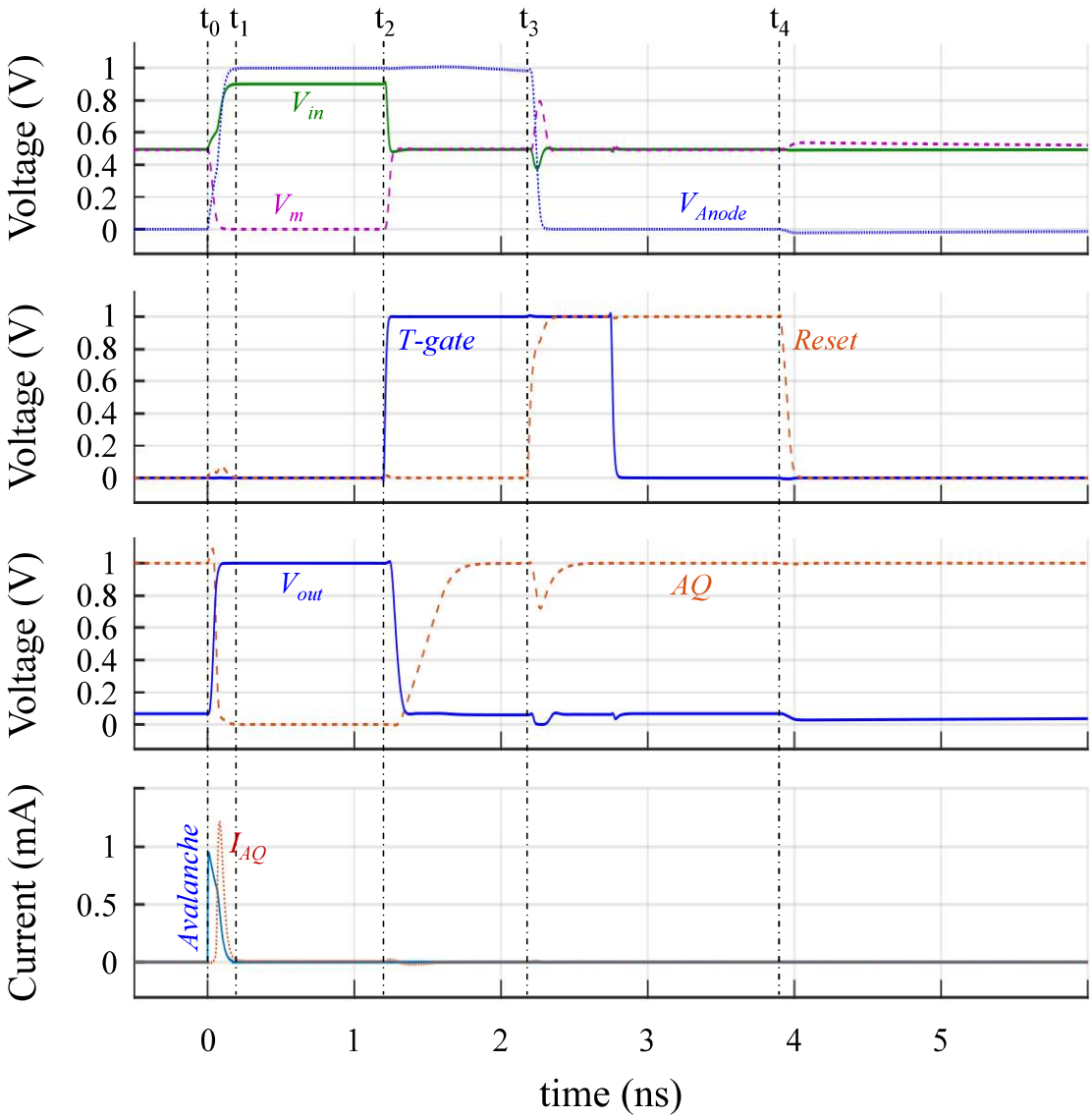


Fig. 3.5 Timing diagram of the proposed AQAR circuit. At t_0 avalanche starts. At t_1, it is quenched. T-gate is switched on at t_2 to re-bias the ticklish detector. Active reset phase starts at t_3 and finishes at t_4, where the SPAD is ready for a new avalanche.

3.3. Native 3D SPAD Active Pixel

Added electronics for avalanche detection and active quenching, reduces the fill factor of the pixel by expanding its inactive (light insensitive) area. There is a proportional relationship between the fill factor and the Photon Detection Efficiency (PDE). Therefore, a trade-off exists between extra circuitry required for the afterpulsing reduction and the PDE. To solve this issue, the 3D stacked structure is proposed in the literature. In this method, SPADs and all the supplementary

electronics are implemented in two different dies and then attached together through specific bonding processes. Thanks to this method, it is possible to use customized technologies for each of the two dies. Even though this solution appreciably enhances the fill factor, it is evident that the cost and the fabrication complexity are much higher than in a standard CMOS process.

Recently, a new SPAD pixel is proposed in FD-SOI CMOS technology in [154]. The pixel allows integration of the electronics over the SPAD in the same die. This is done through an insulator between the SPAD and the associated electronics, yielding a 3D-like structure that we call it native or inherently 3D SPAD pixel. Even though it seems promising in mitigating the trade-off between the afterpulsing, PDE, and the cost, no attempt has been made to implement an active quenching circuit for this pixel. Only two sensing circuit, composed by a capacitive and a resistive divider are introduced in [161]. In the rest of this section, we investigate the challenges of placing electronics over the SPAD active area. For the first time, an AQAR circuit, featuring an ultrafast active detection circuit is presented that is compatible with the native 3D SPAD pixel.

3.3.1. *Native 3D Pixel Structure*

In Fig. 3.6 (a), a cross section of the FD-SOI CMOS technology is presented. In comparison with the Bulk CMOS technology (Fig. 3.6 (b)), in FD-SOI CMOS, transistors are fabricated over a Buried Oxide layer (BOX). This structure yields electrical isolation for the active devices over the BOX which is the key point in design of the inherently 3D SPAD pixel. In the technology that has been used in this work, two different MOSFET families exist; flipped well and regular well. In flipped well MOSFETs, NMOS is fabricated in an N-well and PMOS in a P-well while in the regular well family it is the reverse, see Fig. 3.6 (a). By respecting electrostatic considerations, it is possible to mix these two MOSFET categories and have both NMOS and PMOS in the same well as it is shown in Fig. 3.6 (c). In the same figure, it is also observable that the SPAD is formed at P-well and deep N-well junction. Thanks to the insulating BOX layer, the MOSFETs are electrically isolated from the SPAD while placed over it and surrounded by its active area. In a Back-Side Illumination (BSI) regime, this inherently 3D structure can drastically reduce or remove the dependence of fill factor and PDE on SPAD associated circuit. Even though, to use BSI, a die-thinning process is required, yet the manufacturing process cost of the used FD-SOI CMOS technology is 10% less than the classical bulk CMOS technology [168].

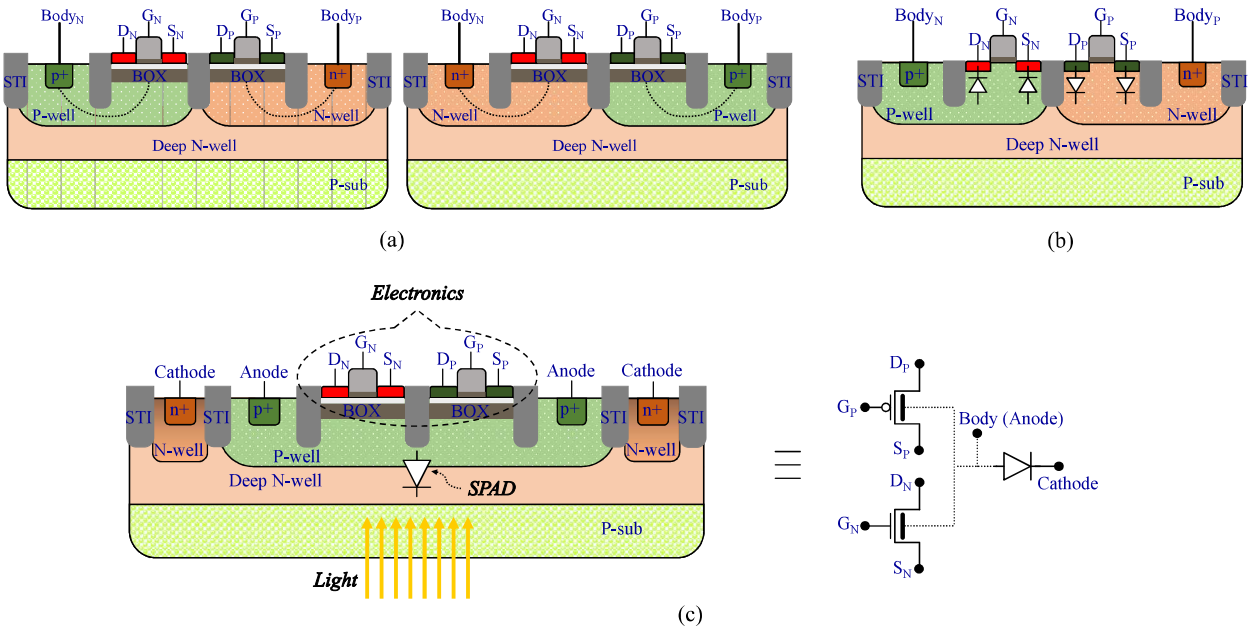


Fig. 3.6 Symbolic cross section of: (a) FD-SOI CMOS technology, regular well (left) and flip well (right), (b) bulk CMOS technology, (C) mix of the flip well and the regular well in the FD-SOI CMOS technology and forming of the SPAD at the junction of P-well and the deep N-well which results in the intrinsic 3D pixel (left) and its schematic counterpart (right).

3.3.2. *Placing the Circuit Over SPAD*

Thanks to the isolation imposed by BOX layer, the parasitic diodes of Source/Well and Drain/Well are not formed. In the bulk CMOS, these parasitic diodes (with a threshold of about 0.6 V), strictly limit the body biasing. While in the FD-SOI CMOS technology, the maximum applicable voltage to the body terminal of a MOSFET is limited by the maximum sustainable voltage by the BOX layer which acts as a second back gate. In the technology used in this thesis, the BOX layer is 25 nm thick, with the maximum applicable voltage to the back gate or body of the MOSFETs is 3 V according to the design rules of the technology.

In the previous chapter, it is shown that it is possible to exploit the body biasing feature to modulate the threshold voltage by applying an external voltage to the body of each transistor and design a fast avalanche detection and quenching circuit. However, in the case of the native 3D pixel, there is no possibility to apply an external body voltage to the transistors over the SPAD. One can see in Fig. 3.6 (c), that by applying a proper fixed bias voltage to the deep N-well -as the cathode of the SPAD-, when an avalanche is generated, the voltage of the P-well -as the anode of the SPAD-, starts to change. This voltage variation in the P-well, which is also the body of MOSFETs placed on top of it, modulates their threshold voltages. This is an uncontrollable body

biasing which is imposed by the SPAD to all the devices that are in the P-well. In the case of the MOSFETs, an anode voltage increase decreases the threshold voltage of NMOS and increases the threshold voltage of PMOS (in absolute value). In the used FD-SOI technology, the body biasing factor is about 0.08, i.e., for an excess bias voltage of 1 V, the threshold voltage changes by 80 mV. Therefore, to have an AQAR circuit compatible with this structure, its detection circuit must be robust against these variations. This undesired body biasing also increases the threshold of the AQ PMOS, which is a little drawback.

Choosing cathode as the dynamic node and fixing the anode at a constant voltage could solve the whole unwanted body biasing issue. However, in this case the realization of native 3D pixel is impossible since BOX layer could not sustain the high breakdown voltage of the SPAD.

One can notice that in this pixel structure, it is also possible to sense the avalanche only through the bodies of the MOSFETs, without any wired connection between the SPAD and the detection circuit. This new approach in the avalanche sensing is called indirect avalanche sensing [161]. Besides its novelty, the most important advantage of this approach is that it can tolerate a voltage more than the technology voltage headroom. Thus, when an avalanche detection circuit exploits the indirect sensing approach, it can tolerate an excess bias voltage higher than the technology headroom. This interesting feature is especially appreciated in more scaled technology nodes since their voltage headrooms are low and can pave the way for denser (higher resolution) SPAD based imagers in ultra-scaled technology nodes. Also, a higher excess bias voltage can reduce the jitter and improve PDP of the pixel. In this case and with the proposed structures, the SPAD cannot be actively quenched with an excess voltage above 1V.

3.3.3. *Mixed Sensing and Indirect Sensing Pixels*

Removing C_c from ticklish detector (Fig. 3.2), and placing it over the SPAD, results in the circuit shown in Fig. 3.7 (a). T-gate biases the inverter at its middle point and then is switched off. Now the circuit is an open loop amplifier biased close to the switching threshold of the inverter. Its input signal is the anode voltage, applied to its body terminals, and its gain is formulated as dV_{out}/dV_{body}. By starting an avalanche, the body (anode) voltage raises, therefore, the threshold voltage of NMOS and PMOS decreases and increases, respectively. This pulls back the VTC of the detection inverter and reduces its switching threshold by 80 m/V, while the input of the inverter is not moving. On the new VTC, the same previous V_{in}, results an output voltage close to zero that

flags avalanche detection. We call this ticklish InDirect Sensing (InDS) circuit since there is no connection between the gate of the MOSFETs and the SPAD. In this regard, since now we call the first detection circuit (Fig. 3.2), ticklish Direct Sensing (DS) circuit, and the AQAR circuit based on it, is named ticklish DS pixel.

The operation principles of the InDS circuit can be expressed in a simpler way. One can suppose VTC is not moving, instead V_{in} is increasing by 80 mV/V. According to the inverter gain at the switching threshold, the output voltage variation is:

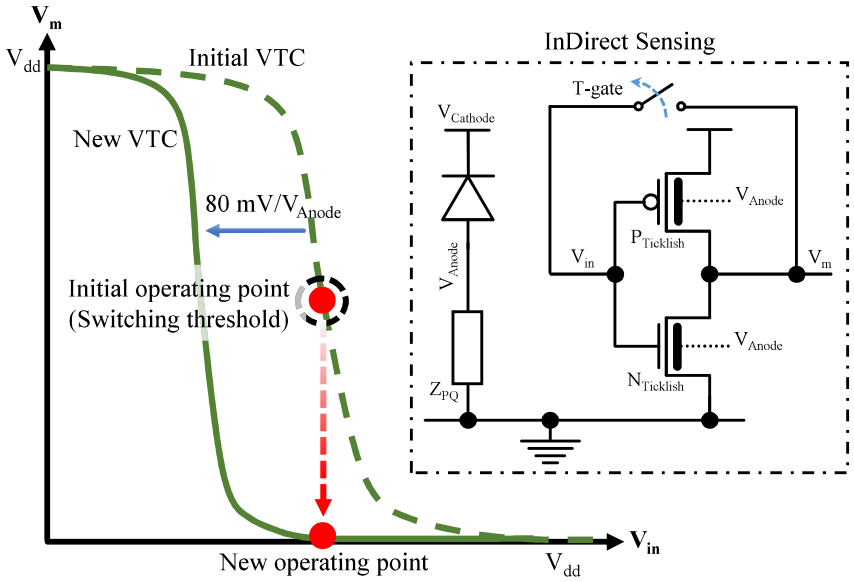
$$\Delta V_m|_{S.th} = \left.\frac{dV_m}{dV_{in}}\right|_{S.th} \times \Delta V_{in}$$

Where, $\left.\frac{dV_m}{dV_{in}}\right|_{S.th}$ is the large signal gain of the inverter at its switching threshold. For a typical inverter in the used technology, this is around 12. Thus, a maximum variation of about 800 mV is expected at the output for 80 mV input variation. Even though this gain is falling due to the nonlinearity, half of this output variation is enough for an avalanche detection.

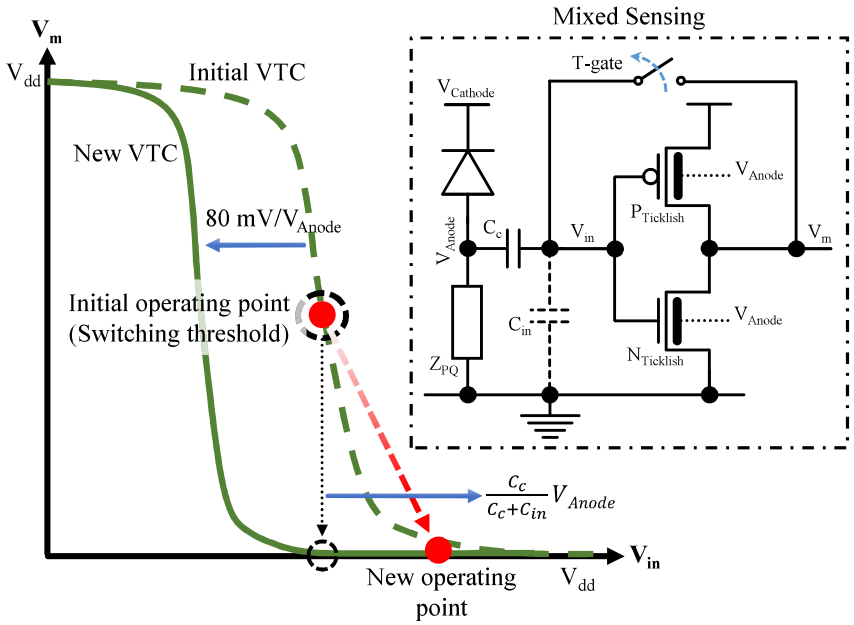
Based on this sensing circuit, a new AQAR circuit is presented in Fig. 3.8 (a), named ticklish InDS pixel. All the body terminals are equal to the anode voltage; however, the circuit is exploiting this uncontrollable body biasing to detect the avalanche, and the ticklish auto biasing feature makes it robust against body voltage variations. Even though the 80 mV/V body biasing factor is not high, the ticklish biasing is sensitive enough to make possible the design of a functional AQAR circuit using this avalanche sensing method. It is obvious that the AQ efficiency of this pixel is less than the DS pixel. Nonetheless, the InDS pixel is compatible with native 3D structure. The rest of the functionality of the pixel is similar to the DS pixel.

Adding C_c to the InDS circuit results in a new avalanche detection circuit shown in Fig. 3.7 (b). This circuit features the two sensing factors of the previous circuits: indirect avalanche sensing through the body and direct avalanche sensing through the gate. Thus, we call this: ticklish Mixed Sensing (MS) circuit. The avalanche occurrence raises the anode voltage, through the bodies, the VTC is pulled back, and through the C_c, the anode voltage rise is transferred to V_{in} and sensed at the gates. Therefore, the MS detection circuit is faster than the other two detection circuits, since InDS and DS detection circuits each suffer from the absence of one of these sensing factors. Even though it is also possible to apply the anode voltage to the bodies of transistors in DS pixel through

the wires and benefits from the imposed body biasing in a more controlled way, however, the resulting pixel will not be a native 3D pixel. Based on MS circuit, the MS AQAR circuit (MS pixel) is presented in Fig. 3.8 (b). Thanks to the ticklish biasing, body biasing, and sensing the avalanche through both gates, MS pixel should be the most efficient pixel in the avalanche detection, quenching, and the afterpulsing reduction. Except the detection part that is already explained, rest of the circuit operation is same as the DS pixel, explained in the section 3.2.2.



(a)



(b)

Fig. 3.7 Schematic and VTC of: (a) InDirect sensing circuit (b) Mixed sensing circuit.

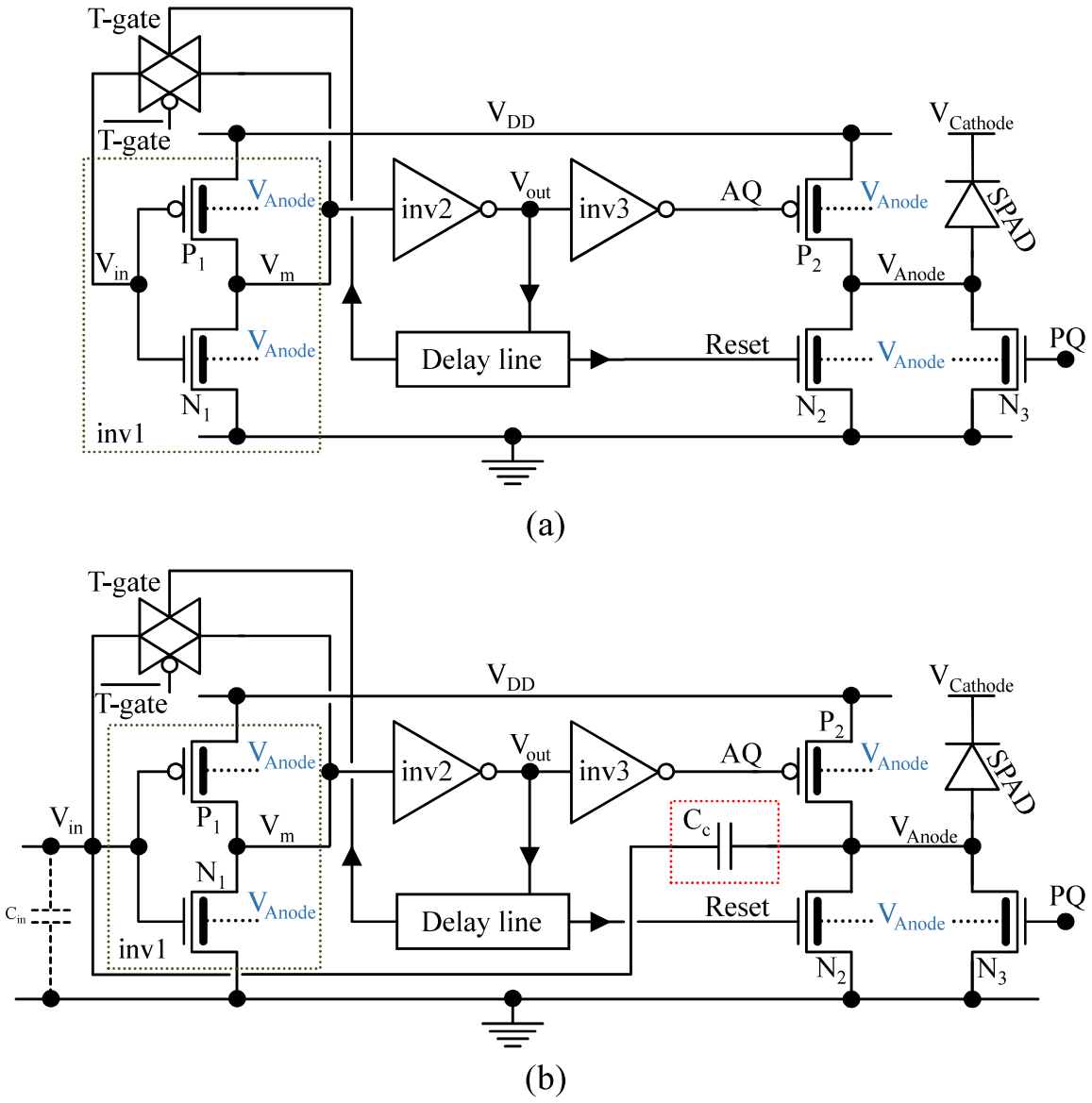


Fig. 3.8 Schematics of the proposed AQAR circuits: (a) InDirect Sensing (b) Mixed Sensing.

3.4. Results and Discussion

3.4.1. *Post Layout Simulation Results*

The proposed DS, MS, and InDS pixels have been designed and laid out in the 28 nm FD-SOI CMOS technology. Fig. 3.9 shows the layout of the pixels, featuring the ticklish AQAR circuits. The used SPAD is a quasi-octagonal SPAD with an active area of 490 μm^2, presented in [154]. Each AQAR circuit has an area of less than 14×15 μm^2. Most of the area is occupied by the delay line (8×15 μm^2), AR NMOS N_2 (2.5×5 μm^2), and AQ PMOS P_2 (2.5×5 μm^2). The detection part (inv1, T-gate, inv2, and inv3) and the PQ NMOS N_3, occupy an area of only 3×4 μm^2. Thus, for the applications that do not need an AQ or AR circuit, the proposed circuit can be smaller. The placement of the electronics over SPAD is observable in Fig. 3.9. In the two native 3D pixels (MS and InDS), the AQAR circuits cover 42 % of the SPAD active area. Thus, it is possible to integrate more electronics to the pixel such as a counter or a time to digital converter. In a BSI regime, these pixels totally remove the dependence of the fill factor and the pitch on the interface electronics.

For measurement purposes, there are also two large switches that allow switching between AQ/PQ and AR/PR. To model the SPAD for simulations, a SPICE macro model is used [167]. The junction capacitance and the series resistance

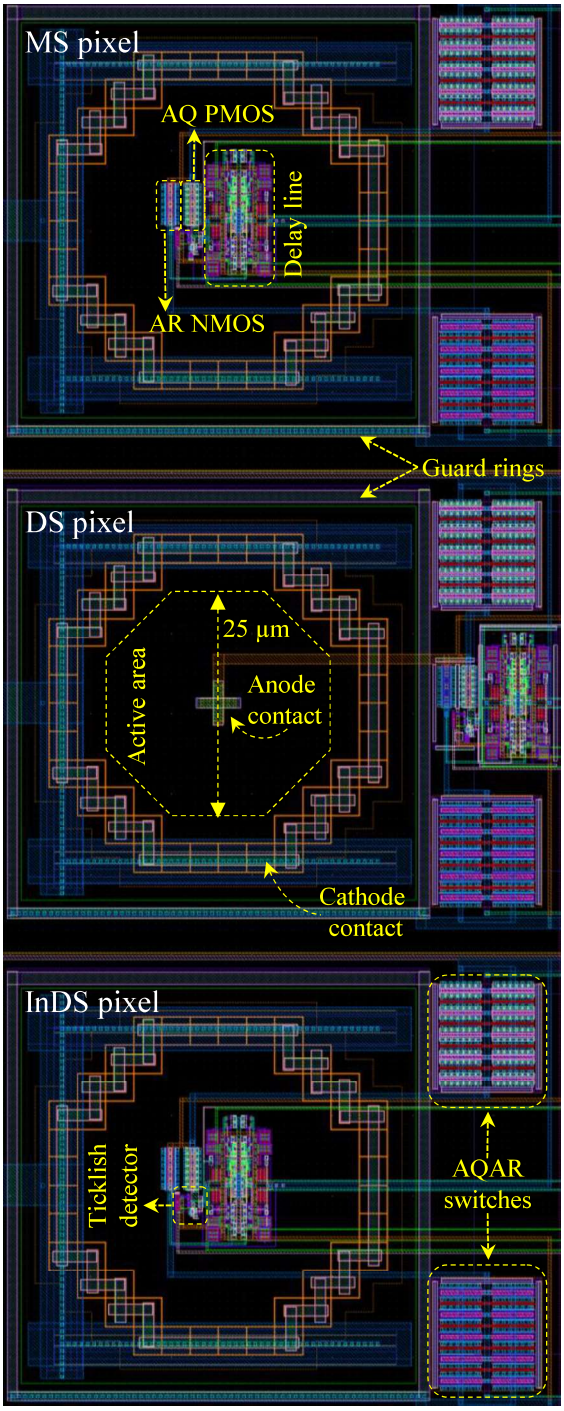


Fig. 3.9 Layouts of mixed sensing, direct sensing, and indirect sensing pixels. The AQAR switches are in series with AQ PMOS and AR NMOS for measurement purposes and are excluded from the pixel area.

of the SPAD are set to 100 fF and 1 kΩ, respectively. The excess voltage is set to 1 V which is the technology headroom. Fig 3.10, shows the simulation results for ticklish DS, MS and InDS pixels. From top to bottom, the anode voltage, avalanche and AQ PMOS (P_2) currents (I_{AQ}), and their charges are shown. To evaluate the functionality of the AQAR circuits, each simulation has been done in both PQ and AQ modes. In DS, about 60 ps after the anode voltage starts to rise, a deviation from the PQ curve is observable in the AQ curve (Fig. 3.10 (a)). It means that the detection circuit, has detected the avalanche sooner than 60 ps and drives P_2 to intervene in the avalanche process. The result of this intervention is a faster quenching as it is seen in both anode voltage and avalanche current curves (Fig. 3.10 (b)). Quenching time in AQ is about 130 ps, while this time is about 480 ps for PQ. Note that in these simulations we consider the point which avalanche current reaches 50 µA, as the end of the quenching phase. The area under the avalanche current curve is equal to the avalanche charge in SCR. In the PQ mode for the DS circuit this value is about 135 fF and in AQ mode is 75 fF, which corresponds to 44 % avalanche charge reduction. The amount of reduced charge is 60 fF which is equal to the charge contribution of the AQ PMOS (P_2). This 44 % charge reduction should result in 44 % afterpulsing reduction.

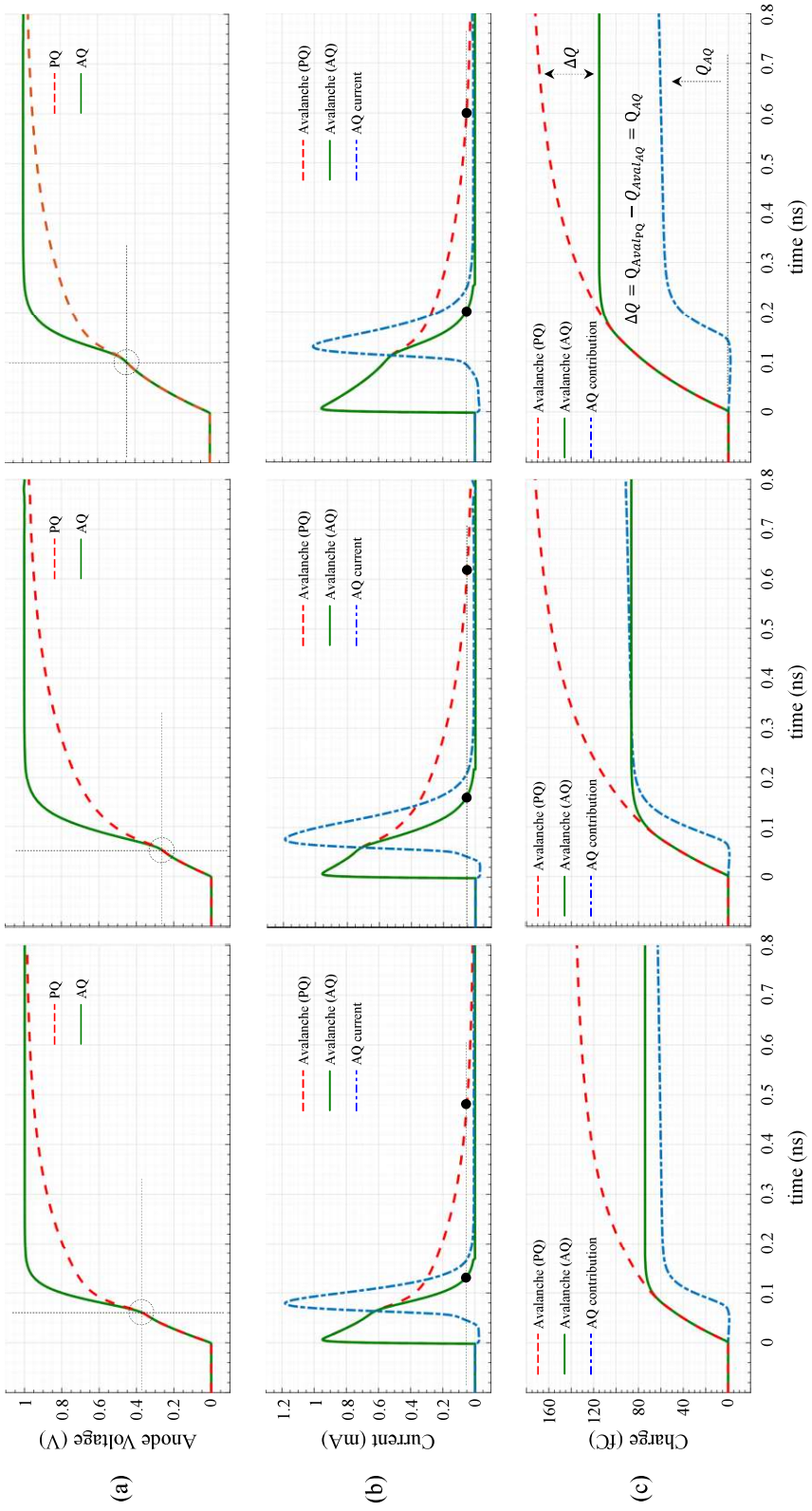


Fig. 3.10 Post layout transient simulation results during the occurrence of an avalanche in both active and passive quenching modes, for DS, MS, and InDS (from left to right, respectively). (a) Anode voltage (b) Avalanche and active quenching currents (c) Avalanche and active quenching charge.

Table 3.1 shows the simulation results for the three pixels. As expected, MS circuit exhibits the shortest detection time while InDS has the longest. This quick detection causes the highest AQ current among other pixels and the longest detection time results in the minimum AQ current and thus, the minimum charge reduction of 34 % in InDS pixel. Despite its shorter detection time, MS pixel has a longer quenching time than the DS and yet despite this, its charge reduction is more. This is because: placing the circuit over the SPAD, adds parasitic capacitances (C_{bd}, C_{bs}, C_{bg}) to anode of the SPAD. Therefore, the anode voltage rise is slowed down and the AQ circuit becomes more effective. This is further observable in Fig. 3.10, by comparing the SPAD PQ charges between the DS and the two native 3D pixels. The PQ charge in the native 3D pixels is higher than the DS pixel due to the combined body and anode. Thus, it is more objective to compare the AQ charge of the native 3D pixels with the PQ charge in DS pixel, instead of their own PQ charge. By doing so, one can see that the MS circuit reduces 36 % of the avalanche charge, and this value is 15 % for the InDS pixel. It can be concluded that for ultra-fast photon detection, the MS circuit is the best candidate. However, its effectiveness in the avalanche charge reduction is affected by the parasitic capacitances at the body terminals of all the transistors, which are placed over the SPAD.

The total static power consumption of the circuit is around 100 µW, which mostly consumed by the two delay lines. The quenching circuit itself consumes around 43 µW which 23 µW of it is consumed only by the detection inverter.

TABLE 3.1
POST LAYOUT SIMULATION RESULTS FOR DS, MS, AND INDS PIXELS, IN ACTIVE AND PASSIVE QUENCHING MODES.

	Direct Sensing	Mixed Sensing	InDS Sensing
Avalanche detection time (ps)	< 60	< 50	< 100
Quenching time in PQ (ps)	480	620	600
Quenching time in AQ (ps)	130	160	200
Avalanche charge in PQ (fC)	135	175	175
Avalanche charge in AQ (fC)	75	86	115
AQ charge contribution (fC)	60	89	60
Avalanche charge reduction (%)	44	51	34

3.4.2. *Experimental Results*

The micrograph of the pixels and the test bench are shown in Fig. 3.11 and Fig. 3.12, respectively. This is the first integration of electronics over SPAD. While the DS pixel can work under a Front-Side Illumination (FSI) regime, the two native 3D pixels should work in a BSI regime. This is how the native 3D pixels remove the dependence of the PDE and the pitch, on the SPAD associated electronics. Thus, the required circuitry for afterpulsing reduction is not a limiting factor anymore. The pixels under the test are integrated in a 1×1 mm^2 chip, which is wire bonded on a daughter board. The daughter board is plugged into the same mother board used and described in the previous chapter. On the daughter board an SMA connector is directly connected to the output of a wideband analog voltage buffer that enables us to monitor the anode voltage transitions. The analog output is measured by a 13 GHz wideband oscilloscope. To perform the measurements, the test bench is placed inside a thermal chamber in the dark

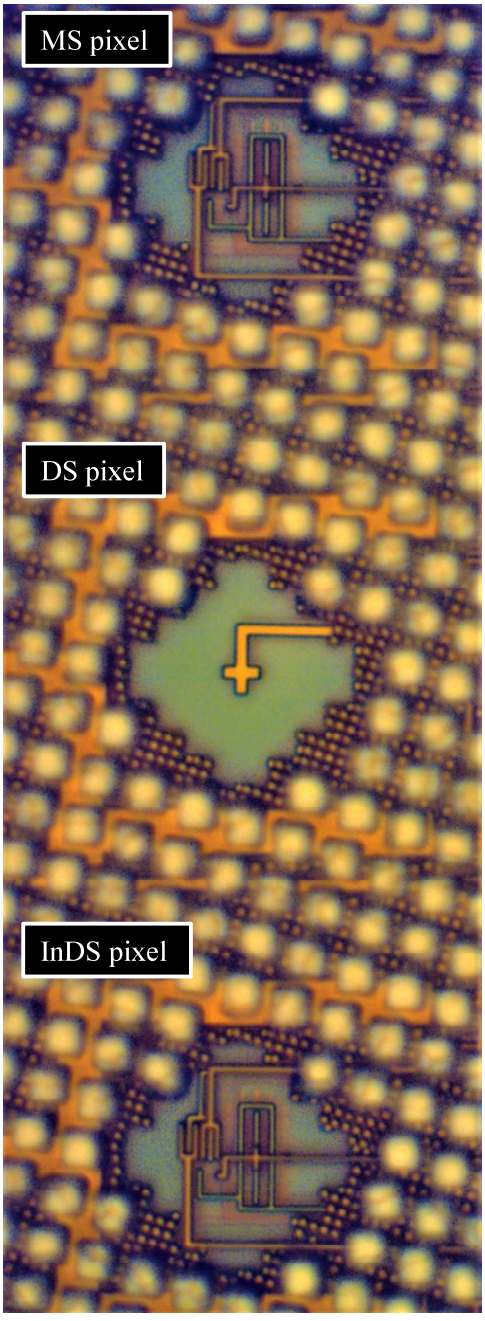


Fig. 3.11 Micrograph of the pixels. From top to bottom: Mixed Sensing, Direct Sensing, and Indirect Sensing.

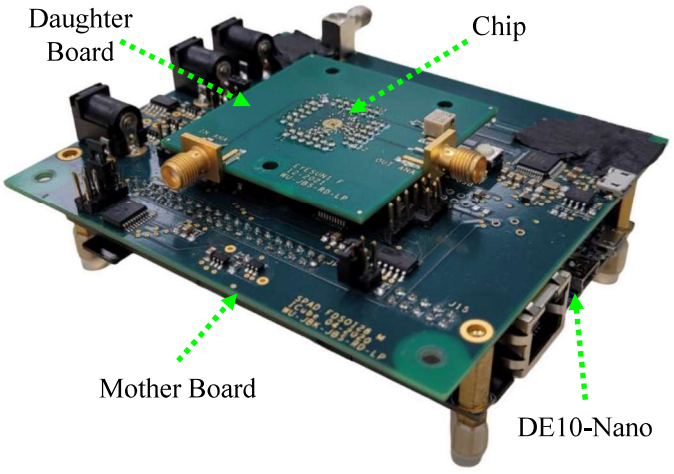


Fig. 3.12 Test bench setup, including a 1 mm^2 chip in a 28 nm FD-SOI CMOS technology, daughter board, mother board, and a FPGA DE10-Nano board.

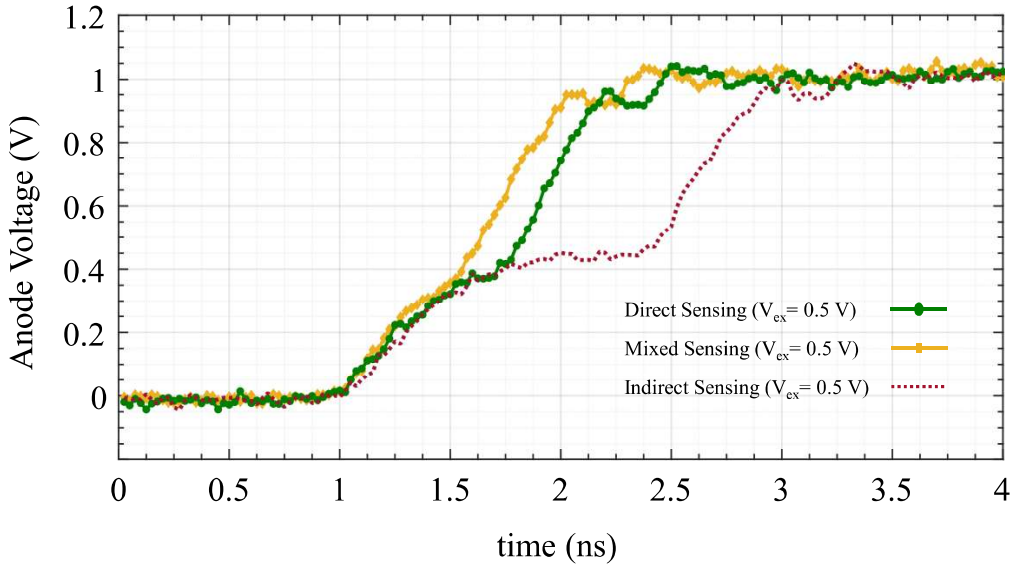


Fig. 3.13 Measured anode voltage variations of Direct Sensing, Mixed Sensing, and Indirect Sensing pixels during an avalanche in AQ mode at 60 °C with 500 mV of excess bias voltage. Experimental counterpart of Fig. 3.10 (a).

condition. All the reported results are based on the DCR measurements while the SPADs are in the free running mode.

Fig. 3.13 shows the anode voltage transition of the SPAD during an avalanche in the AQ mode when the pixels are placed in the chamber at 60 °C. A slope change is observed as in the simulations (Fig. 3.10 (a)). For 500 mV of V_{ex}, the fastest avalanche detection time, as expected, belongs to the MS pixel which is less than 400 ps. Due to this fast avalanche detection, the native 3D MS pixel quenches the SPAD in only 625 ps. Thus, the reaction time of the circuit is as fast as 225 ps. To the best of our knowledge, it is the fastest avalanche detection and quenching time ever reported in the literature.

According to the measurements, the MS and InDS pixels cannot sustain an excess bias voltage as high as the DS pixel. In the PQ mode, without the assistance of the AQ circuit, the DS pixel can operate until 1 V of V_{ex}, while, in the same conditions, the native 3D pixels are saturated by noise for an excess bias voltage of slightly higher than 250 mV. This is due to the differences in their structure; placing electronics over the SPAD electrically and physically affects the device. This happened through adding the parasitic capacitances at the body terminals and Shallow Trench Isolators (STI), around each transistor over the SPAD. The STI, which is a silicon oxide, can be a source of traps in the SPAD structure. Also, the generation rate around STI is very high [155]. The

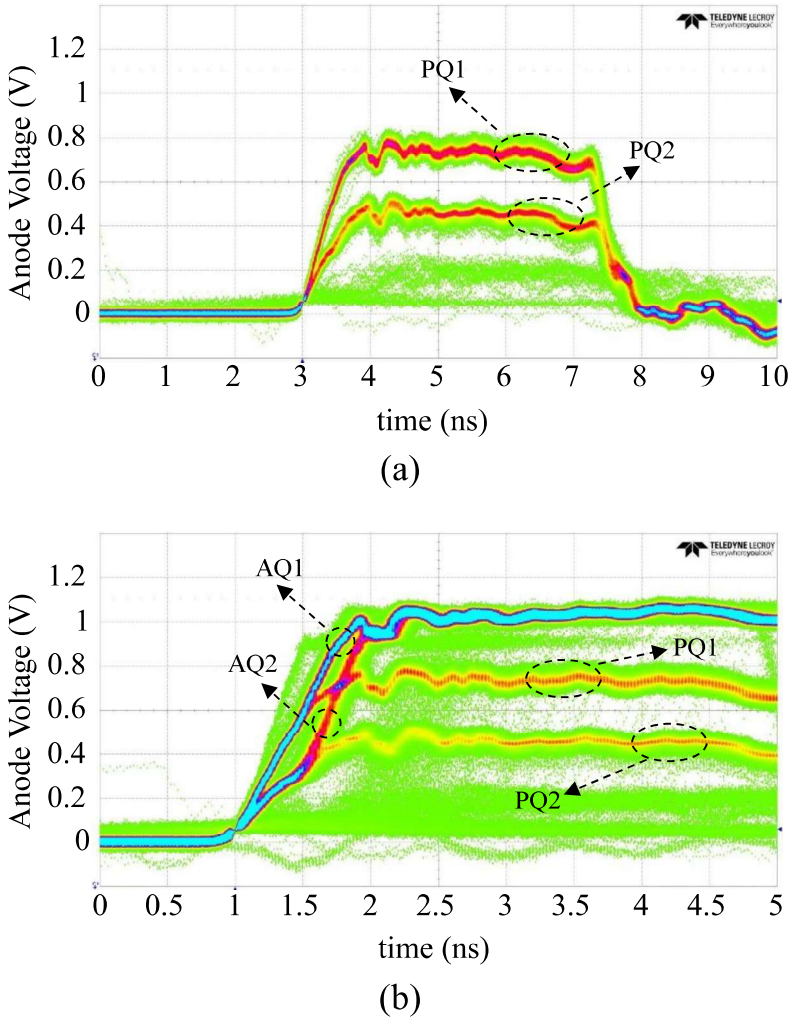


Fig. 3.14 Oscilloscope screen shot of the anode voltage transition of Direct Sensing pixel under 800 mV of V_{ex} at 60 °C when the oscilloscope is in the persistence mode: (a) passive quenching mode (b) when the active quenching circuit is enabled.

active area of the used SPAD is surrounded by STIs, and some are also placed inside the active area to respect the design rules of the used technology. Therefore, a very high DCR and afterpulsing is expected. Also, in this configuration, the STIs result in the Premature Edge Breakdown (PEB) as it is shown in [155]. In Fig. 3.14 (a), anode voltage variations of the DS pixel in the PQ mode are shown, while the oscilloscope is in the persistence mode. V_{ex} is set to 800 mV and the temperature is fixed at 60 °C. At this excess bias voltage several types of avalanches are observable: a main avalanche of 800 mV amplitude, another one with an almost same statistical frequency as the main one, with about 500 mV amplitude and a less frequent one with an amplitude of about

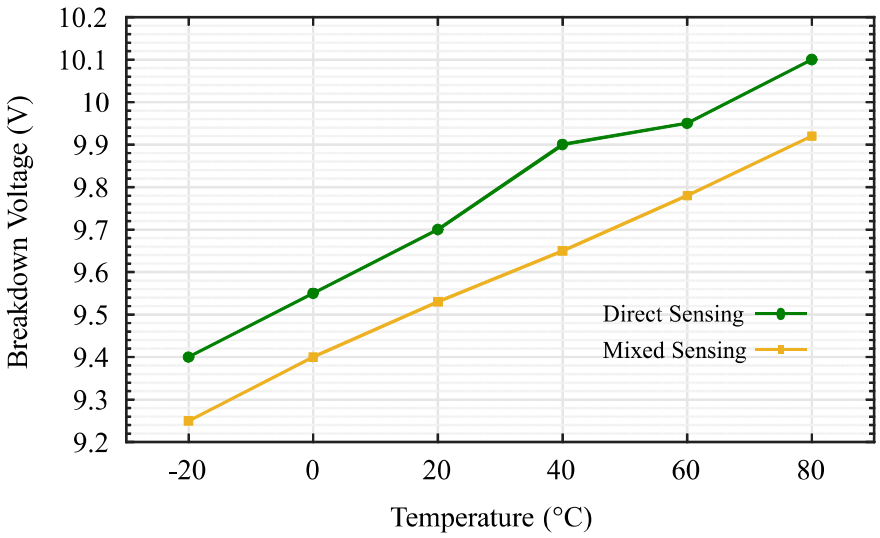


Fig. 3.15 Breakdown voltages of the SPAD in different temperatures for Direct Sensing and Mixed Sensing pixels.

200 mV. This is the evidence of PEB effect due to the STIs around the SPAD. Fig. 3.14 (b) shows the same condition as Fig. 3.14 (a), except that the ticklish AQAR circuit is enabled this time. The circuit is capable of detecting and quenching the different avalanches, and since the deviation of the AQ signal from the PQ signal is visible, it also efficiently reduces the afterpulsing effect.

Thus, increasing the number of STIs in the center of the active area due to the inherently 3D structure, results in a same effect as PEB, but in the center, we call it “premature breakdown”. Fig. 3.15 shows the breakdown voltage variations versus temperature for DS pixel and MS pixel. The breakdown voltage of the native 3D structure at every temperature is at least about 150 mV less than the non-3D pixel. These results are in agreement with the aforementioned explanations that STIs results in the premature breakdown.

Fig. 3.16 (a) shows that the MS pixel is saturated by the noise for a V_{ex} of higher than 250 mV, i.e., it is always triggered by the noise in the PQ mode. By activating the AQ circuit, at the same V_{ex}, the SPAD works normally (Fig. 3.16 (b)). This proves not only the effectiveness of the proposed circuit in the afterpulsing reduction, but also its vital role for the native 3D pixel. In other words, placing electronics over the SPAD in the used FD-SOI CMOS technology, increases the noise such that the pixel is not functional in higher excess bias voltages, nevertheless, the ticklish AQAR circuit is able to compensate for this added noise. For the InDS pixel, the situation is quite the same, since the only structural difference between two pixels is a coupling capacitor (C_c).

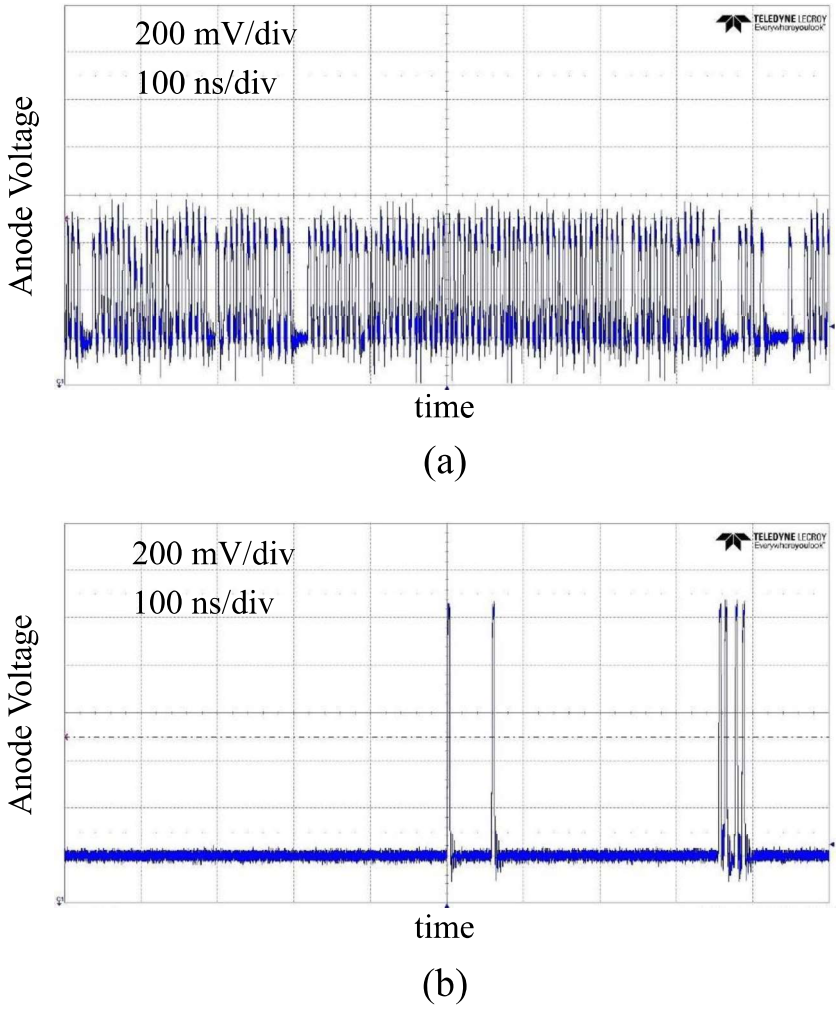
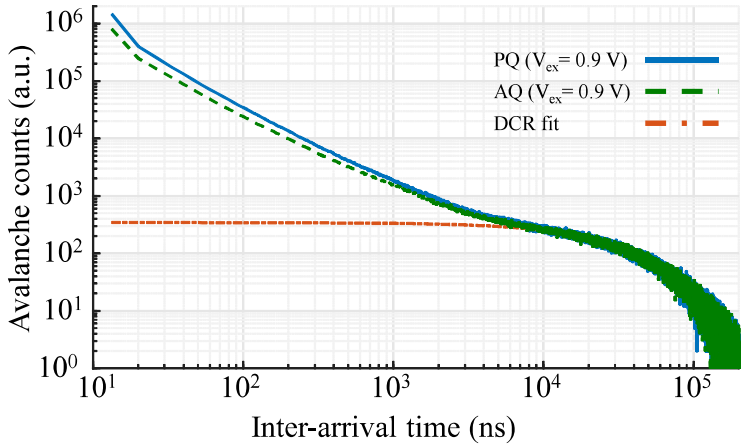


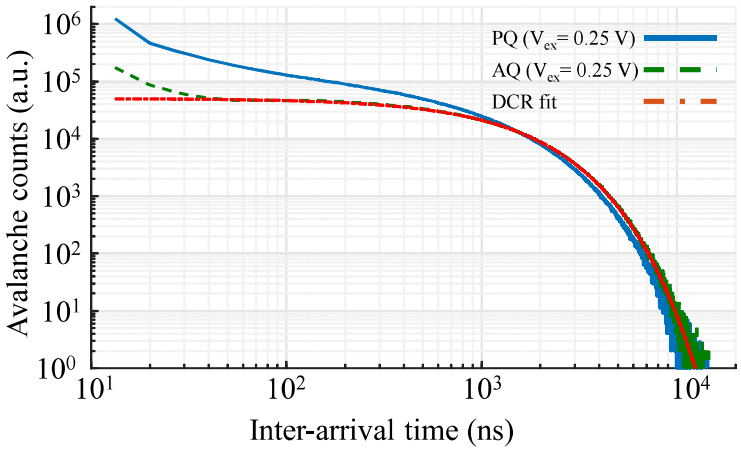
Fig. 3.16 Oscilloscope screenshots of avalanches in the Mixed Sensing pixel under 500 mV of V_{ex} at 60 °C: (a) passive quenching mode (b) active quenching mode.

However, since the sensitivity of InDS is much less than the MS pixel, only a higher excess bias voltage can be detected by it. Therefore, the InDS pixel is only functional in AQ mode for an excess bias voltage of around 500 mV. Regarding the 80 mV/V body biasing factor in the used technology, this translates to a detection threshold of about 40 mV for the ticklish detection circuit. Less than this excess bias voltage the event cannot be detected and more than that the SPAD becomes saturated by the noise.

These observations also are another justification for this hypothesis that STI is increasing the afterpulsing effect as well as the DCR, since the AQ circuits are not able to modify the DCR but can only reduce the afterpulsing effect. It can be concluded as well that the contribution of the



(a)



(b)

Fig. 3.17 Avalanche inter-arrival time histogram and DCR fit in passive and active quenching mode for Direct Sensing and Mixed Sensing pixels with 8 ns dead time at 60 °C. (a) Direct Sensing at 900 mV of excess bias voltage (b) Mixed Sensing at 250 mV of excess bias voltage.

afterpulsing effect in the total noise of the used SPAD is much more than the DCR, since the afterpulsing reduction saves the SPAD while the DCR has remained unchanged. This claim is further substantiated with the following afterpulsing measurements.

The inter-arrival time histograms of the pixels are shown in Fig. 3.17. The pixels are at 60 °C and a dead time of about 8 ns is set for them. Afterpulsing effect is measured by subtracting the fitted DCR curve from the inter-arrival time histogram. To obtain the fitted -afterpulsing-free- DCR curve, it is considered that after a certain time i.e., a few microseconds, the afterpulsing effect is very low, thus, the part of the histogram after this time can be used for the fitting. In the DS pixel

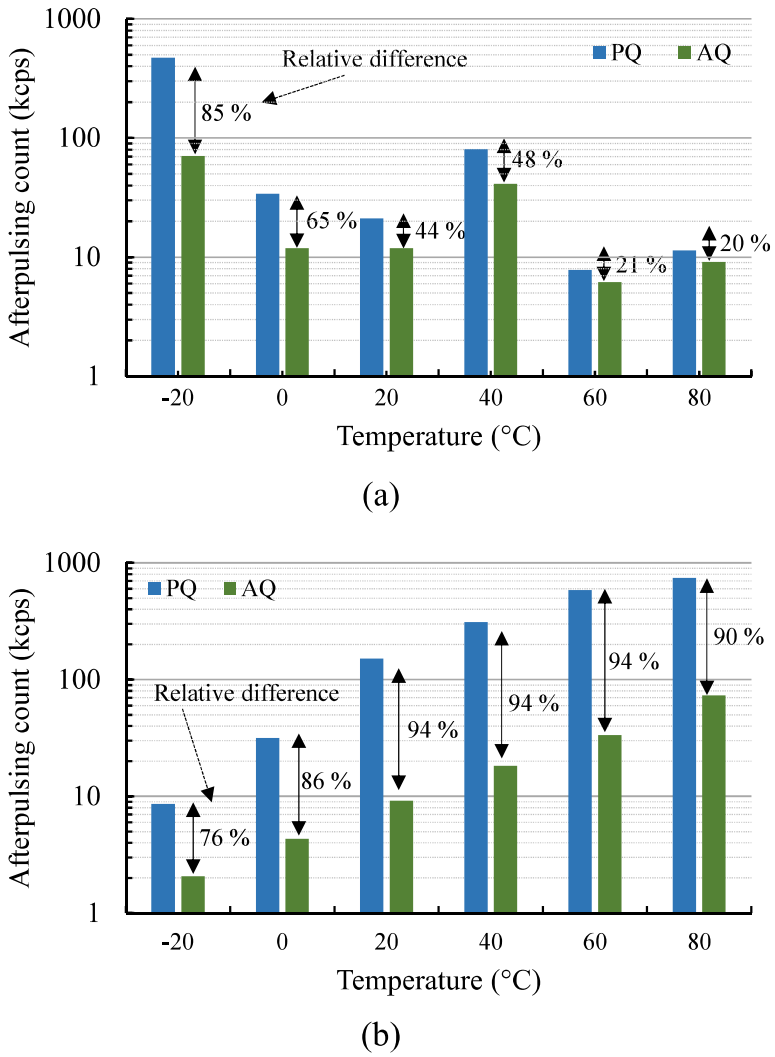


Fig. 3.18 Normalized afterpulsing counts of passive quenching mode (blue bars) and active quenching mode (green bars) at different temperatures with 8 ns dead time. (a) Direct Sensing at V_{ex} = 800 mV, (b) Mixed Sensing at V_{ex} = 250 mV.

(Fig. 3.17 (a)), the afterpulsing reduction is observable in the histogram even though the afterpulsing value is very high. It can be seen that there are traps which hold the carriers for longer than 10 µs. This histogram is also evidence of the small contribution of DCR (even with a high V_{ex} of 900 mV) in the total noise of the used SPAD in comparison with the afterpulsing effect. In the MS pixel (Fig. 3.17 (b)) this afterpulsing reduction is more visible. However, the comparison between AQ and PQ is limited to 250 mV of V_{ex}. Fig. 3.17 also demonstrates the very high impact of the top placed electronics on the DCR and afterpulsing effect; even though, the excess bias voltage in the MS pixel is more than three times lower than the DS pixel, its DCR and afterpulsing

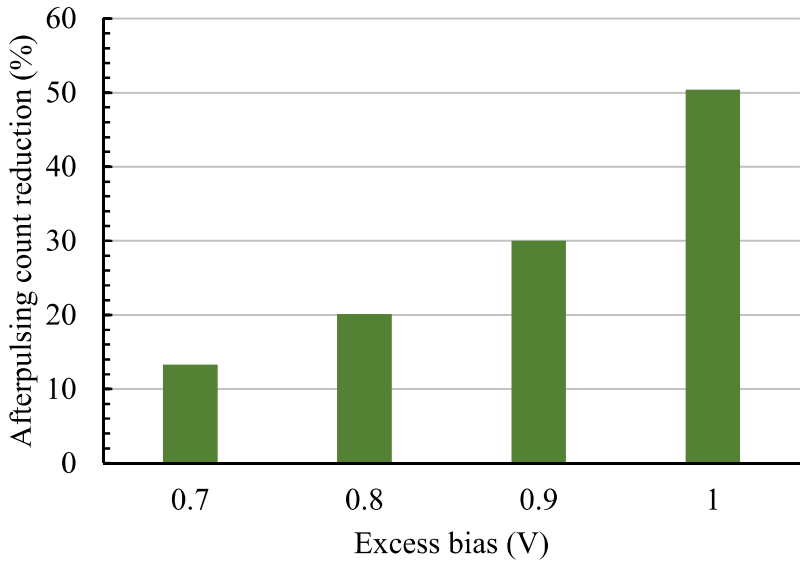


Fig. 3.19 Afterpulsing count reduction (relative difference between afterpulsing counts in passive quenching and active quenching modes) of Direct Sensing pixel with 8 ns dead time at 80 °C for different excess bias voltages.

is about 100 times and 70 times more than DS pixel, respectively.

Same measurements have been done in different temperatures with different V_{ex}. In Fig. 3.18, the normalized afterpulsing counts in PQ and AQ are shown for DS and MS pixels under 800 mV and 250 mV of V_{ex}, respectively. At every temperature both circuits are effective in the afterpulsing reduction. In the DS pixel in Fig. 3.18 (a), 85 % of AP reduction at -20 °C is seen. In the MS pixel (Fig. 3.18 (b)) the minimum afterpulsing reduction is 76 % while at most of the temperatures this value is more than 90 %. Fig. 3.19 shows the afterpulsing reduction versus V_{ex} for DS pixel at 80 °C. By increasing V_{ex}, the ticklish AQAR circuit reduces the afterpulsing effect more effectively. For 1 V of V_{ex}, more than 50 % afterpulsing reduction is achieved.

In Fig. 3.20, a comparison between the DCR values of DS and MS pixels is performed at different temperatures and excess bias voltages. Even though the excess bias voltage of the MS pixel is less than the DS pixel, its DCR is about one hundred times higher than the DS pixel. This is also observable through the DCR fitted curves in the histograms of Fig. 3.17. These results further confirm the statement that placing electronics over the SPAD, increases the DCR as well as the afterpulsing effect. While the dominant part of the noise (afterpulsing effect), is reduced drastically (up to 94 %) by the proposed ticklish AQAR circuits. Also, with some modifications in

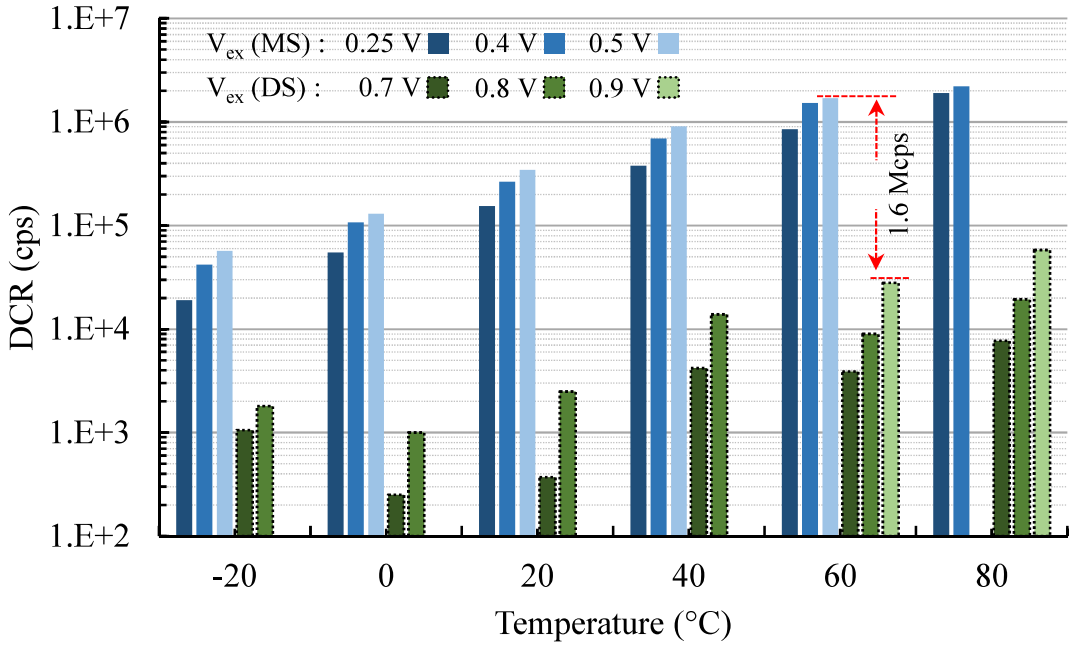


Fig. 3.20 Dark Count Rates (DCR) of SPAD in Direct Sensing and Mixed Sensing pixels with 8 ns dead time in different temperatures and excess bias voltages.

the layout of the SPAD, the DCR can be reduced. The DCR analysis and the modifications to improve it, are discussed more in [154]–[157].

Chapter 4
Pre-Biased Inverter and Body Voltage Independent Biasing Circuit

4.1. Introduction

In the previous chapter it is discussed that biasing an inverter at its switching threshold results in a very high sensitivity as in the ticklish detection circuit. At this biasing point, the input voltage is equal to the output voltage of the inverter, both NMOS and PMOS transistors are conducting in the saturation region. Therefore, there is a path between V_{dd} and gnd that allows flow of a DC current. In this operating region, the power dissipation of the MOSFETs is maximal. Thus, using the ticklish detector would result in a huge power consumption in a dense SPAD array which prevents the realization of megapixel SPAD arrays. At the pixel level also, in the inherent 3D structure, this considerable quiescent current consumption can heat the SPAD and increases the DCR.

In this chapter, a new avalanche detection circuit is presented that is not notably less sensitive than the ticklish detector, however, its power dissipation is drastically less than the ticklish inverter. Based on this detection circuit, an AQAR circuit that detects the avalanche at its early stages and significantly reduces afterpulsing effect is presented. In the following, the AQAR circuit is developed to be compatible with the inherently 3D structure and exploits its distinguishing feature in the fill factor improvement.

4.2. Pre-biased Inverter for Avalanche Detection

4.2.1. *Pre-biasing Idea*

In the digital operation, the static power dissipation of an inverter is almost zero (neglecting the leakage current of the transistors). However, this operation mode is not suitable for the low voltage signal detection as it is required in the avalanche detection for afterpulsing reduction purposes. To achieve the maximum sensitivity to the input voltage variation in an inverter, the

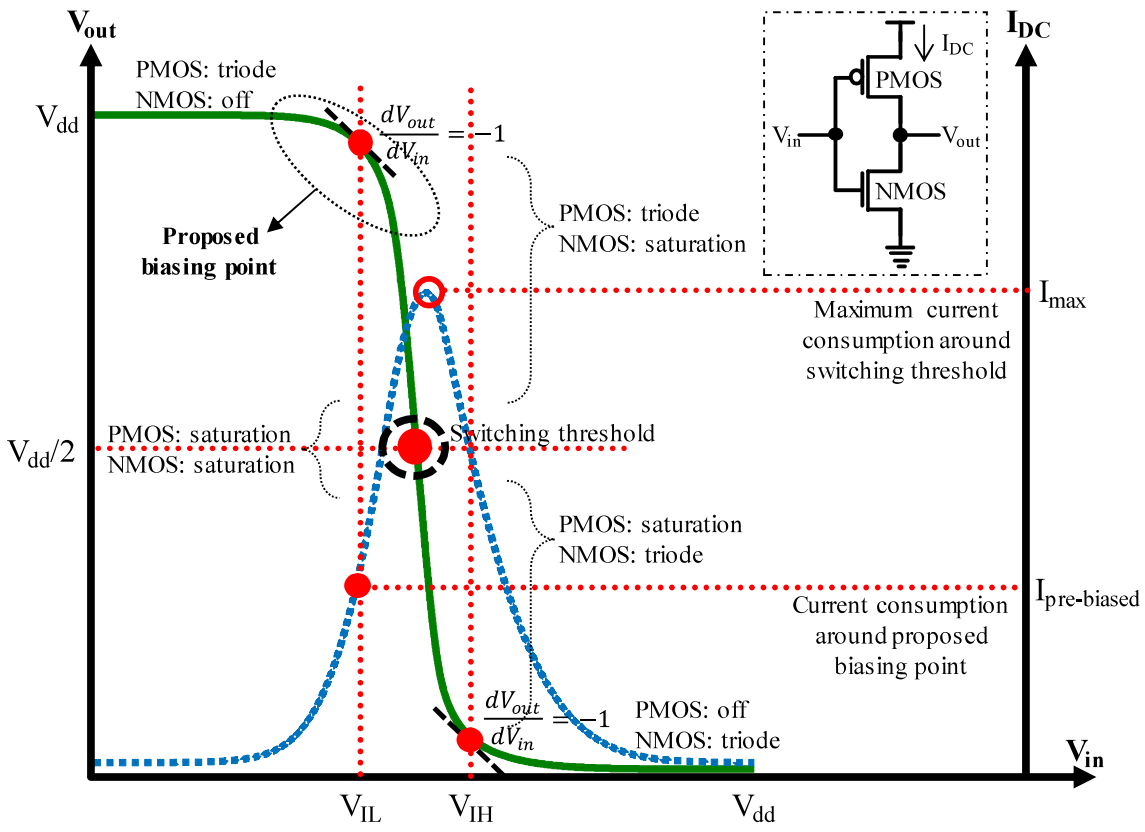


Fig. 4.1 Voltage Transfer Characteristic (VTC) – green curve - and Current Transfer Characteristic (CTC) – blue dotted curve - of a typical inverter. Operating points of the transistors is shown for each part of the VTC. Current consumption for the proposed biasing point is marked on the CTC.

ticklish inverter is designed and biased at the middle point of the VTC of the inverter by mean of a T-gate. In this design, the switching delay is minimum, and the sensitivity is maximum at the cost of the maximum static power consumption.

Fig 4.1 shows the VTC and Current Transfer Characteristic (CTC) of a typical inverter. The linear part of the VTC is narrow in comparison with the nonlinear parts. This part is limited by the points where $dV_{out}/dV_{in} = -1$, named as input low voltage V_{IL} and input high voltage V_{IH}. V_{IL} is the maximum V_{in} which V_{out} is still considered a logic high and V_{IH} is the minimum V_{in} which V_{out} is still considered a logic low. At V_{IL}, NMOS is at the edge of the cut-off and the saturation region, therefore, consumes a low current while PMOS is in the triode region. Thus, on the CTC, it can be seen that, at V_{IL}, the static current passing through the NMOS and PMOS is low as well. By increasing V_{in}, NMOS goes further into the saturation region and so the static current is increased until the switching threshold V_S where both NMOS and PMOS are in saturation region.

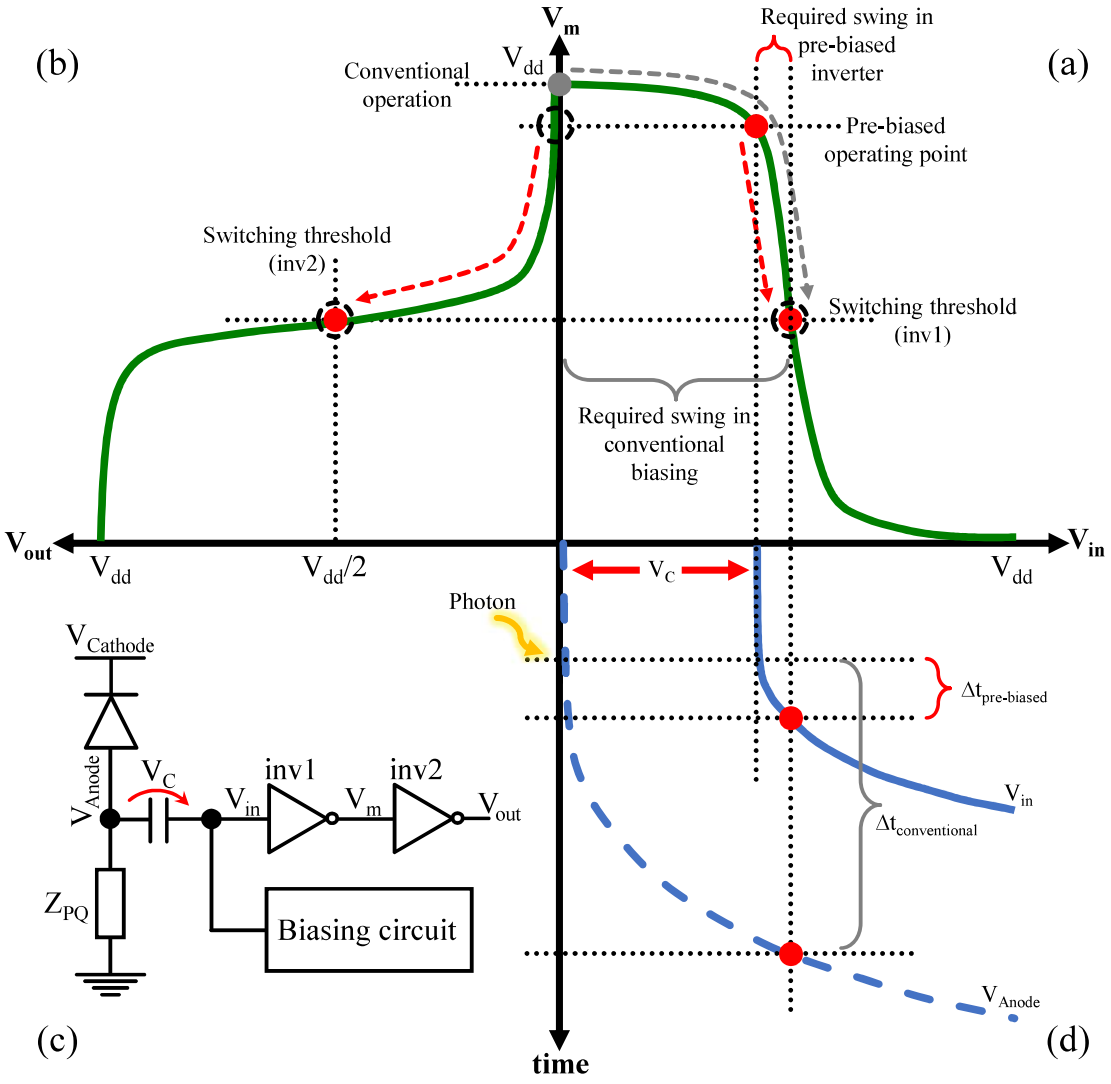


Fig. 4.2 Illustration of the pre-biased detector. (a) VTC of a typical inverter biased around the V_{IL} (b) VTC of the following inverter (c) Symbolic schematic of the proposed pre-biased detector (d) Anode voltage (V_{Anode}) and input voltage (V_{in}) of the pre-biased inverter versus time during an avalanche.

Thus, the static current and power consumption are the maximum. By a proper design, V_{IL} and V_S can be very close.

Based on these observations, it is proposed to pre-bias the inverter around V_{IL} or just below it. Under this biasing condition, the current consumption is much less than the ticklish inverter while the sensitivity is still high enough for the fast avalanche detection. Fig. 4.2 illustrates the avalanche detection based on this strategy. In Fig. 4.2 (c), a symbolic schematic of the proposed pre-biased detection circuit is shown. The first inverter (inv1) is biased around the V_{IL} by the pre-biasing

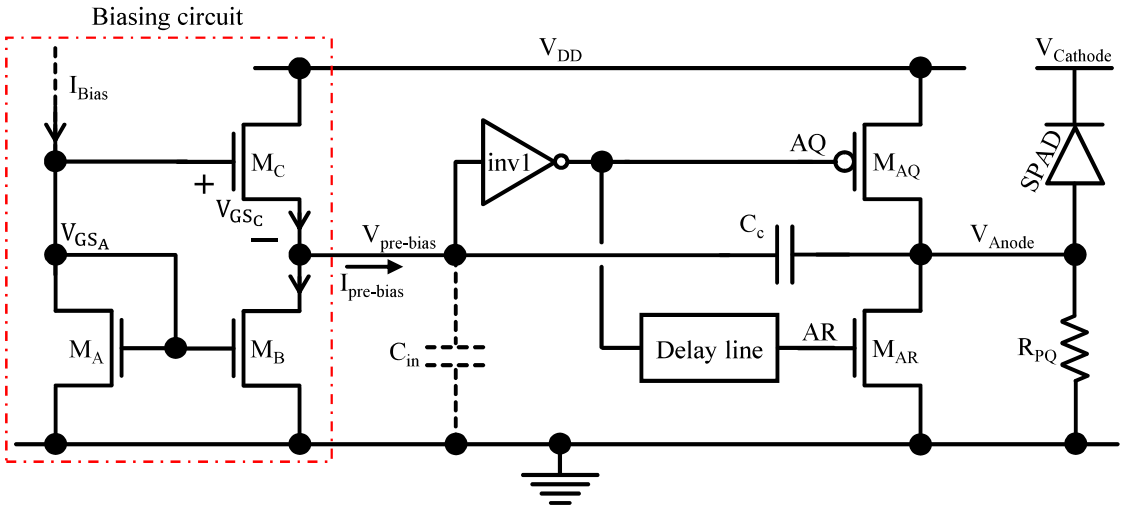


Fig. 4.3 Schematic of the pre-biased avalanche detection and quenching circuit. The biasing circuit is inscribed in the red rectangle. The dashed capacitor is representing the total capacitance seen at the input of inv1 (parasitic and input capacitances).

circuit. The second inverter (inv2) generates a digital signal for each avalanche. A coupling capacitor is mandatory to separate the DC voltage of the anode and the input of inv1. With the occurrence of an avalanche, the anode voltage starts to increase. As shown on the VTCs of inv1 and inv2 in Fig. 4.2 (a) and (b), respectively, just a few millivolts of this voltage increase are required for the pre-biased detection circuit to detect the avalanche. This small voltage swing is translated to a short detection time as demonstrated in Fig. 4.2 (d).

4.2.2. *Biasing Circuit*

The schematic of the AQAR circuit, based on the proposed biasing circuit, is presented in Fig. 4.3. The detection inverter is biased around V_{IL} thanks to M_A, M_B and M_C. M_A and M_B form a current mirror biased by a constant current source I_{Bias} of about 600 nA. These two transistors are chosen long and narrow ($W/L \ll 1$) to generate a relatively high voltage $V_{GS\,(A)}$ on the gate of M_C. There are biased in strong inversion regime and $V_{GS\,(A)}$ can thus be evaluated by using the conventional SPICE level 1 quadratic model:

$$V_{GS_A} = V_{th_{long}} + \sqrt{\frac{2 \cdot I_{Bias}}{K_N} \frac{L_A}{W_A}} \tag{4.1}$$

where $V_{th\,(long)} \approx 410$ mV is the threshold voltage of long NMOS transistors, $K_N = 540\ \mu A/V^2$ is the transconductance coefficient of NMOS, and $W_A = 100$ nm and $L_A = 1500$ nm are the width and

the length of M_A. Since M_B has the same dimensions as M_A, the current drained by this transistor is equal to I_{Bias}. The more, using long transistors ensures that M_B behaves as a quasi-perfect current source with a very high output impedance. The gate of M_C is thus maintained to the constant voltage V_{GS_A} while M_B forces the current that flows through it to be equal to I_{Bias}. However, contrary to M_A and M_B, this transistor is chosen wide and short ($W/L \gg 1$). M_C is thus biased in the week inversion and its gate-source voltage V_{GS_C} can thus be evaluated by using the EKV model:

$$V_{GS_C} = V_{th_{short}} + u_T \cdot \ln\left(\frac{I_{Bias}}{2 \cdot K_N \cdot u_T^2}\frac{L_C}{W_C}\right) \tag{4.2}$$

where $V_{th\ (short)} \approx 410$ mV is the threshold voltage of short NMOS transistors, $u_T = k \cdot T/q$ is the thermal voltage (~ 27 mV at 40 °C), and $W_C = 200$ nm and $L_C = 30$ nm are the width and the length of M_C. Finally, by combining equations (4.1) and (4.2), one can find the quiescent input voltage V_{pre_bias} of the detection inverter:

$$V_{pre-bias} = V_{GS_A} - V_{GS_C}$$
$$V_{pre-bias} = \Delta V_{th} + \sqrt{\frac{2 \cdot I_{Bias}}{K_N}\frac{L_A}{W_A}} - u_T \cdot \ln\left(\frac{I_{Bias}}{2 \cdot K_N \cdot u_T^2}\frac{L_C}{W_C}\right) \tag{4.3}$$

where $\Delta V_{th} = V_{th\ (long)} - V_{th\ (short)} = 90$ mV is the difference between threshold voltages of long and short NMOS transistors in this technology. Finally, the biasing circuit allows to set the quiescent voltage at the input of the detection inverter at V_{IL} around 310 mV, below its switching threshold voltage (410 mV).

Since the biasing current of M_C is very low, its output impedance is relatively high and doesn't prevent any voltage change. Thereby, when an avalanche occurs, the anode of the SPAD rises and a fraction of this rise is transferred to the input of the detection inverter (inv1) thanks to C_c:

$$\Delta V_{pre-bias} = \frac{C_c}{C_c + C_{in}} \cdot \Delta V_{Anode} \tag{4.4}$$

where C_{in} is the effective capacitance at the input of the detection inverter (inv1), $V_{pre\text{-}bias}$ is its input voltage, and V_{Anode} is the SPAD anode voltage. The gate-source voltage of M_C is decreased, and will even become negative, turning this transistor off. This can happen quickly since the M_C is

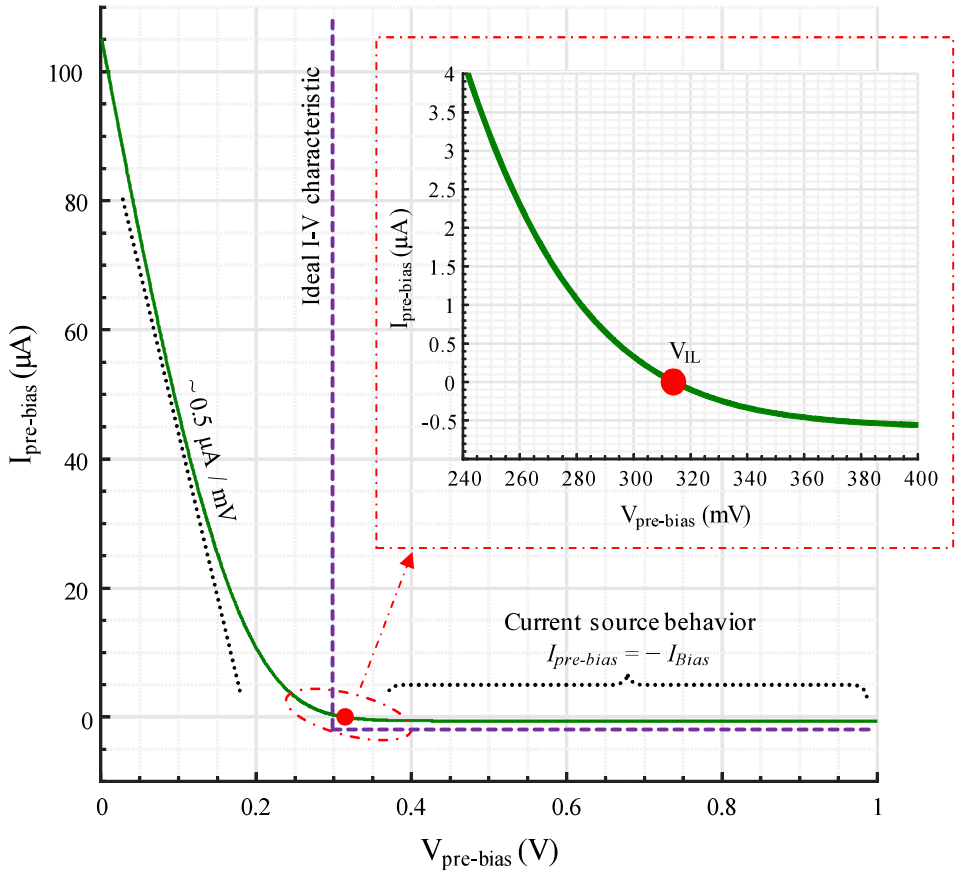


Fig. 4.4 Simulated output current of the biasing circuit as a function of its output voltage. When the voltage is over the desired biasing point, the circuit behaves as a constant current source of −600 nA to pull down the input voltage of the detection inverter after an avalanche event. On the contrary, when the voltage goes below the desired biasing point, the current increases dramatically, preventing the voltage to go too low during the reset phase.

biased in the weak inversion. Thanks to the pre-biasing, only a few tens of millivolts of voltage rise is needed at the input of the detection inverter to drive the AQ transistor M_{AQ} and a standard inverter to signal the avalanche detection. In the quenching phase, the biasing circuit behaves as a constant current source sinking a current equal to I_{Bias}. This can be seen in Fig. 4.4, which shows the I-V characteristic of the biasing circuit. This constant current will thus discharge the input capacitance of the detection inverter (C_{in}) to pull down the input voltage of the inverter just below its threshold voltage to V_{IL}. By starting the reset phase, which is controlled by the delay line, the anode voltage starts to fall and so the input voltage of the detection inverter. The input voltage of the detection inverter goes below the switching threshold and

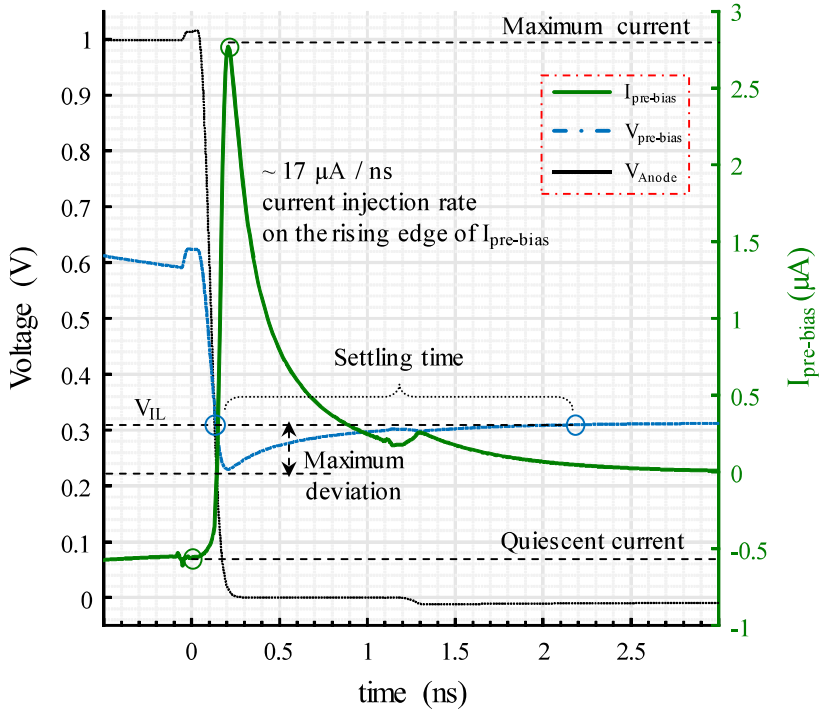


Fig. 4.5 Output current and voltage of the biasing circuit at the falling edge of the anode voltage during the reset phase. When $V_{pre\text{-}bias}$ starts to go below the biasing point, the biasing circuit inject a fast-rising current of 17 µA/ns that results in a maximum voltage deviation of less than 80 mV with a settling time of less than 2 ns for the biasing point.

consequently, the AQ transistor M_{AQ} is turned off. Now, the anode voltage is no longer maintained to V_{DD} and will decrease to the ground potential thanks to the AR transistor M_{AR} and PQ resistor R_{PQ}. A fraction of this voltage change is also transferred to the input of the inverter through the coupling capacitor C_c. However, this increases the gate-source voltage of M_C, moving it from the weak inversion to the strong inversion. Thus, its current increases dramatically, preventing the input voltage ($V_{pre\text{-}bias}$) to deviate from the desired quiescent voltage close to V_{IL}. Fig. 4.4 illustrates that below the desired biasing point, the current injection rate of the biasing circuit can reach to 0.5 µA/mV. This ensures a fast and stable biasing for the next coming avalanche and reduces the detection delay and jitter. Actually, it does not need to reach this current rate, since the C_{in} is fully recharged sooner. Fig. 4.5 shows the transient behavior of the biasing circuit on the falling edge of the anode voltage. Returning to its initial biasing point for the biasing circuit takes only about 2 ns while the maximum deviation from the biasing point is less than 80 mV. To achieve these, the biasing circuit injects the current which rises at the speed of 17 µA/ns and reach a maximum current

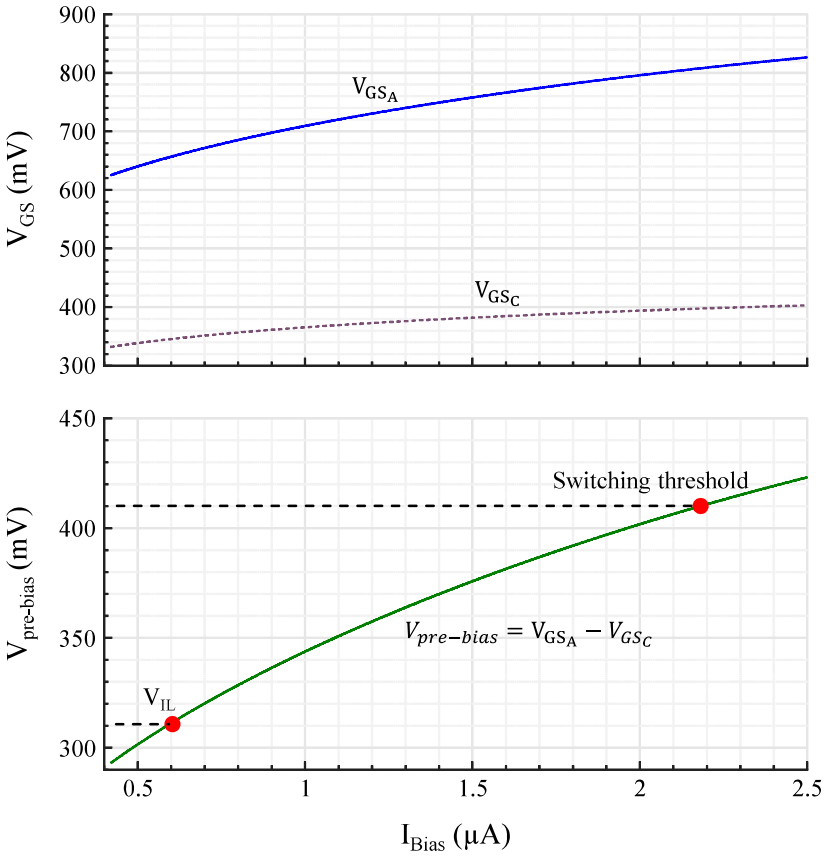


Fig. 4.6 V_{GS} variations of M_A and M_C due to the I_{Bias} change which allows tuning of the biasing point. For less than 2 µA of current variation, $V_{pre\text{-}bias}$ changes 100 mV. This feature can be used for increasing the sensitivity of the detection circuit or for robustness against the PVT variations.

of less than 3 µA. This is in agreement with the simulation results presented in Fig.4.4.

Another feature of the biasing circuit is the adjustability. By changing I_{Bias} it is possible to fine tune the pre-biasing voltage. Increasing I_{Bias}, increases both V_{GS_A} and V_{GS_C}, however the rate is higher in V_{GS_A}. Therefore, $V_{pre\text{-}bias}$ is rising and thus a more sensitive and faster detection circuit is achieved. This is illustrated in Fig. 4.6 where V_{GS_A}, V_{GS_C} and $V_{pre\text{-}bias}$ versus I_{Bias} are shown. It is possible to increase the sensitivity up to the one of the detection circuits in the previous chapter, but that will be at the cost of a higher static current consumption. This feature is also essential for robustness against the PVT variations. For instance, temperature variations modulate the threshold voltages of the MOSFETs, so both the switching threshold of the detection inverter and $V_{pre\text{-}bias}$ value change. Tuning I_{Bias} changes $V_{pre\text{-}bias}$ and compensates for these unwanted variations. For less than 2 µA of I_{Bias} variations, $V_{pre\text{-}bias}$ changes 100 mV.

4.3. Native 3D Pixel & Body Biasing Issue

As discussed in chapter 3, in order to achieve the native 3D pixel, the electronics must be robust under the unwilling body biasing induced by the pixel structure. Otherwise, the biasing point changes after each avalanche event, affecting the performance of the detection circuit or resulting in its malfunctioning. In the case of the ticklish detector, its self-biasing feature ensures a rugged and stable biasing point under any body voltage. Whatever the body voltage is, the t-gate forces the ticklish inverter to find its new equilibrium point. Now the compatibility of the proposed biasing circuit with the native 3D structure regarding the body voltage modulation should be investigated.

4.3.1. *Body Effect in FD-SOI Technology*

By placing electronics over the SPAD, body terminals of all the transistors and the anode voltage are combined. Thus, at each avalanche, when the anode voltage moves from 0 V to the excess bias voltage, the transistors undergo a compulsive body biasing. This body biasing modulates the threshold voltage of the transistor. In the FD-SOI technology the threshold voltage of NMOS and PMOS can be expressed as follows [166]:

$$\begin{cases} V_{th_N} = V_{th_{N_0}} - \dfrac{C_{si} \cdot C_{ox,b}}{C_{ox,g} \cdot \left(C_{si} + C_{ox,b}\right)} \cdot V_{body} \\ V_{th_P} = V_{th_{P_0}} - \dfrac{C_{si} \cdot C_{ox,b}}{C_{ox,g} \cdot \left(C_{si} + C_{ox,b}\right)} \cdot V_{body} \end{cases} \tag{4.5}$$

where $V_{th_{N_0}}$ and $V_{th_{P_0}}$ (negative value) are the threshold voltages of the NMOS and PMOS when their body voltages are zero, C_{si} is the silicon film or the channel capacitance, $C_{ox,b}$ is the box or the back gate capacitance, $C_{ox,g}$ is the gate capacitance and V_{body} is the body voltage. This equation shows a linear relationship between the threshold voltage and the body voltage. By taking the derivative of equation (4.5) with respect to V_{body}, the body effect coefficient or the body factor is extracted:

$$\frac{dV_{th}}{dV_{body}} = -\frac{C_{si} \cdot C_{ox,b}}{C_{ox,g} \cdot \left(C_{si} + C_{ox,b}\right)} \equiv -\alpha \tag{4.6}$$

The body factor α of the used technology in this work is about 85 mV/V and is almost identical for NMOS and PMOS.

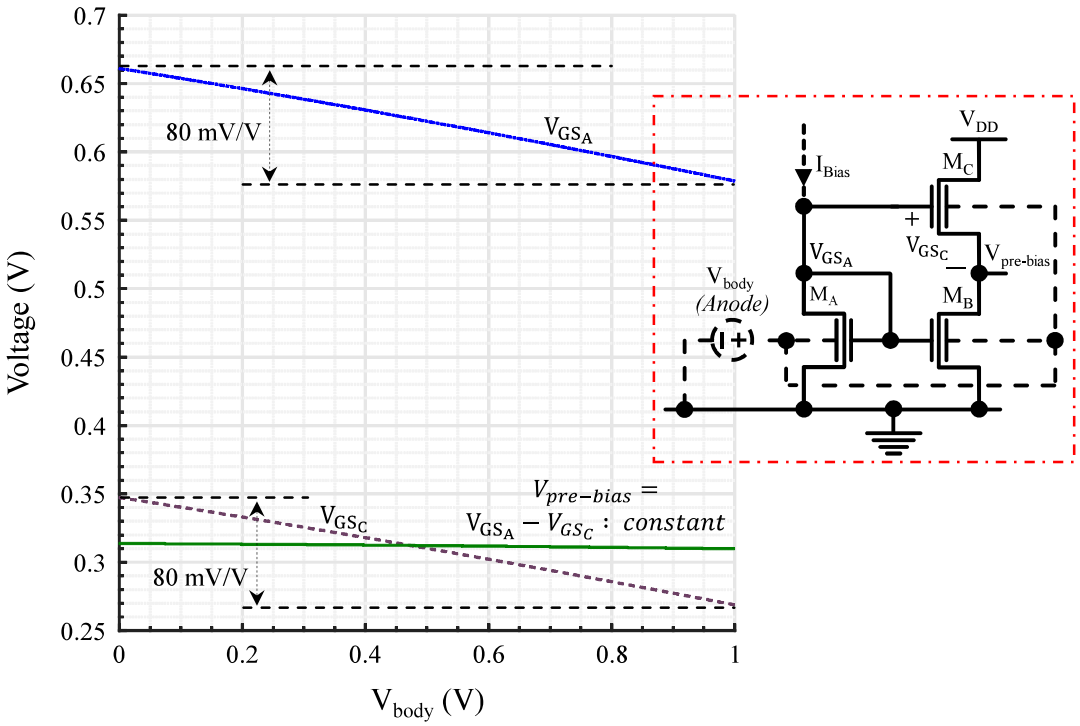


Fig. 4.7 Simulated gate-source voltage variations of M_A and M_C versus their body voltage. The identical linear relationship between the gate-source voltages and the body voltage of these two transistors makes it possible to obtain a bias voltage independent of the body voltage by subtracting the gate-source voltages of the two transistors. Inset: simulation setup.

4.3.2. *Body Independent Biasing Circuit*

In the native 3D pixel, based on the two last equations, the anode voltage rise decreases the threshold voltage of the NMOS while it increases the threshold voltage of the PMOS in absolute value. Note that the $V_{th_{P_0}}$ is a negative value. Nevertheless, in the proposed biasing circuit (Fig. 4.3), this does not modify the behavior of the current mirror made of M_A and M_B, since these two transistors are identical and share the same body voltage. This is further guaranteed by choosing a long length for them that reduces the mismatch. However, to maintain I_{Bias} in the current mirror, V_{GS_A} is decreased due to the threshold decrease of M_A. Yet this is not changing the output voltage $V_{pre\text{-}bias}$, since the threshold voltage of M_C and therefore V_{GS_C} are also decreased by the same amount as the V_{GS_A}. Note that, I_{Bias} is also passing through M_C and there is a linear relationship between the body voltage and the threshold voltage, and also between the threshold voltage and the gate-source voltage. Fig. 4.7 shows the V_{GS_A} and V_{GS_C} variations under the body biasing and

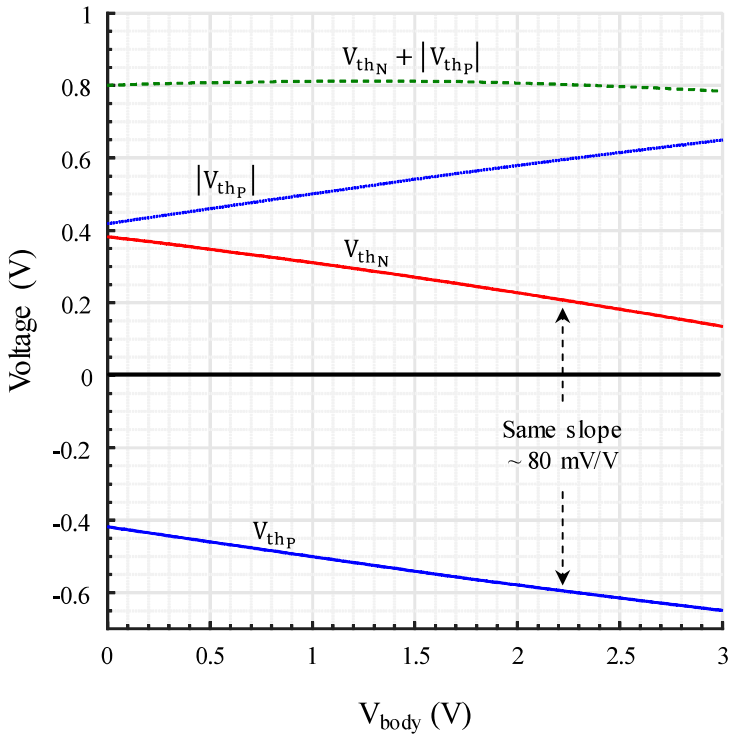


Fig. 4.8 Linear dependence of the threshold voltage on the body voltage with a body factor of 80 mV/V for both NMOS and PMOS in the FD-SOI technology. Because of this identical linear relationship, the NMOS threshold voltage variations compensate for the absolute value of the PMOS threshold voltage variations when added together.

their subtraction which in the proposed biasing circuit gives a voltage independent of the body biasing.

4.3.3. *Body Independent Current Source*

Yet, there is a body voltage dependent element in the equation (4.3). The proposed circuit needs to be biased by a constant current. However, since the body voltage is not constant, the threshold voltages of the transistors change, making the design of a constant current source complicated. Substituting the equation (4.6) in the equation (4.5), one can notice that the threshold voltage changes of both NMOS and PMOS are identical:

$$\begin{cases} V_{th_N} = V_{th_{N_0}} - \alpha \cdot V_{body} \\ V_{th_P} = V_{th_{P_0}} - \alpha \cdot V_{body} \end{cases} \tag{4.7}$$

Since $V_{th_{N_0}}$ is positive while $V_{th_{P_0}}$ is negative, their absolute values vary linearly in opposite directions as depicted in Fig. 4.8. This observation allowed us to design a circuit delivering a

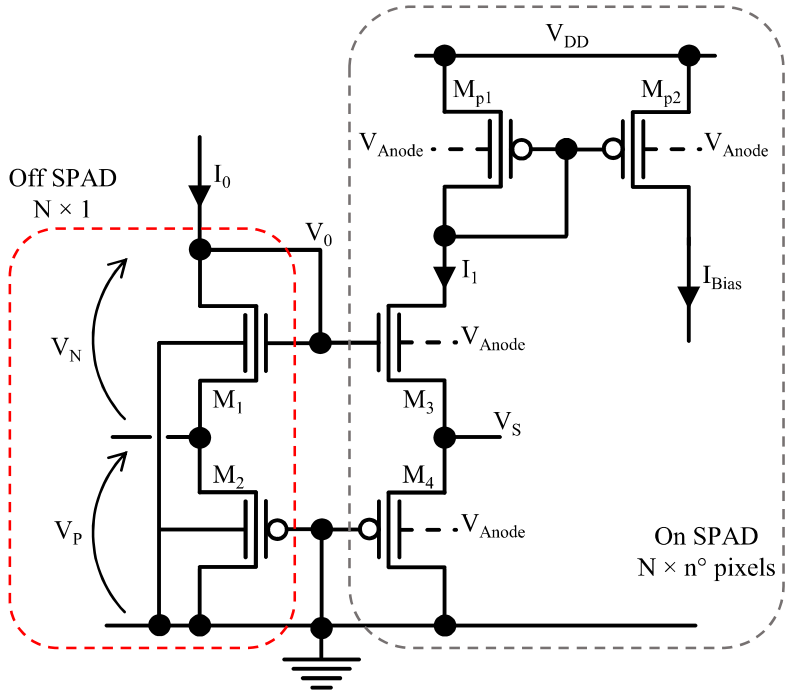


Fig. 4.9 Schematic of the body voltage independent current source. The transistors inside the red rectangle are implemented once outside the active area of the SPAD. V_0 is shared with all the pixels of the array and transistors in the gray rectangle are placed over the SPAD of each pixel.

current whose intensity does not depend on the body voltage of its transistors. This circuit is presented in Fig. 4.9. It consists of an NMOS and a PMOS current mirrors (M_1 to M_4) stacked together. The PMOS current mirror is biased in weak inversion to minimize the voltage drops. The left side of the current mirrors (M_1 and M_2) is intended to be placed outside the pixel to generate a single constant voltage V_0, shared with all the current sources used to bias each pixel in an array. Note that each current source must feed one pixel in the SPAD array while respecting the pre-biasing concept regarding the low power consumption. Since M_1 and M_2 are diode connected transistors, their gate-source voltages V_N and V_P only depend on their current I_0, which can be set by an external trimmer or an external current source:

$$\begin{cases} V_N = V_{th_N}(V_{body}) + \sqrt{\dfrac{2 \cdot I_0}{K_N}\dfrac{L_1}{W_1}} \\ V_P = \left|V_{th_P}(V_{body})\right| + u_T \cdot \ln\left(\dfrac{I_0}{2 \cdot K_N \cdot u_T^2}\dfrac{L_2}{W_2}\right) \end{cases} \tag{4.8}$$

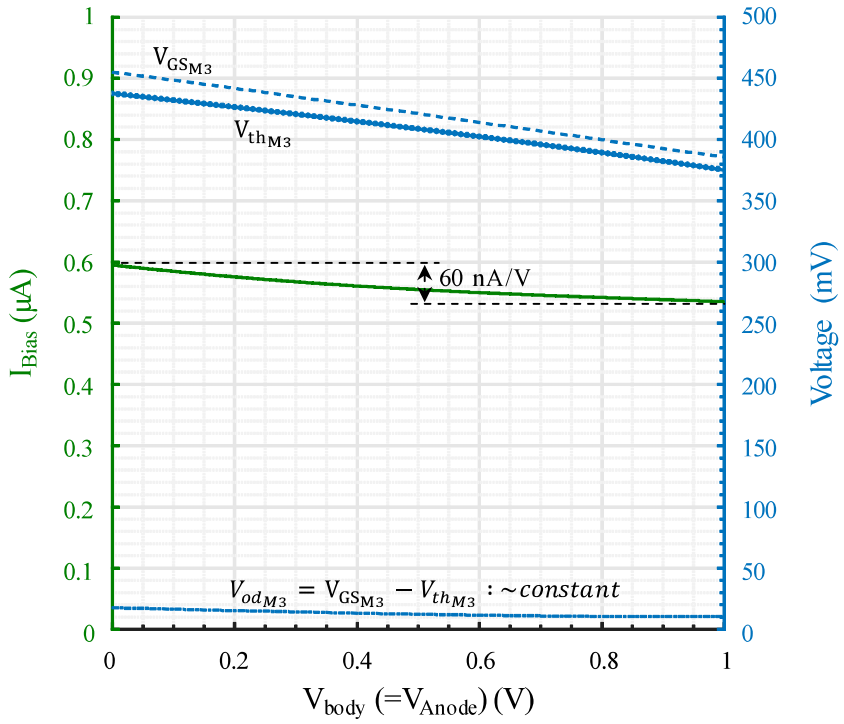


Fig. 4.10 I_{Bias} variations versus the body voltage, which in the native 3D pixel is equal to the anode voltage. As body biasing changes the threshold voltage of M_3, its gate-source voltage also changes in almost same amount and direction, resulting in a constant overdrive voltage and so a constant current.

Where W_x and L_x are the width and length of the M_x transistor. The voltage V_0 is thus the sum of V_P and V_N:

$$V_0 = V_N + V_P$$

$$V_0 = V_{th_N}(V_{body}) + \left|V_{th_P}(V_{body})\right| + \sqrt{\frac{2 \cdot I_0}{K_N}\frac{L_1}{W_1}} + u_T \cdot \ln\left(\frac{I_0}{2 \cdot K_N \cdot u_T^2}\frac{L_2}{W_2}\right) \tag{4.9}$$

It should be noted that even though the bodies of M_1 and M_2 are both connected to a different voltage than the anode, the voltage V_0 does not depend on their body voltage as long as they are equal. This is because V_0 depends on the addition of $V_{th_N}(V_{body})$ and $\left|V_{th_P}(V_{body})\right|$, canceling the body voltage dependent terms of each other as depicted in Fig. 4.8. This is also true for the right side of the current mirrors (M_3 and M_4) which are placed over the SPAD. When the anode voltage V_{Anode} rises, the absolute value of the threshold voltage of M_3 decreases and the one of M_4 increases. Their source voltage V_S are increased by the same amount to maintain the current copying function.

This keeps the overdrive voltages of M_3 and M_4 constant. Therefore, these transistors drain a constant current whatever the anode voltage (the body voltage) is. Finally, this body voltage independent current is fed to the body voltage independent biasing circuit through a PMOS current mirror (M_{P1}, M_{P2}). For a proper operation, the lengths of the two PMOS (M_2 and M_4) must be equal as well as the lengths of the two NMOS (M_1 and M_3), and it is possible to set the gain of the current mirror ($\beta = I_1/I_0$) by sizing the widths of M_3 and M_4 ($W_3 = \beta W_1$ and $W_4 = \beta W_2$). Fig. 4.10 shows the simulated behavior of the proposed body voltage independent current source under the body voltage variations.

4.4. Pre-Biased Active Quenching Active Reset Circuits

Based on the proposed pre-biased inverter as the detection circuit, three different AQAR circuits are introduced. As in chapter 3, they are categorized as the Direct Sensing (DS) pixel, the Mixed Sensing (MS) pixel and the Indirect Sensing (InDS) pixel. The full schematic of the pre-biased DS AQAR circuit is presented in Fig. 4.11. The DS pixel is a non-3D structure where the electronics are placed beside the SPAD and thus all the body terminals of the transistors can be connected to their nominal voltages. Though it is possible to use a conventional current source for biasing the pre-biasing circuit in this pixel, the body independent current source is used to have faire comparison between the pixels. Fig. 4.12 shows the timing diagram of the DS circuit operation. Thanks to the pre-biasing circuit, the detection inverter (inv1) is biased close to its switching threshold. By arrival of a photon, at the onset of the resultant avalanche t_0, the anode voltage slightly increases. This is enough for the pre-biased inverter to detect the avalanche (t_1) through the coupling capacitor C_c and drives the AQ transistor M_{AQ}. This wide PMOS injects a fast-rising high current to the anode until it reaches 1 V, which is equal to excess bias voltage of the SPAD. Thus, it is driven out of the breakdown region, the impact ionization process is extinguished and therefore the avalanche is quenched at t_2. After a certain hold-off time, set by the delay line, the active reset transistor M_{AR} discharges the anode voltage to zero volt (t_3 - t_4) and pulls the SPAD back to the breakdown region, where it waits for a new photon.

In the pre-biased MS AQAR circuit, the same circuit topology as the DS pixel is used over the SPAD. Placing the biasing circuit over the SPAD has already been discussed in section 3. The rest of the circuit behaves as explained in chapter 3. As the anode voltage rises, the threshold voltages

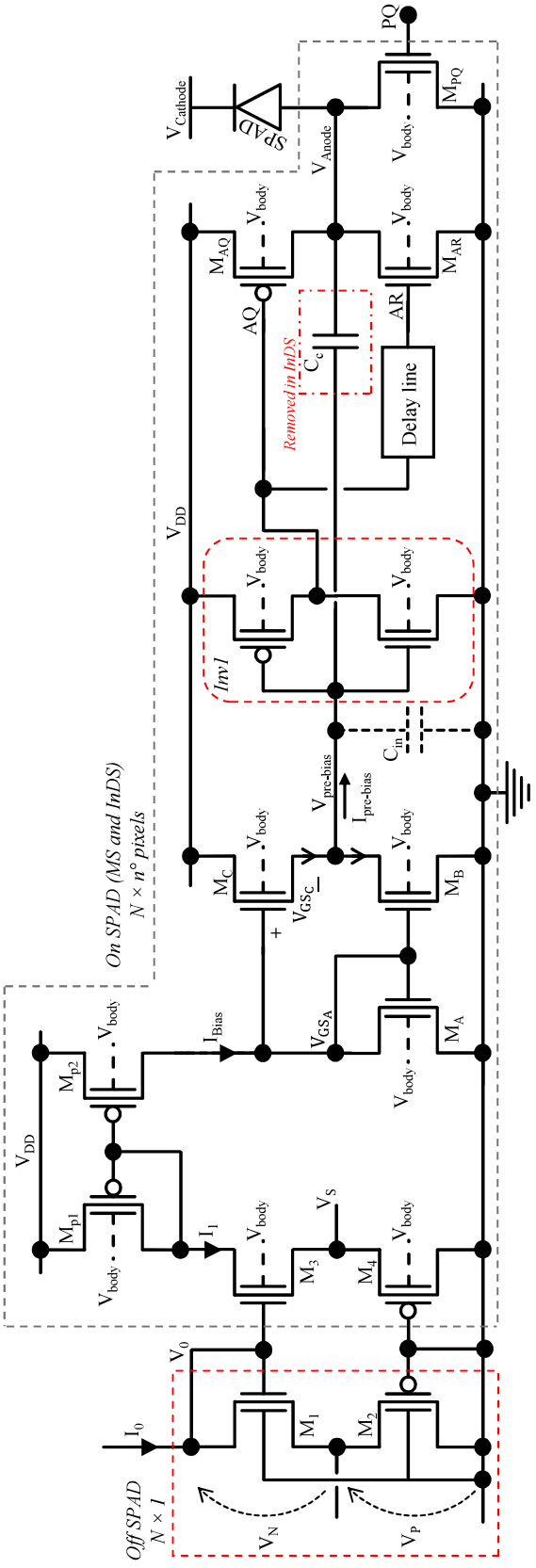


Fig. 4.11 Full schematic of the proposed pre-biased AQAR circuit. In the DS pixel all body voltages (V_{body}) are connected to ground. In the MS and the InDS pixels, all of them are equal to the anode voltage. For the InDS pixel, the coupling capacitor must be removed from the schematic. M_1 and M_2 are implemented once outside of SPAD and are used for all the pixels.

of

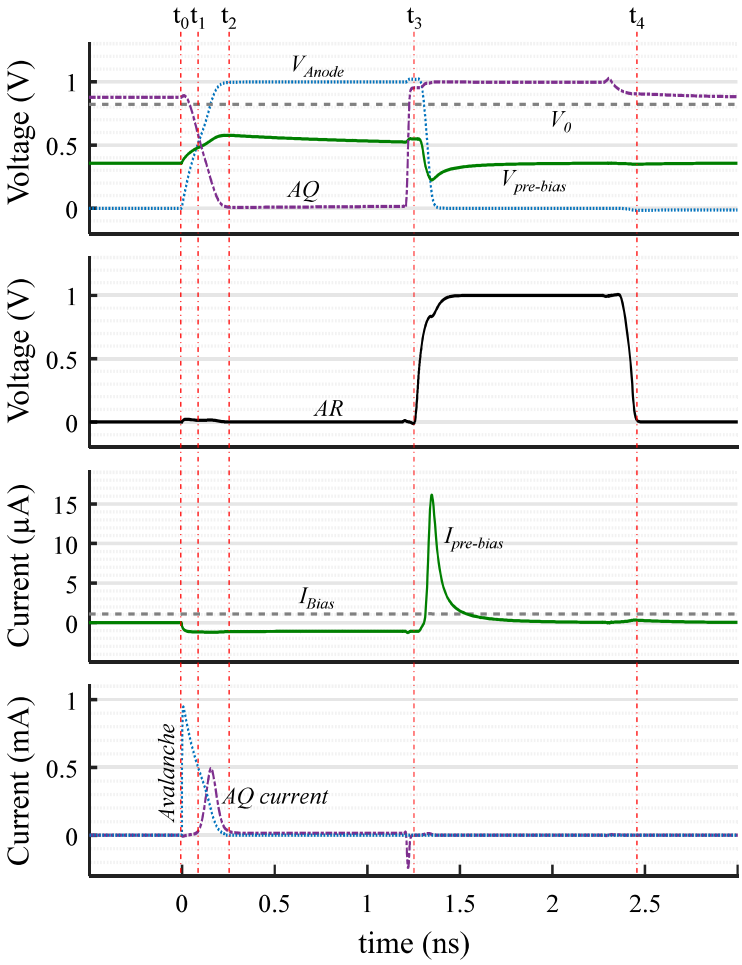


Fig. 4.12 Timing diagram of the circuit in Fig. 4.11. At t_0 avalanche starts. At t_1, it is detected by the detection inverter. Thanks to the active quenching current, it is quenched at t_2. At t_3, the active reset phase starts and finish at t_4, where the SPAD is ready for a new avalanche.

the NMOS and PMOS decreases and increases respectively in absolute value. Consequently, the switching threshold of the detection inverter inv1 is decreased as well, resulting in an increased sensitivity. This means that in a same biasing point, the MS pixel can detect the avalanche sooner than the DS pixel and thus be more efficient in the afterpulsing reduction.

Removing the coupling capacitor C_c from Fig. 4.11, brings in the pre-biased InDS pixel. Consider that in the last two pixels (MS and InDS) the body terminals are connected to the anode. In the InDS pixel, the avalanche is supposed to be sensed only through the bodies of the detection inverter. In the absence of C_c, the anode voltage variations are not sensed at the gates of the detection inverter. In fact, the gates are fixed at $V_{pre\text{-}biase}$, and this is only the VTC of the detection

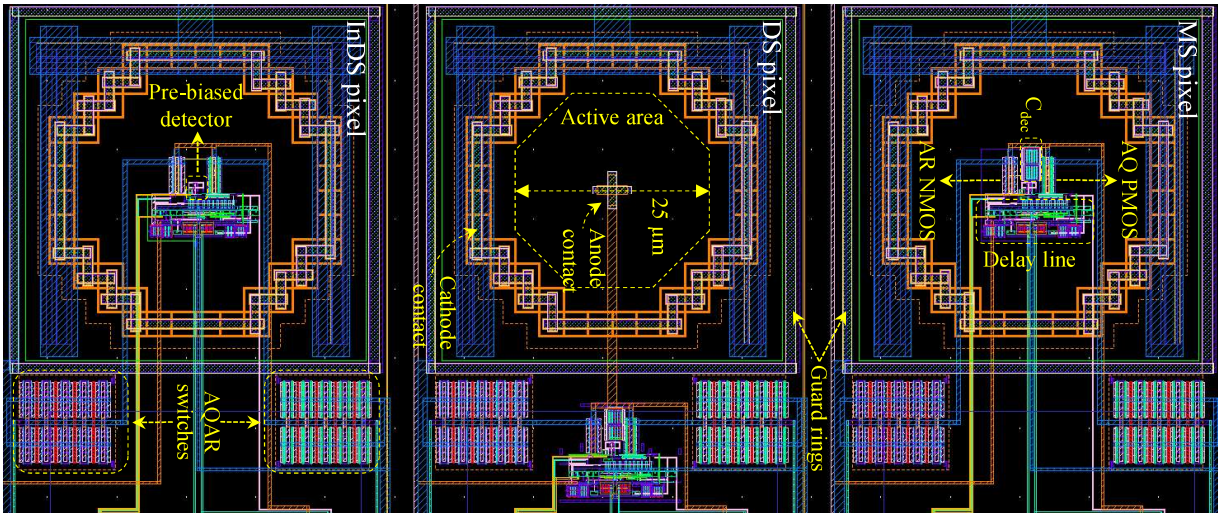


Fig. 4.13 Layouts of mixed sensing, direct sensing, and indirect sensing pixels. The AQAR switches are in series with AQ PMOS and AR NMOS for measurement purposes and are excluded from the pixel area.

inverter that moves. When the switching threshold passes the pre-biased point $V_{pre\text{-}bias}$ the avalanche is detected. It is obvious that the detection time is longer than the DS and MS pixels due to the less sensitivity. Also, the 85 mV/V body factor necessitates raising the pre-biased point a little higher than the V_{IL} to ensure that it is reachable by the switching threshold. Since the InDS pixel allows applying a higher V_{ex} than the voltage headroom of the technology, in an application that does not need the AQ, V_{IL} still can be chosen as the biasing point.

4.5. Results and Discussion

4.5.1. *Post Layout Simulation Results*

Fig. 4.13 presents layouts of the three proposed AQAR circuits in the used 28 nm FD-SOI CMOS technology. The same quasi-octagonal SPAD with an active area of 490 μm^2 is used as in the chapter 3. Each circuit has an area of about 145 μm^2. This area is less in InDS pixel due to the absence of the coupling capacitor. Most of this area is filled with the delay line (5×16 μm^2) and the wide AQ and AR MOSFETs (5×5 μm^2). The detection circuit area is less than 3×3 μm^2, which makes it a compact circuit for the applications that only need the detection part. In the two native 3D pixels (MS and InDS), all the electronics are placed over the SPAD inside the active area. Yet there are places (70% of the active area) to fit more advanced electronics such as a counter or a compact time to digital converter. These layouts illustrate how, in a BSI mode, the fill factor can be maximized and be independent of the circuit. The area of the DS pixel is 29 % (145/490) of the

active area and the pixel can be used in the front side illumination mode. There are also two big switches outside of each pixel that enable us to switch between the PQ, AQ, PR and AR modes for measurements purposes and can be excluded from the final design.

Monte Carlo simulations are performed for the DS pixel at different biasing points between the V_{IL} and the switching threshold of the detection inverter V_{Sth}, considering mismatches and all the process variations to investigate the reliability of the circuit at different sensitivity levels for a SPAD array. The standard deviations are reported for $V_{pre\text{-}bias}$, V_{AQ}, V_{Anode}, active quenching current I_{AQ}, and I_{Bias}, which are the most critical voltages and currents of the circuit. Table 4.1 shows the Monte Carlo simulation results for all the biasing points. At 310 mV, 340 mV, and 360 mV (until 50 mV below the switching threshold), all the voltages and currents have adequate standard deviations and therefore, the circuit is robust against mismatches and process variations. These results ensure the functionality of the proposed circuits for each pixel of a SPAD array at the mentioned biasing points. However, at 390 mV the standard deviation of the V_{Anode} exceeds the tolerable value. Though, the mean value of 178.9 mV for V_{Anode} shows that there are only a few samples which are affected by the process variations.

According to the Monte Carlo simulation, M_3 and M_4 have the highest contribution in the mismatch, 13 % and 18 % (3 times higher than the contribution of the other MOSFETs), respectively. The length of M_4 is 40 nm (L_{min} = 30 nm), so a slight mismatch in the fabrication process can cause a threshold voltage difference between the samples. A threshold voltage variation at this transistor changes its current and so the I_{Bias} current, which can result in the malfunctioning of the circuit. Since these transistors are only generating a static current, it is proposed to increase their width and length (W and L). This would not change the behavior of the circuit while making it more robust against the mismatch and process variations. Therefore, it enables the circuit to reach a higher sensitivity in an array of SPADs.

The static power consumption of the DS AQAR circuit at each of the biasing points are also reported at table 4.1. This value can be as low as 5.8 µW. It means, for an array of 1000 SPAD, the quiescent consumption is 5.8 mW, while, in the ticklish AQAR circuit this is 43 mW, 43 µW per pixel, about 7.4 times lower consumption. As is shown in table 4.1, for a higher sensitivity more power is needed. However, at $V_{pre\text{-}bias}$ = 360 mV, the power consumption is still 3 times less than the ticklish circuit.

TABLE 4.1
MONTE CARLO SIMULATION RESULTS AND THE STATIC POWER CONSUMPTION FOR DIRECT SENSING PIXEL AT DIFFERENT BIASING POINTS.

	$V_{pre\text{-}bias}$ = 390 mV (V_{Sth} – 20 mV)	$V_{pre\text{-}bias}$ = 360 mV (V_{Sth} – 50 mV)	$V_{pre\text{-}bias}$ = 340 mV (V_{Sth} – 70 mV)	$V_{pre\text{-}bias}$ = 310 mV (V_{Sth} – 100 mV)
V_{Anode} (V): SD[1]	427.1 m	7.4 m	90.3 µ	20.4 µ
Min[2] *– Max*[3]	-15 m – 999.2 m	-241.7 µ – 69.5 m	-242.1 µ – 579.1 µ	-241.8 µ – -133.2 µ
Mean	178.9 m	828.5 µ	-162.4 µ	-184.7 µ
$V_{pre\text{-}bias}$ (V): SD	58.41 m	11.59 m	11.15 m	10.3 m
Min – Max	293.2 m – 615.1 m	333.1 m – 391.7 m	313.1 m – 368.9 m	287.9 m – 339.1 m
Mean	400.8 m	361 m	339.8 m	312.8 m
V_{AQ} (V): SD	286.7 m	60.4 m	28.8 m	12.8 m
Min – Max	3.7 m – 997.7 m	607.9 m – 993.5 m	818.5 m – 971.6 m	919 m – 986 m
Mean	659.2 m	871.4 m	928.8 m	966.1 m
I_{AQ} (A): SD	80.6 µ	924.9 n	2.7 n	126.1 p
Min – Max	822.2 p – 787.5 µ	278.1 p – 12.8 µ	135.3 p – 23.5 n	79.8 p – 904.8 p
Mean	9.5 µ	60.8 n	984.9 p	194.9 p
I_{Bias} (A): SD	149.7 n	123.8 n	102.7 n	75.4 n
Min – Max	1.1 µ – 2 µ	851.5 n – 1.5 µ	645.7 n – 1.2 µ	452 n – 857.2 n
Mean	1.6 µ	1.2 µ	934.9 n	633 n
SPC[4] *AQAR circuit (µW)*	22.8	14.5	10.3	5.8
SPC detection inverter (µW)	16.2	11	7.6	4.1

[1]: Standard Deviation
[2]: Minimum value
[3]: Maximum value
[4]: Static Power Consumption

To simulate the SPAD and the avalanche behavior, the same SPICE macro model as in the previous chapters is used, with 100 fF junction capacitance and 1 kΩ series resistance. The excess bias voltage is set to the voltage headroom, which is 1 V. Fig. 4.14 shows the transient simulation of the anode voltage, AQ and avalanche currents, and avalanche charge in the SCR at different sensitivity level ($V_{IL} < V_{pre\text{-}bias} < V_{Sth}$) for the DS pixel. Thanks to the AQAR switches, a comparison is made between the PQ and the AQ to evaluate the avalanche detection and quenching ability of the pre-biased AQAR circuit. According to the anode voltage curve Fig. 4.14 (a), the avalanche is detected in less than 140 ps and 80 ps for the least sensitive ($V_{pre\text{-}bias}$ = 310 mV) and

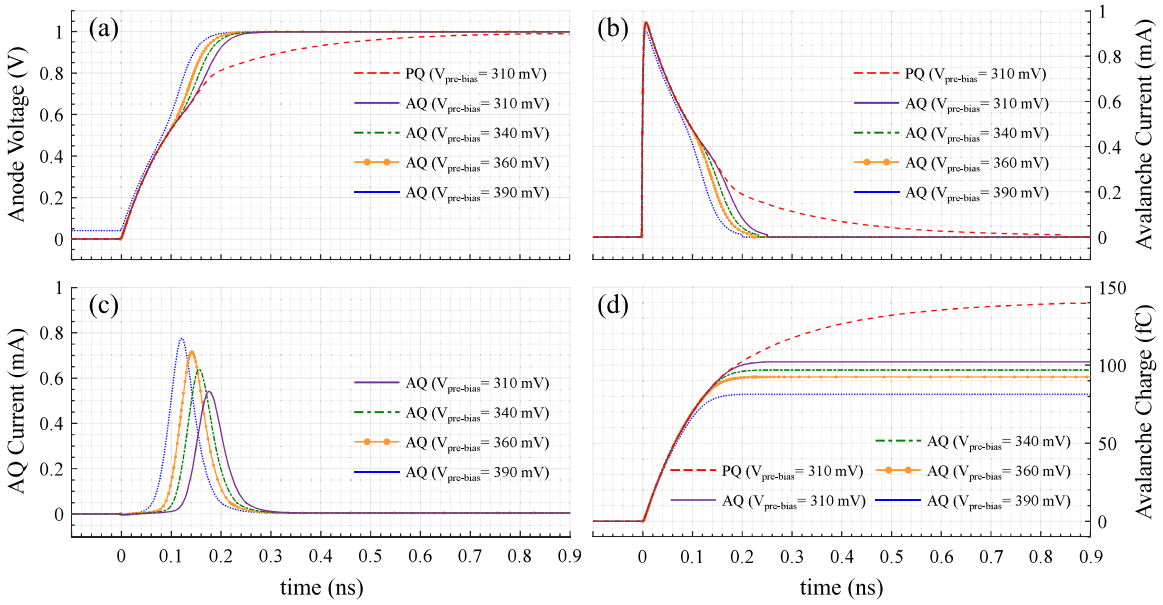


Fig. 4.14 Post layout transient simulation results during the occurrence of an avalanche in both active and passive quenching mode at different sensitivity level. (a) Anode voltage (b) Avalanche current (c) Active quenching current (d) Avalanche charge.

the most sensitive case ($V_{pre\text{-}bias}$ = 390 mV), respectively. The outcome of this detection is observable in Fig. 4.14 (b) where the avalanche current is shown. The sooner the avalanche is detected, the faster the quenching and the lower the avalanche current. Therefore, less avalanche charges flow in the SCR as shown in Fig. 4.14 (d). This charge reduction is due to the contribution of the AQ transistor in the quenching process, see Fig 4.14 (c). These results are shown in table 4.2. At $V_{pre\text{-}bias}$ = 360 mV, in AQ mode, the avalanche is quenched in 180 ps (50 µA of avalanche current) and the avalanche charge is reduced by 34 % in comparison with the PQ mode. In the ticklish circuit, these values are 166 ps and 45 %. Thus, the ticklish circuit is 14 ps faster than the pre-biased circuit in the quenching. The efficiency of the latter in the avalanche charge reduction is 11 % less than the ticklish while its power consumption is about 70 % less than the ticklish (3 times improvement).

The same simulations are performed for the three proposed pixels (DS, MS, and InDS). $V_{pre\text{-}bias}$ is set to 360 mV, which is the most efficient and robust biasing point according to the Monte Carlo simulations and table 4.2. The results are shown in Fig. 4.15 and table 4.3. As it was expected,

TABLE 4.2
POST LAYOUT SIMULATION RESULTS FOR DIRECT SENSING PIXEL AT DIFFERENT BIASING POINTS IN ACTIVE AND PASSIVE QUENCHING MODE.

	Quenching time (ps)	Avalanche charge (fC)	Avalanche charge reduction (%)
PQ ($V_{pre\text{-}bias}$ = 310 mV)	500	140	_
AQ ($V_{pre\text{-}bias}$ = 310 mV)	220	102	27
AQ ($V_{pre\text{-}bias}$ = 340 mV)	200	96	31
AQ ($V_{pre\text{-}bias}$ = 360 mV)	180	92	34
AQ ($V_{pre\text{-}bias}$ = 390 mV)	160	81	42

TABLE 4.3
POST LAYOUT SIMULATION RESULTS FOR DS, MS, AND INDS PIXELS AT $V_{PRE\text{-}BIAS}$ = 360 MV, IN ACTIVE AND PASSIVE QUENCHING MODES.

	Direct Sensing	Mixed Sensing	InDirect Sensing
Quenching time in PQ (ps)	500	558	500
Quenching time in AQ (ps)	180	164	394
Avalanche charge in PQ (fC)	142	165	158
Avalanche charge in AQ (fC)	92	84	142
AQ charge contribution (fC)	49	80	19
Avalanche charge reduction (%)	35	48	10

the fastest and the most efficient AQAR circuit is the MS pixel with a quenching time (50 µA of avalanche current) of about 164 ps and an AQ charge contribution of 80 fC that results in 48 % avalanche charge reduction. While at $V_{pre\text{-}bias}$ = 360 mV, only 50 mV is needed until the switching threshold of the detection inverter, the body biasing effect also reduces the switching threshold by 80 mV/V, and the fast-rising anode voltage that is coupled to the input of the detection inverter is increasing $V_{pre\text{-}bias}$. These three factors: pre-biasing, Body biasing and Gate biasing, make the MS pixel more efficient than the other two pixels, which each suffers from the absence of one of the mentioned factors.

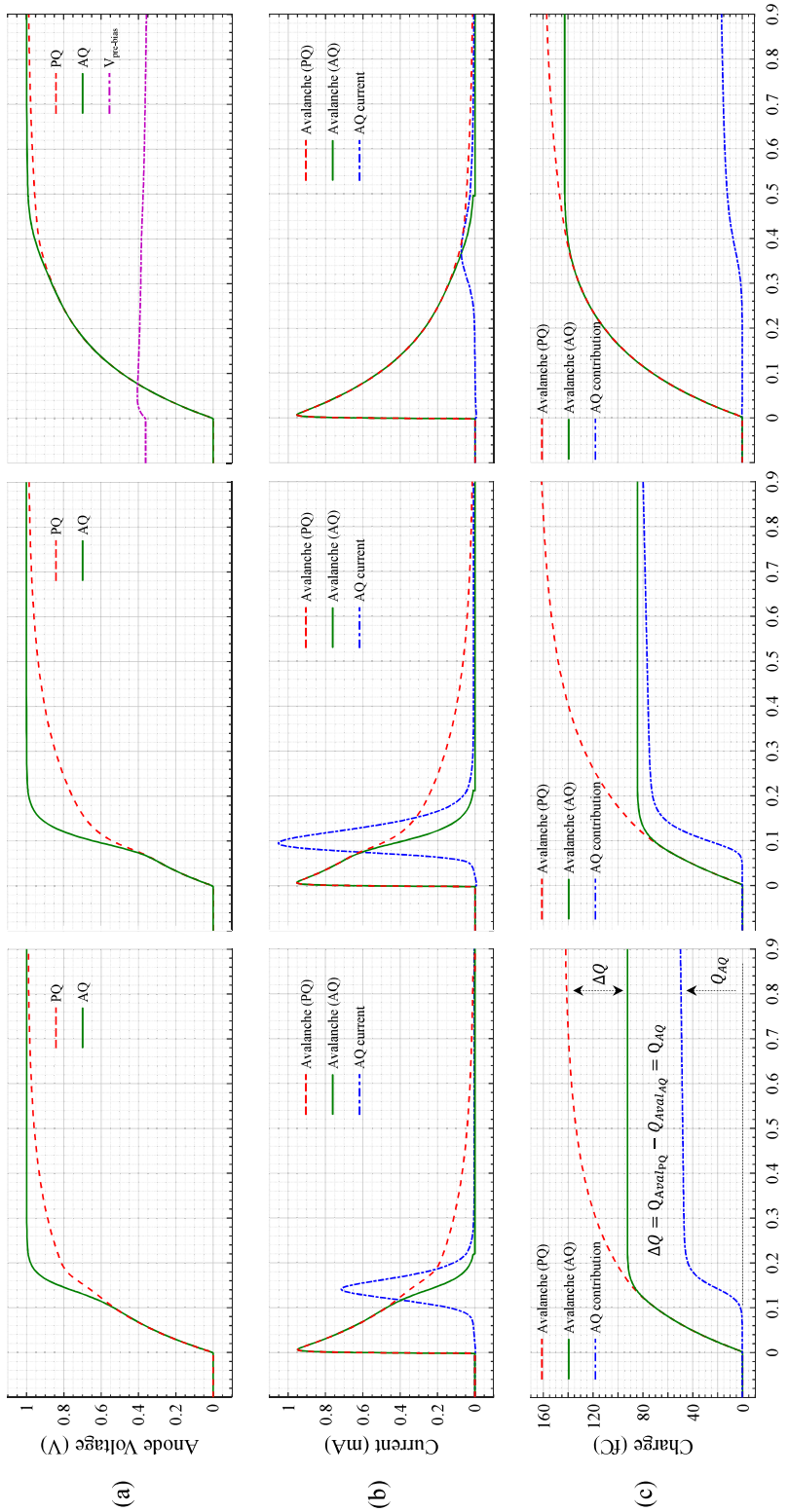


Fig. 4.15 Post layout transient simulation results during the occurrence of an avalanche in both active and passive quenching mode at $V_{pre\text{-}bias}$ = 360 mV for DS, MS, and InDS (from left to right, respectively). (a) Anode voltage (b) Avalanche and active quenching currents (c) Avalanche and active quenching charge.

In the InDS pixel, due to the absence of the Gate biasing, an anode voltage of 625 mV (50 mV/80 m) minimum is required to reduce the switching threshold of the detection inverter by 50 mV, to reach the pre-bias voltage. Nonetheless, the remaining 30 mV of switching threshold modulation is not enough to obtain a strong zero at the output of the detection inverter.

Thus, the AQ PMOS is not fully conducting. The late avalanche detection and the small AQ PMOS contribution make the InDS pixel the least efficient circuit. A higher biasing point increases the sensitivity of the InDS pixel on the cost of reliability and power consumption. In the DS pixel, even though the body effect does not exist, the two others are enough to quench the avalanche in 180 ps and reduce the charge by 35 %. It is notable that there is also a possibility of body biasing in DS pixel. Rather than the capacitively coupled body biasing, which is ubiquitously imposed in the inherently 3D pixels, here, a wire is required to connect the anode selectively to the body terminals of the detection inverter. This can make the DS pixel as efficient as the MS pixel in reducing the avalanche charge, without further investigation of body biasing effect on the rest of the circuit. Indeed, the MS pixel still has the advantage of maximized fill factor in the BSI mode.

There are a few remarks in Fig. 4.15 that are worth mentioning. In the anode voltage figures of Fig. 4.15 (a), the AQ curves deviate from their PQ counterparts at a certain point. The sooner this deviation occurs, or equivalently, the lower the anode voltage at the deviation point, the sooner the avalanche has been detected. In consequence, the AQ PMOS starts to conduct quicker, which allows this transistor to generate more current, since its drain-source voltage is higher at a lower anode voltage. This is observable in the avalanche and AQ PMOS current curves in Fig 4.15 (b). In the MS pixel, the deviation starts at the lowest anode voltage among the other pixels, thus, the AQ PMOS current is the highest in this pixel. It can be seen in Fig. 4.15 (c) that almost all the avalanche charge difference between the AQ and PQ modes (ΔQ) is provided by this current. The higher the AQ current, the shorter the quenching time and the lower the avalanche charge.

In the PQ mode the avalanche charge is different for each pixel, because of the same reason as explained in the previous chapter. In Fig. 4.15 (c), the PQ charge of the InDS pixel is slightly lower than of the MS pixel due to absence of C_c. As in chapter 3, to be fair, the comparison of avalanche charge reduction is conducted again, in regard with the PQ avalanche charge in DS pixel. Therefore, the avalanche charge reduction in the MS pixel is 40 % which still is higher than the DS

pixel. In the InDS pixel this value is around zero which is expectable due to its late detection and long quenching time.

By taking a closer look at Fig. 4.15, one can see a slope change in both anode voltage and avalanche current of DS and MS pixels in PQ mode. This change in the gradient of the curves, which is similar to the AQ effect, is due to C_{gd} (and other parasitic capacitances) of the ultra-wide switch between AQ PMOS and V_{dd}. As mentioned before, for the measurement purposes, a very wide switch (10 times wider than the AQ PMOS) is placed in series between V_{dd} and the AQ PMOS. Although, in the PQ mode, this switch cut the path between V_{dd} and AQ PMOS, however, C_{gd} of the switch still exist between V_{dd} and AQ PMOS. Also, the AQ signal is applied to the gate of the AQ PMOS, even in the PQ mode. Thus, at the detection of an avalanche, its gate falls quickly and allows C_{gd} to be discharged through AQ PMOS. This AQ-like effect slightly increases the anode voltage and consequently, reduces the avalanche current and charge, in PQ mode. When the quenching phase is finished, the AQ signal rises quickly and injects charge to C_{gd}. Therefore, at each new avalanche this loop of charging and discharging C_{gd} continues. Due to this effect, in these simulations, the avalanche charge reduction efficiency of the proposed AQAR circuits is slightly underestimated since the PQ is also assisted by some charge injection.

4.5.2. *Experimental Results*

In this section the experimental results of the proposed AQAR pixels are presented. The measurements setup is similar to the previous chapter, as these two groups of pixels are fabricated in the same chip, and all the measurements are performed in dark. The used SPAD in these two last chapters also is the same quasi-octagonal SPAD. Fig. 4.16 shows the micrograph of DS, MS, and InDS pixels, realized in a 28 nm FD-SOI CMOS technology. One can see the circuits placed on top of the SPAD in the two monolithic 3D pixels (MS and InDS), while in the DS pixel, the AQAR circuit is placed beside the SPAD, even though it is not visible due to the metal filling in the used technology.

In order to set and tune the biasing point ($V_{pre\text{-}bias}$) and the sensitivity level of the AQAR circuits, a digital potentiometer is integrated on the mother board in series with the body independent current source to generate I_0 (see Fig. 4.9 and Fig. 4.11). A higher resistor value generates less current (I_0) and thus lowering $V_{pre\text{-}bias}$ and decreasing the circuit sensitivity, while a lower resistor value increases the sensitivity (see Fig. 4.6). This tunability feature is also used for compensating the

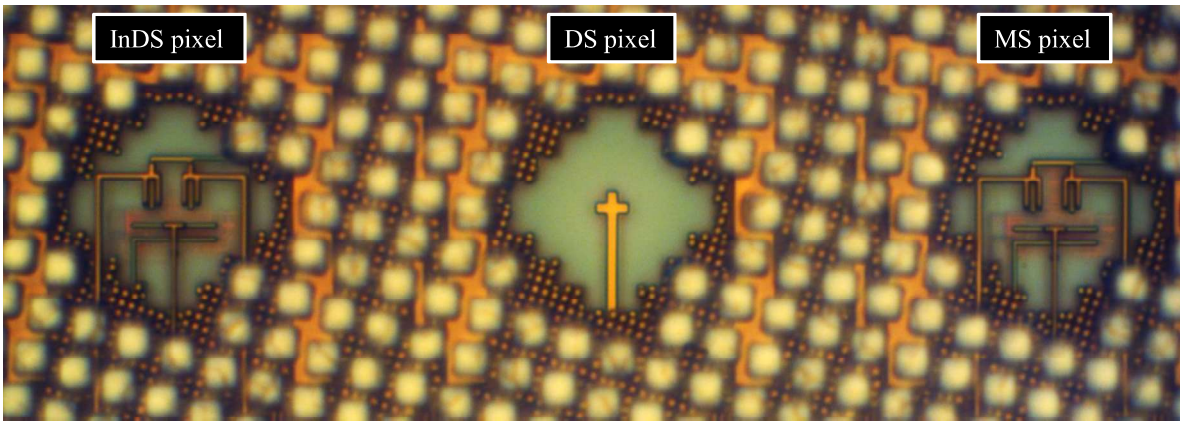


Fig. 4.16 Micrograph of the pixels. From left to right: Indirect Sensing, Direct Sensing, and Mixed Sensing pixels.

threshold voltage dependence of MOSFETs on the temperature and insures it works at every temperature.

4.5.2.1. SPAD voltage transient behavior

The anode voltage of the SPAD is accessible through the wideband analog buffer. Fig. 4.17 presents the anode voltage variations of the DS pixel during an avalanche at 60 °C. The excess bias voltage is around 700 mV, and the digital potentiometer value ($R_{sensitivity}$) is 533 kΩ. Fig 4.17 (a) shows the anode voltage variation in a pure passive quenching mode of operation. The oscilloscope is in persistence mode to observe the different avalanche amplitudes and their occurrence probability. Similar to what is shown in chapter 3, multiple avalanches are observed here: a main avalanche at V_{ex} = 700 mV, another avalanche with around 400 mV of amplitude and a statistical frequency a bit smaller than the first one. A less frequent avalanche at around 100 mV is observable. These multiple avalanches are results of the PEB effect due to the STI on top and around the SPAD and multiple sharp edges in the SPAD structure as it can be seen in Fig 4.16. In Fig. 4.17 (b), the AQ circuit is enabled while the oscilloscope is still in persistence mode. All the avalanches, even the weakest ones, are detected and quenched by the AQ circuit. Therefore, it can be concluded that the proposed AQAR circuit not only can reduce the intrinsic afterpulsing of the SPAD, i.e., afterpulses generated by the main avalanche, but also can reduce the afterpulsing imposed by the parasitic effects. The PQ time is about 1.1 ns. By activating the AQ circuit, this time is reduced by 600 ps, resulting in a total quenching time of 500 ps in the AQ mode of operation. The avalanche is detected in less than 300 ps. Probably, this time is much shorter, however it is not easy to find the precise anode slope variation in the measurements. This faster

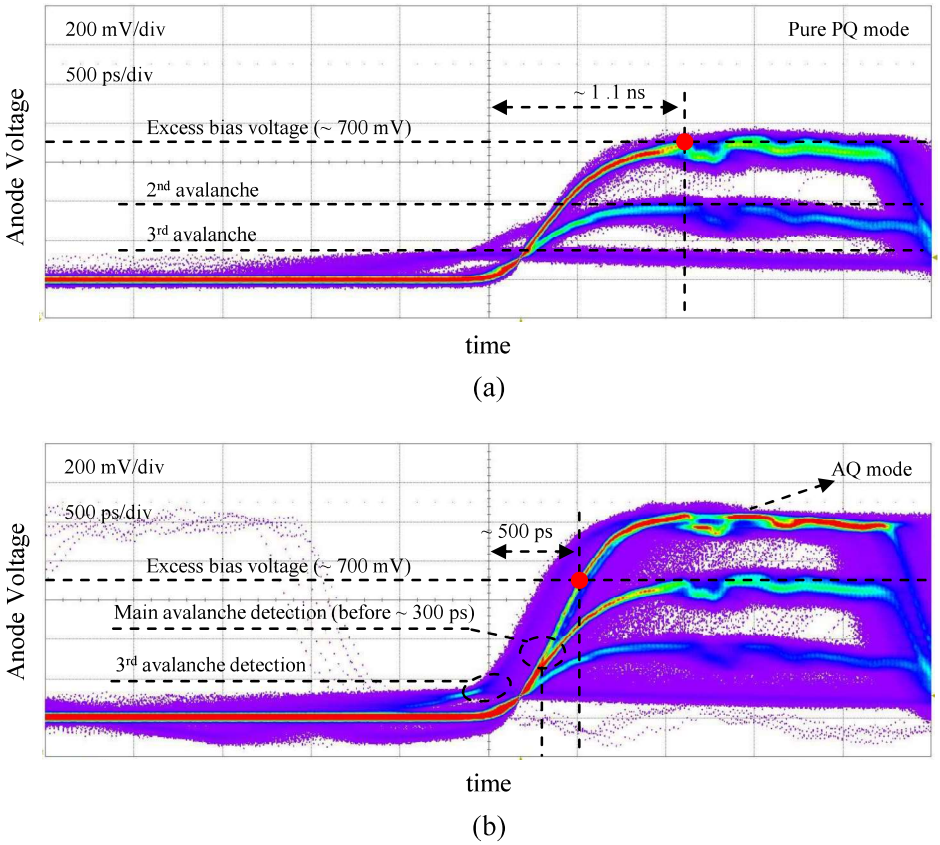


Fig. 4.17 Anode voltage transition in direct sensing pixel under 700 mV of excess bias voltage and $R_{sensitivity}$ = 533 kΩ, at 60 °C when the scope is in persistence mode: (a) PQ mode (b) when AQ is enabled.

quenching time in comparison with the quenching time in the PQ mode, demonstrates the effectiveness of the proposed AQAR circuit in avalanche charge and afterpulsing reduction.

Due to the added STIs in the monolithic 3D pixels (MS and InDS pixels), the noise level is expected to be higher than the DS pixel. Fig. 4.18 shows the anode voltage variations of MS pixel in PQ and AQ modes of operation with 6 ns dead time. The excess bias voltage is 500 mV, and the temperature is set at 40 °C. The digital representative of the analog anode voltage signal (yellow) is shown in pink. In the pure PQ mode, Fig. 4.18 (a), the SPAD is almost saturated by the noise. Many afterpulses are observable that the pixel is unable to distinguish all of them in the PQ operation. As a result, the digital signal has different pulse widths. That makes difficult the precise estimation of the SPAD noise. It is the actual behavior of the SPAD when the electronics are placed over it, when the circuit is not participating in the quenching process (AQ switch is off), while in the DS pixel, the SPAD is functional even with a higher excess bias voltage in the PQ mode. This

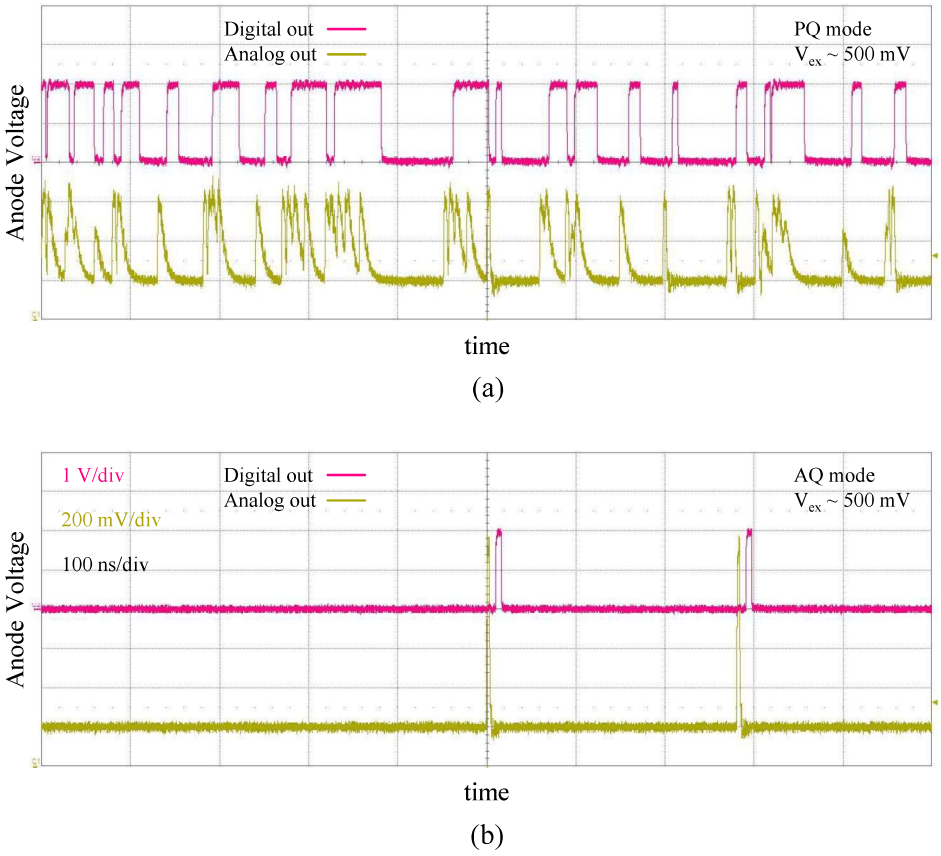


Fig. 4.18 Avalanches in the mixed sensing pixel under 500 mV of excess bias voltage at 40 °C with 6 ns dead time: (a) PQ mode (b) when AQ is enabled.

comparison clearly shows the impact of placing circuit on top of the SPAD. However, in the exact same conditions, when the AQ mode of operation is enabled, the SPAD goes back to an exploitable behavior and it is not saturated by the noise anymore, see Fig. 4.18 (b). Therefore, one can conclude that the added electronics on top of the SPAD increase the noise such the SPAD cannot operate well in its PQ mode of operation. However, with the assist of the electronics in the AQ mode, the SPAD becomes usable again, whereas the circuit does not occupy any pixel area in a BSI configuration. It is how the monolithic 3D SPAD pixel can mitigate the trade-off between the PDE and the SPAD noise.

Fig. 4.19 shows the anode voltage transition of MS pixel during an avalanche. V_{ex} is set to 500 mV and the sensitivity trimmer is set to 630 kΩ. The temperature is 60 °C and the oscilloscope is in persistence mode. The quenching time in PQ mode is around 1.1 ns and thanks to the AQ circuit it becomes as short as 500 ps in the AQ mode. The timing performance of the MS pixel is

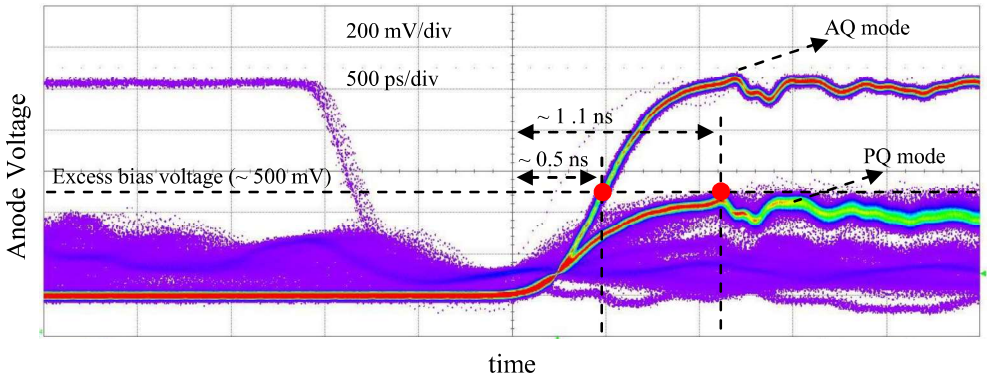


Fig. 4.19 Anode voltage transition in mixed sensing pixel under 500 mV of excess bias voltage and $R_{sensitivity}$ = 630 kΩ, at 60 °C when the scope is in persistence mode for both PQ and AQ modes of operation.

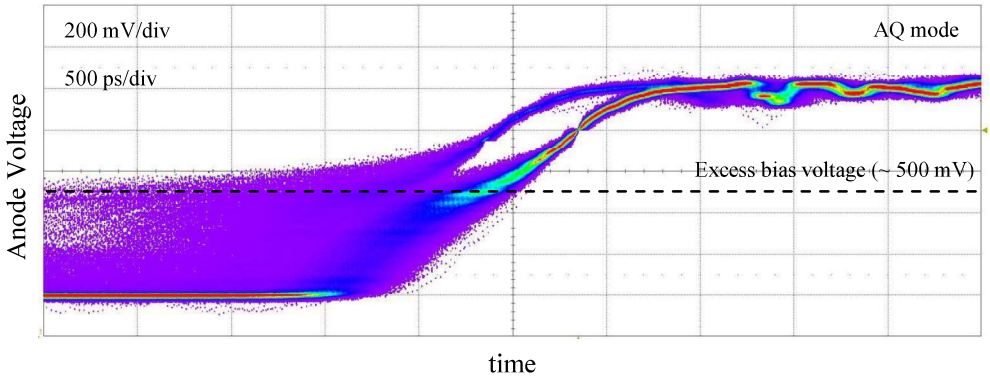


Fig. 4.20 Anode voltage transition in indirect sensing pixel under 500 mV of excess bias voltage and $R_{sensitivity}$ = 480 kΩ, at 60 °C when the scope is in persistence mode for AQ modes of operation.

similar to the DS pixel, even though it was expected to be slightly faster regarding the simulation results. One reason for this different result is the different excess bias voltages of the two pixels. Also, it is not easy to precisely tune both pixels at the same $V_{pre\text{-}bias}$ since there is no direct access to this point. Nevertheless, this transient response, is evidence of the effectiveness of the MS pixel in avalanche charge reduction.

The same measurements have been performed for the InDS pixel for the same V_{ex} and temperature, see Fig. 4.20. However, due to the lower sensitivity of this circuit, a lower trimmer value ($R_{sensitivity}$ = 480 kΩ) is chosen to increase the sensitivity in comparison with MS pixel. One can see that the AQ phase starts almost after the SPAD is passively quenched. This result agrees with the very low effectiveness of the InDS pixel in afterpulsing reduction in the simulation results

(almost zero). Here V_{ex} is lower, so the sensitivity of the circuit is even worse. It can be expected with a higher V_{ex}, the InDS circuit contributes more to the quenching process. Nonetheless, the circuit is able to detect avalanches and can handle an excess voltage 3 times higher than the technology headroom thanks to its indirect sensing method. In a less noisy SPAD, this feature can increase the PDP and jitter of the pixel.

4.5.2.2. Afterpulsing measurements

Fig. 4.21 shows the histograms of the avalanche inter-arrival time for DS pixel at three different temperatures (20, 40, and 60 °C). V_{ex} is 800 mV and the dead time is set to 6 ns. The measurements are performed in both PQ and AQ modes to evaluate the effectiveness of the DS circuit in afterpulsing reduction. At each temperature, measurements are performed at different sensitivity levels. At all the temperatures, the difference in afterpulsing value is clearly observable between PQ mode and AQ modes. This demonstrates the ability of the proposed AQ circuit in avalanche charge and afterpulsing reduction. It can be seen also that increasing the sensitivity level of the detection circuit (lower resistor value), reduces the afterpulsing, while a lower sensitivity level ensures a low power consumption. At 20 °C most of the SPAD noise is composed of afterpulsing effect. That makes it difficult to fit the SPAD DCR and this is why several DCR fits are observable in the histogram. It can be observed in Fig. 4.21 that by increasing the temperature the contribution of afterpulsing in the noise is decreased.

Fig. 4.22 shows the same measurements for the MS pixel. In this set of measurements, V_{ex} is set to 250 mV. Higher than this voltage, the SPAD is saturated by noise in the PQ mode due to the reasons explained before. Although the SPAD can operate normally up to 500 mV with the assistance of AQ circuit. Nevertheless, to evaluate the effectiveness of MS pixel in afterpulsing reduction, a comparison between AQ and PQ modes is mandatory thus, it limits the V_{ex} to 250 mV. Afterpulsing reduction is observable at each temperature and is more significant in comparison with the DS pixel. One can see in PQ mode, despite normal SPADs, the predominant source of the noise is afterpulsing effect. This unnormal phenomenon makes it challenging to find a DCR fit due to the significant contribution of afterpulsing effect in the overall noise. As soon as switching to AQ mode however, the SPAD behavior returns to a more conventional behavior and the DCR distribution over the time is observable.

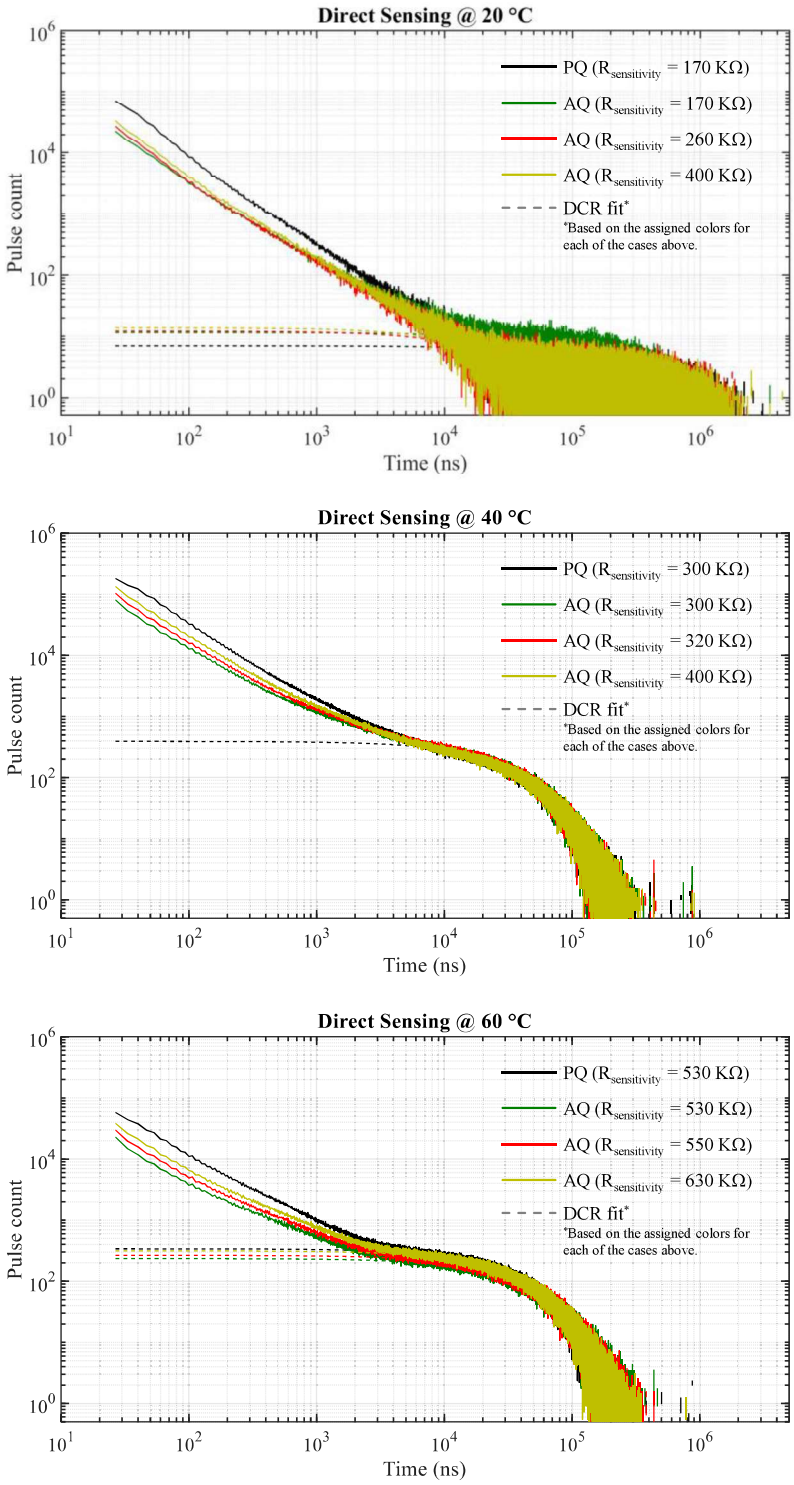


Fig. 4.21 Avalanche inter-arrival time histograms and DCR fits for direct sensing pixel at different sensitivity levels and temperatures: 20, 40, and 60 °C (from top to bottom). Dead time is 6 ns and excess bias voltage is 800 mV.

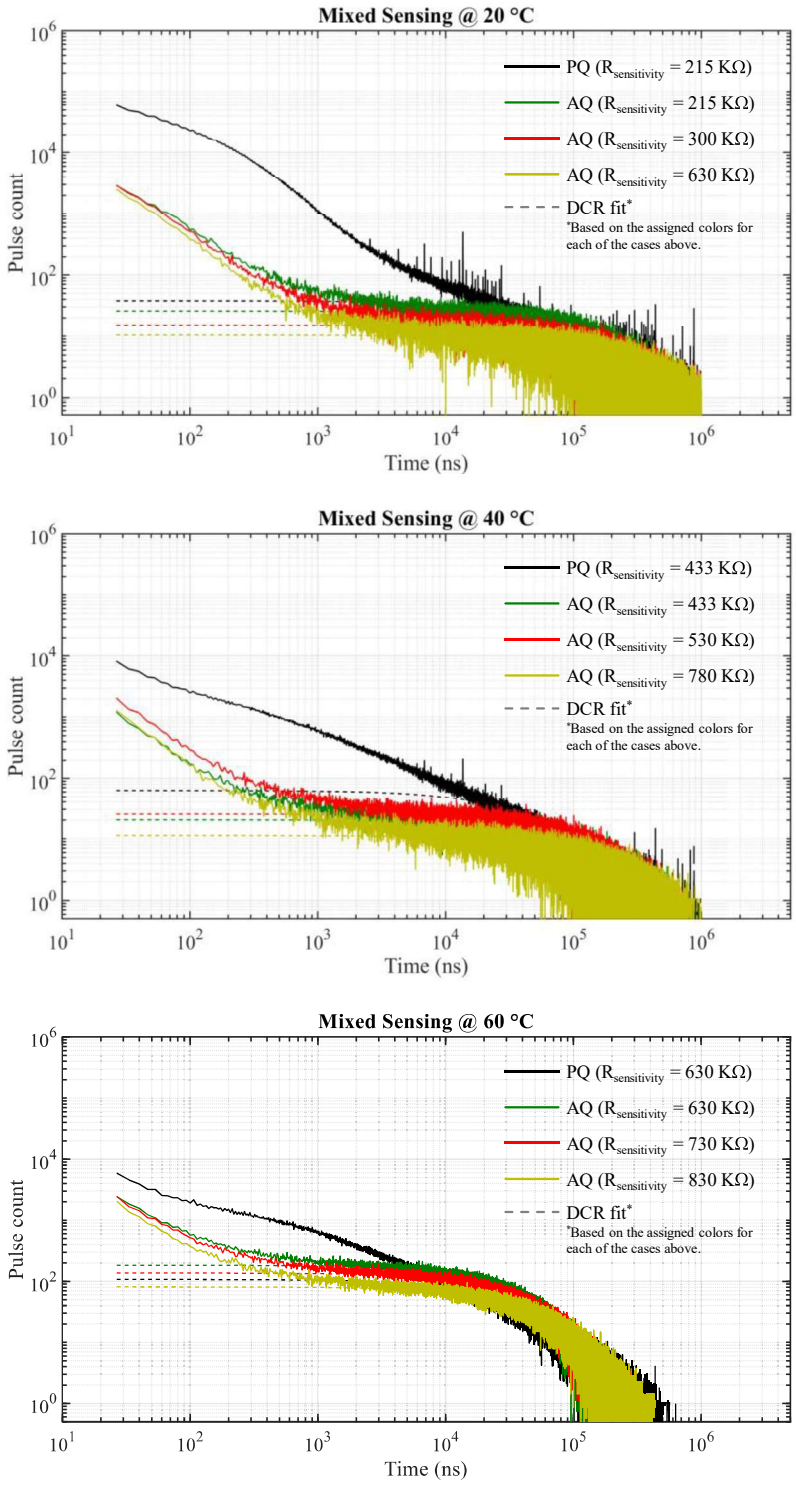


Fig. 4.22 Avalanche inter-arrival time histograms and DCR fits for mixed sensing pixel at different sensitivity levels and temperatures: 20, 40, and 60 °C (from top to bottom). Dead time is 6 ns and excess bias voltage is 250 mV.

Here also the difference in the histograms at different sensitivity levels is visible. In some cases, this difference is not clearly distinguishable at the peak of afterpulsing part, though different DCR levels for each histogram are discernible. The higher the sensitivity level (lower resistor values), the higher the observed DCR level, because more events can be detected by the detection circuit. In fact, with a lower sensitivity, the weak avalanches can be missed.

Fig. 4.23 presents the pure afterpulsing rate (DCR subtracted) of the two pixels at their different sensitivity levels and temperatures. The afterpulsing reduction as the relative difference between the Afterpulsing Count Rate (ACR) in PQ mode and the ACR in the different AQ modes is calculated and presented on the figures:

$$Afterpulsing\ reduction_x = \frac{ACR_{PQ} - ACR_{AQ(R_{sensitivity\ x})}}{ACR_{PQ}}$$

In the DS pixel (Fig. 4.23 (a)), the minimum afterpulsing reduction of 25 % is achieved at 20 °C with the lowest sensitivity. The maximum afterpulsing reduction is 62 % at 60 °C with the highest possible sensitivity of the pixel. This value corresponds to at least 62 % avalanche charge reduction. As expected, by increasing the sensitivity, afterpulsing further reduces. While, by decreasing the sensitivity, the afterpulsing reduction is significant, however, the power consumption reduces drastically. It implies a trade-off between the power consumption and the afterpulsing reduction.

In the MS pixel (Fig. 4.23 (b)), the difference in ACR between PQ and AQ modes is remarkably high (minimum 90 % relative difference). Therefore, the figures are presented in logarithmic scales. The maximum afterpulsing reduction is 96 %. In the MS pixel the trend in afterpulsing reduction versus sensitivity levels is not as clear as the DS pixel. One reason can be the different DCR levels.

Using APP, as another indicator of the quenching circuit efficiency in afterpulsing reduction, reveals the afterpulsing reduction versus sensitivity levels trend even for MS pixel. Fig. 4.24 shows the APP of the pixels at different temperatures and sensitivity levels. It is clear that by increasing the sensitivity, APP is reducing. In the MS pixel at 60 °C, from an APP of more than 40 % in PQ mode, the pixel achieves an APP of around 3 % by activating the AQ circuit ($R_{sensitivity}$ = 630 kΩ). In DS pixel at 20 °C, the APP in PQ mode is not respecting the trend. The reason can be the uncertainty in its DCR fit as explained before.

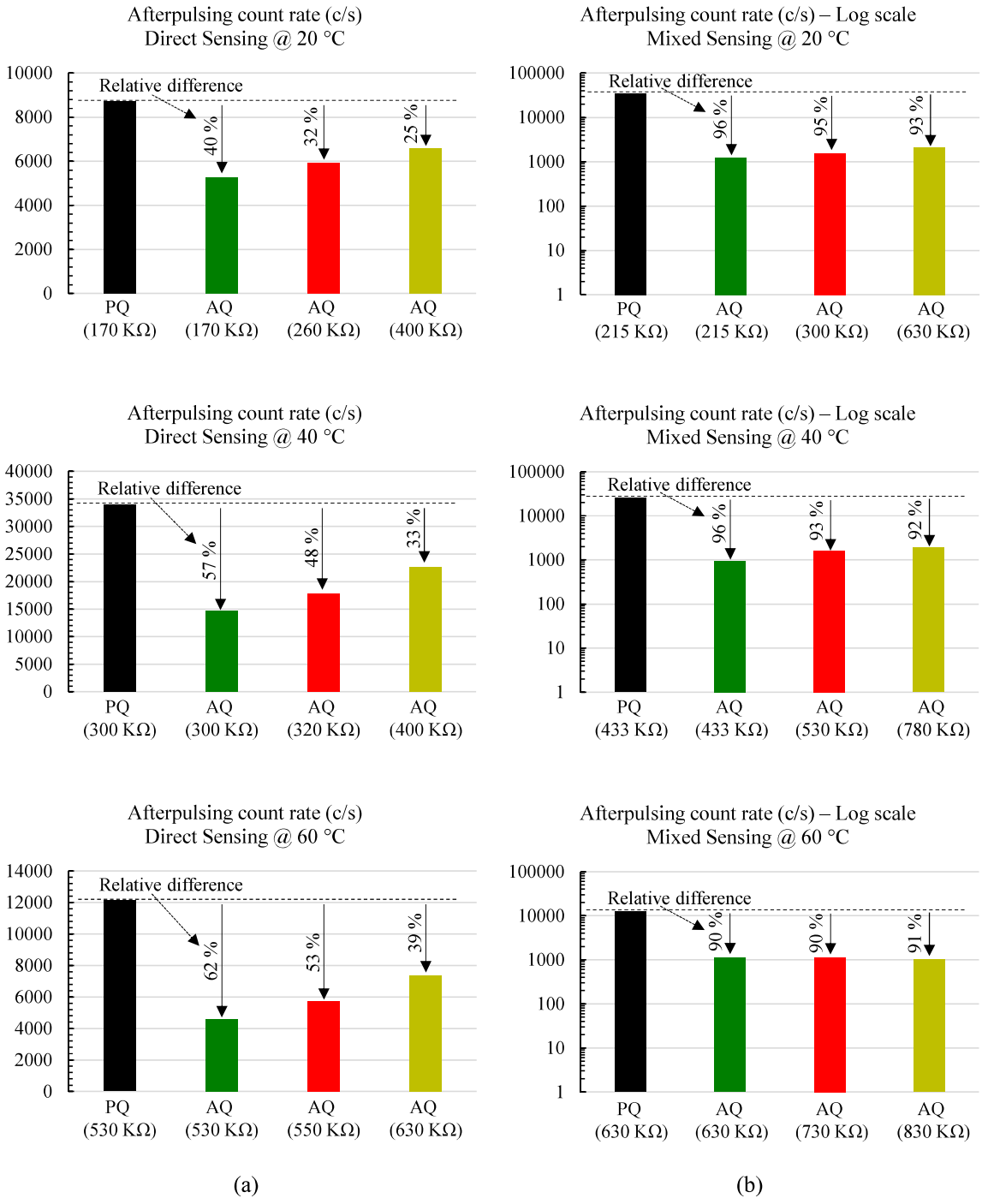


Fig. 4.23 Afterpulsing count rates in PQ and AQ modes at different sensitivity levels with 6 ns dead time for different temperatures: 20, 40, and 60 °C (from top to bottom). (a) direct sensing pixel under 800 mV of excess bias voltage (b) mixed sensing pixel under 250 mV of excess bias voltage. The relative difference in afterpulsing count rates between PQ and AQ modes are indicated on the figures. The resistor values on the x axis are different values of $R_{sensitivity}$.

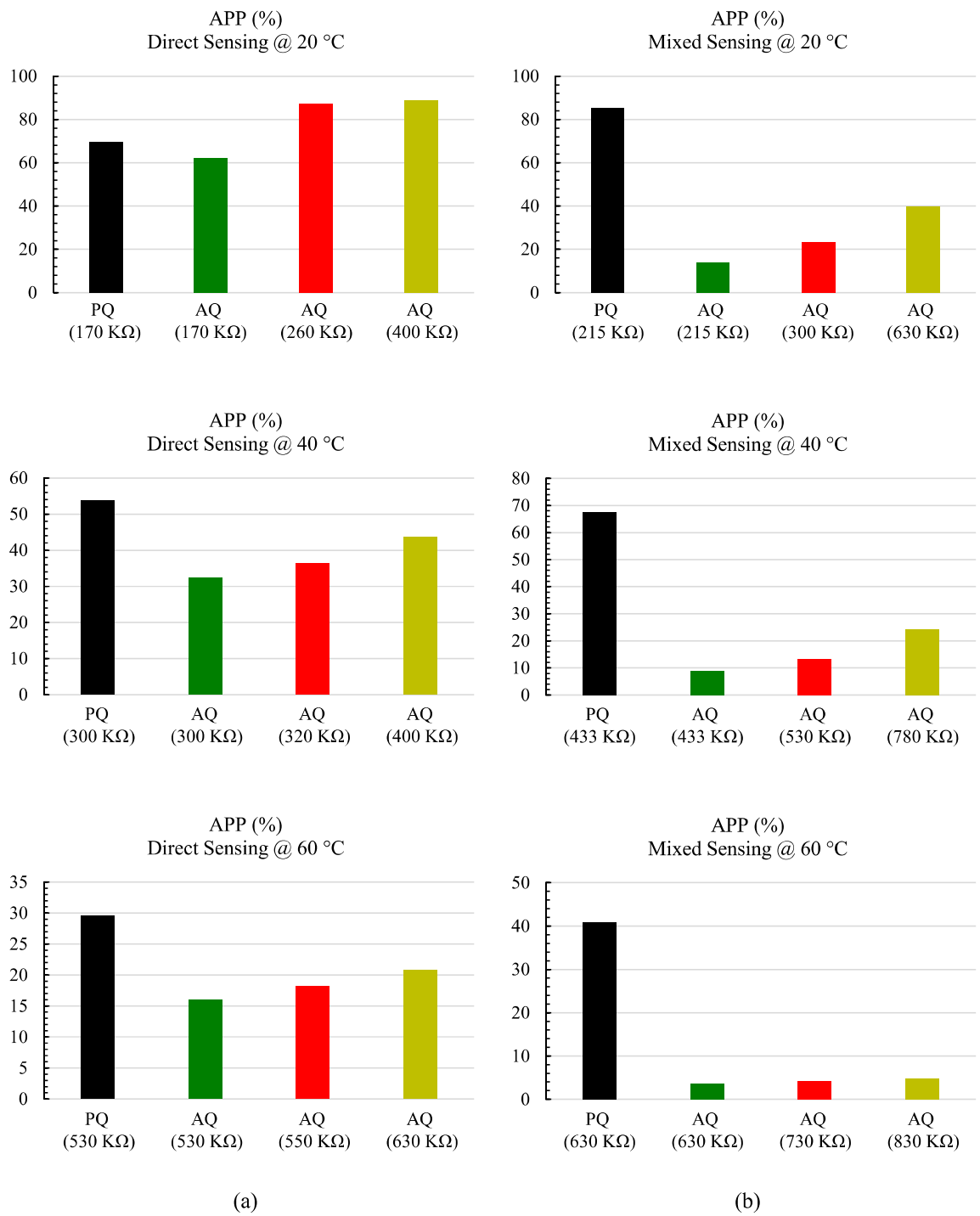


Fig. 4.24 Afterpulsing Probability Percentage (APP) in PQ and AQ modes at different sensitivity levels with 6 ns dead time for different temperatures: 20, 40, and 60 °C (from top to bottom). (a) direct sensing pixel under 800 mV of excess bias voltage (b) mixed sensing pixel under 250 mV of excess bias voltage. The resistor values on the x axis are different values of $R_{sensitivity}$.

Chapter 5
Conclusions and Perspectives

5.1. Conclusions

The objective of this work is to design and implement avalanche detection and AQ circuits for the first SPAD realized in an FD-SOI CMOS technology that improves the performance parameters of the SPAD such as noise and PDE. As the first of its kind, the used SPAD in this work is an unoptimized device. A sensitive avalanche detection circuit also can help to characterize the SPAD in order to identify the required optimizations.

In chapter 1, a review on the SPAD physics and its operation principles is made. The avalanche breakdown in a reverse biased pn junction is explained as the fundamental phenomenon in SPAD which can convert a single photon to a measurable avalanche of charge carriers (avalanche current) in the SCR of the SPAD. Then the most important parameters of SPAD such as PDP, jitter, DCR, and afterpulsing are introduced and some of the existing techniques in the literature to improve them are presented.

Afterpulsing effect is a correlated noise source of the SPAD that puts a limit on the maximum achievable photon count rate of a SPAD pixel. AQ circuits as the most efficient solution in reducing the afterpulsing noise are discussed and their advantages and disadvantages are investigated. An AQ circuit is able to reduce the afterpulsing effect through the avalanche charge reduction while not deteriorating the maximum count rate of the pixel. However, the required circuitry to realize such a strategy is area consuming and thus degrades the fill factor and PDE of the pixel. Thus, there are trade-offs between afterpulsing effect, maximum photon count rate, fill factor, and PDE. Stacked 3D pixels as a solution to mitigate these trade-offs are introduced in the literature. However, they suffer from the excessive cost of fabrication as they need two separate dies. Therefore, the aim of this thesis is to find an ultimate solution to reduce SPAD noise while keeping the maximum count rate, fill factor, and PDE as high as possible in a more cost-effective way than the stacked 3D technology.

Recently, the idea of a monolithic 3D SPAD pixel in the FD-SOI CMOS technology has been presented in the literature. In this technology a Buried Oxide (BOX) layer, isolates the devices from their corresponding wells, thus allowing a monolithic integration of SPAD and the devices placed on top of the SPAD over the BOX layer. The authors have shown it is possible to sense the avalanche through a simple circuitry (a voltage divider) over the SPAD thanks to the body biasing effect. However, no quenching circuit yet has been integrated into this inherent 3D structure. The integration of an AQ circuit in this monolithic 3D structure is another objective of this work which aligns with the first objective (mitigating the trade-offs between noise, count rate and PDE). Placing a functional and efficient AQ circuit on top of the SPAD active area in a BSI configuration, can strongly reduce the dependence of FF on the pixel electronics and thus alleviate the mentioned trade-offs.

In chapter 2, the first attempt to design an AQAR circuit in a 28 nm FD-SOI CMOS is presented. The efficiency of an AQ circuit in afterpulsing reduction strongly depends on the avalanche detection time. This is why in many of the works, researchers use bulky and very power consuming comparators as a detection circuit which highly degrades PDE. Therefore, we propose an inverter as a detection circuit to be as compact as possible. In the literature, inverters are used as detection circuits and their sensitivity is increased by lowering their switching threshold through a proper sizing. The novelty of our design is to mix the sizing method with body biasing technique to further increase the sensitivity of the inverter and shorten the detection time. In FD-SOI CMOS technology the BOX layer prevents forming of the parasitic diodes between drain/source and body of the MOSFETs. This allows applying a body voltage up to the maximum sustainable voltage of the back-gate oxide (BOX layer) which is around 3V in the used technology. Through the body voltage it is possible to modulate the threshold voltage of the MOSFETs. The rate of this variation is defined as the body factor which is around 85 mV/V in this technology. It means that by applying 3 V body voltage, the threshold voltage changes to 255 mV. In an inverter decreasing and increasing the threshold voltages of NMOS and PMOS respectively results in decreasing its switching threshold by the same amount. Thus, in combination with a proper sizing, the proposed body biased inverter can achieve an ultra-low switching threshold as a detection circuit. According to the post layout simulation results, with a body voltage of 2.5 V the avalanche is detected in less than 40 ps and quenched actively in only 150 ps which results in about 42 % avalanche charge reduction in AQ mode in comparison with the pure PQ mode. Afterpulsing measurements show

more than 50 % APP reduction in AQ mode with 2 V of body voltage in comparison with the PQ mode, with a dead time as short as 5 ns. Thus, the proposed body biased inverter based AQAR circuit is able to effectively mitigate the trade-off between the afterpulsing noise and the maximum count rate. In this chapter, the target is not the monolithic integration, even though the circuit can be very compact. Due to the different MOSFET families in the used technology, it is possible to implement the inverter in one shared well to avoid well separation for body biasing requirements. That results in a very compact detection circuit that does not severely impact the fill factor.

Chapter 3 focuses on the pursuit of two objectives at the same time: Design an even faster AQ circuit which has the possibility to be integrated with the monolithic 3D structure simultaneously. As long as an inverter operates as a digital gate, it suffers from its propagation delay. The body biased inverter strongly decreases this delay, however, it cannot completely remove it. In this chapter we propose to bias the inverter close to its switching threshold to make this delay minimal. In this case, the minimum voltage variations at the input of the inverter are multiplied by the maximum gain of the inverter and changes its output.

To achieve the monolithic 3D pixel, this pre-biasing must be done in a way to be immune against the body voltage variations of the transistors over the SPAD. In fact, the body of MOSFETs in this structure is the dynamic node of the SPAD (anode). Occurring an avalanche increases the anode voltage of the SPAD or in other words the body voltage of the MOSFETs. Therefore, an avalanche imposes an unwanted body biasing to the transistors on top of the SPAD. To solve this issue, a T-gate is placed in parallel with the detection inverter. The T-gate biases the inverter at its switching threshold by shorting its input and output at any body voltage of its transistor. Also choosing an inverter as the detection circuit is an advantage for AQ circuit in the monolithic 3D structure. That is, as it is explained in chapter 2, applying body voltage to an inverter modulates its switching threshold. In this case, anode voltage rise decreases the switching threshold of the detection inverter and results in an even faster detection and thus a faster quenching. Based on this detection circuit three AQAR circuits are presented. In Direct Sensing (DS) pixel, the electronic is not over the SPAD and it is designed to prove the functionality of the proposed AQ circuit in the standard pixels. It is named DS since the avalanche is sensed through the gate of the MOSFETs of the detection inverter. The two other AQAR circuits can be integrated in the monolithic 3D structure. The first is Mixed Sensing (MS) pixel which is the first ever reported monolithic 3D

active pixel. In this circuit the avalanche is sensed through the gate and body of the MOSFETs of the inverter simultaneously. This is a new avalanche sensing method that can also be used in the DS pixel through a wired connection between the anode and the body of MOSFETs. Another pixel is sensed the avalanche only through the body of its transistors, named Indirect Sensing (InDS) pixel. In this pixel, it is possible to apply an excess bias voltage three times higher than the technology voltage headroom that improves the PDP and jitter of the pixel. According to the post layout simulation results, MS pixel, as the fastest and most efficient pixel, detects an avalanche sooner than 50 ps and quenches it actively in 160 ps. This results in more than 50 % avalanche charge reduction compared to the operation of the pixel in PQ mode. In the DS case, the value of this reduction is around 44 %. The static power consumption is around 100 µW, mostly consumed by the delay lines. The quenching circuit itself consumes around 43 µW due to the biasing of its detection inverter.

In the measurements, the effects of placing electronics over the SPAD are investigated for the first time. It is observed that the added Shallow Trench Isolators (STIs) due to the circuitry placement on the SPAD, increases both DCR and afterpulsing noises. Furthermore, it leads to an increase in premature breakdown across the SPAD, resulting in the breakdown voltage of the MS pixel being at least 150 mV lower than that of the DS pixel. Consequently, it limits the excess bias voltage of the MS pixel to be around three times less than DS pixel in the PQ mode in order to avoid the pixel saturation by the noise sources. However, by assisting the proposed AQAR circuit, the noise level is lowered such that the MS pixel can operate normally in the higher excess bias voltages in its AQ mode of operation. In other word, the added electronics on top of the SPAD, increase the noise level in a way that SPAD is not functioning well in high excess bias voltages, however, the proposed AQ circuit is highly efficient in noise reduction that compensates the added noise, while not consuming any pixel area in a BSI configuration. In DS pixel 85% of afterpulsing count reduction compared to the PQ mode is observed at 800 mV of excess bias voltage. In MS pixel, to make a comparison between AQ and PQ modes, excess bias voltage is limited to 250 mV. In this case an afterpulsing count reduction of 94 % is measured.

After realizing the first integration of an AQAR circuit in a monolithic 3D SPAD pixel, in chapter 4 a power efficient AQAR circuit is presented which is almost as sensitive as the presented circuit in chapter 3 an also compatible with the monolithic 3D structure. Instead of biasing the

detection inverter at its switching threshold where it consumes the maximum static power, biasing the inverter at a voltage slightly below this point can strongly reduce the static power consumption while the sensitivity is still high enough. The proposed biasing circuit has three essential features: during an avalanche it is not fighting against the anode voltage rise, during the reset phase, it brings back the biasing point to its initial value as soon as possible, and in order to be compatible with the monolithic 3D structure, it is body voltage independent. Therefore, we present for the first time a body voltage independent current source to bias the proposed AQAR circuits. In this chapter, same as chapter 3, the proposed circuit is designed in three different forms: DS, MS, and InDS pixels. According to the post layout simulation results, DS and MS pixels quench the avalanche actively in 180 ps and 164 ps respectively which results in 35 % and 48 % avalanche charge reduction in comparison with their PQ mode of operation. In experimental results a quenching time of about 500 ps is measured for DS pixel with 700 mV of excess bias voltage, which is 600 ps faster than the PQ mode. Almost the same timing is measured for MS pixel at 500 mV of excess bias voltage. Here also the same effects of placing electronics over the SPAD are observed as chapter 3: increased noise level in the monolithic 3D structure and compensating for this noise thanks to the AQ circuit. Afterpulsing measurements show 62 % and 96 % afterpulsing count reduction for DS and MS pixels, respectively. In MS pixel, thanks to the proposed AQ circuit, an APP of more than 40 % in PQ mode, reduces to around 3 % in AQ mode. One can conclude that the proposed AQAR circuits are vital for the monolithic 3D pixels and can turn a noisy SPAD to almost a well behaved SPAD. All these are achieved with a static power consumption lower than 15 µW, three times less than the proposed circuit in chapter 3.

Table 5.1 compares the state of the arts with the proposed pixels in this work. The proposed pixels have the shortest measured quenching time while consuming a low power. Unfortunately, there are not a lot of works which compare the afterpulsing effect between AQ and PQ modes. They mostly report the Afterpulsing Probability (APP) in the AQ mode. However, with only this value it is not possible to evaluate the effectiveness of the AQ circuit. Nevertheless, among the other works which present the APP reduction or avalanche charge reduction, the proposed pixels have the highest afterpulsing reduction.

TABLE 5.1
STATE OF THE ARTS COMPARISON WITH THE PROPOSED PIXELS IN THIS WORK.

Work	Technology (nm)	Dead time (ns)	Quenching time (ns)	Afterpulsing/avalanche charge reduction (%)	Circuit area (μm^2)	Power (mW)
[113]	CMOS 350	9.5	1.04*	67 ‡	130×134	4.8
[142]	-	500	~ 4	80	-	-
[143]	BiCMOS 350	-	<2	-	100×370	30 @ 1 Mc/s
[145]	CMOS 350	20	3	-	672	0.5 @ 50 Mc/s
[146]	CMOS 180	4	0.7*	-	306	-
[148]	CMOS 350	7.86	1.7	34 ‡	236×108	13.7 @ 33 Mc/s
[150]	BiCMOS 350	10	~2	36 ‡⁂	37800	58 @ 40Mc/s
BBI	FD-SOI CMOS 28	5	450 (V_{ex}= 0.5)	50 ‡	24×20	0.007
Ticklish MS	FD-SOI CMOS 28	8	625 (V_{ex}= 0.5)	94	14×15	0.1 (0.04)
Ticklish DS	FD-SOI CMOS 28	8	825 (V_{ex}= 0.5)	85	14×15	0.1 (0.04)
Pre-bias MS	FD-SOI CMOS 28	6	500 (Vex= 0.5)	96 (90 ‡)	145	0.014
Pre-bias DS	FD-SOI CMOS 28	6	500 (V_{ex}= 0.7)	62 (50 ‡)	145	0.014

†Representing the medium diagonal or height. *Simulation results. ‡Representing Afterpulsing Probability (APP) reduction. ⁂Estimated value (no measurement or simulation).

5.2. Perspectives

We have presented three different AQAR circuits and proved their functionality and effectiveness in the afterpulsing reduction. Also, the first ever monolithic 3D SPAD active pixel is designed, and measurements justify its operation. A new avalanche sensing method named mixed sensing was also introduced that improves the avalanche detection time. An innovative body voltage independent current source is introduced as well for biasing purposes in the monolithic 3D pixel in FD-SOI CMOS technology. In this section the possible improvements of the proposed solutions in this thesis and the probable future works in line with the current works are discussed.

- In chapter 2, for realizing the body biased inverter as an avalanche detection circuit, since both NMOS and PMOS need the same body voltage to reduce the switching threshold of the inverter, it is possible to use NMOS and PMOS from FW and RW families and mixed them to form an inverter in a same well. This minimizes the circuit

area and thus increases the fill factor and PDE of the pixel. This is the exact solution that is used for realizing detection inverters and all other circuitry in monolithic 3D pixel in chapter 3 and 4.

- In the InDS pixels, since the detection inverter can tolerate a voltage higher than the technology voltage headroom, it is possible to use cascode transistors in series with AQ and AR transistors in order to achieve a higher excess bias voltage in AQ mode that can also improve the PDP and jitter of the pixel while having all the benefits of an AQ circuit such as afterpulsing reduction. In fact, the cascode technique can be used for all the DS and MS pixels in chapters 3 and 4 as well thanks to the coupling capacitor between anode and the gates of the detection inverters. Also, the sensitivity of the InDS pixels and MS pixels depends on the excess bias voltage since the anode voltage is actually their body voltage. Thus, a higher excess bias voltage can result in a more sensitive detection circuit.

- The MS approach can also be applied to the direct sensing pixels. One just needs to connect the body of the detection inverter in the DS pixel to the anode voltage by wire. The resulting pixel, even though is not a 3D pixel, would have almost the same and even slightly higher sensitivity than the MS pixels.

- In order to have the maximum achievable count rate with the proposed pixel, it is possible to use the same feedback path as the AQ feedback for the AR path with the minimum possible delay, instead of a finely tunable delay lines as it is used in this work for the measurements purposes. It will not degrade the noise performance of the pixel, because the AQ circuit strongly reduces the afterpulsing effect.

- In general, an optimized SPAD could improve the circuit characterization and strongly increase the pixel performance. A less noisy SPAD allows applying a higher excess bias voltage that can increase the sensitivity of the detection inverters.

- In the proposed monolithic 3D pixels still are places for further circuit integration. In the worst case, the circuit covers 42 % of the SPAD active area. Therefore, it is still possible to add a compact TDC or other more advanced electronics.

- Since the monolithic 3D pixels are designed to work in a BSI configuration, thus, it is mandatory to apply a die thinning for further characterizing the pixel such as PDP measurements.

References

[1] S. M. Sze and M. K. Lee, *Semiconductor Devices: Physics and Technology, 3rd Edition: Physics and Technology*. Wiley-Blackwell, 2012. [Online]. Available: https://books.google.fr/books?id=WdcbAAAAQBAJ

[2] J. Y. Tang and K. Hess, "Impact ionization of electrons in silicon (steady state)," *Journal of Applied Physics*, vol. 54, no. 9, pp. 5139–5144, Sep. 1983, doi: 10.1063/1.332737.

[3] A. G. Chynoweth, "Ionization Rates for Electrons and Holes in Silicon," *Phys. Rev.*, vol. 109, no. 5, pp. 1537–1540, Mar. 1958, doi: 10.1103/PhysRev.109.1537.

[4] K. G. McKay and K. B. McAfee, "Electron Multiplication in Silicon and Germanium," *Phys. Rev.*, vol. 91, no. 5, pp. 1079–1084, Sep. 1953, doi: 10.1103/PhysRev.91.1079.

[5] K. G. McKay, "Avalanche Breakdown in Silicon," *Phys. Rev.*, vol. 94, no. 4, pp. 877–884, May 1954, doi: 10.1103/PhysRev.94.877.

[6] O. Christensen, "Quantum efficiency of the internal photoelectric effect in silicon and germanium," *Journal of Applied Physics*, vol. 47, no. 2, pp. 689–695, Aug. 2008, doi: 10.1063/1.322635.

[7] V. I. Kovalev, "Nature of Photoelectric Effect in a Ge-on-Si SPAD at Ultralow Energy in Incident Pulsed Laser Radiation," *Optics*, vol. 2, no. 1, Art. no. 1, Mar. 2021, doi: 10.3390/opt2010004.

[8] P. Seitz and A. J. Theuwissen, Eds., *Single-Photon Imaging*, vol. 160. in Springer Series in Optical Sciences, vol. 160. Berlin, Heidelberg: Springer, 2011. doi: 10.1007/978-3-642-18443-7.

[9] F. Zappa, S. Tisa, A. Tosi, and S. Cova, "Principles and features of single-photon avalanche diode arrays," *Sensors and Actuators A: Physical*, vol. 140, no. 1, pp. 103–112, Oct. 2007, doi: 10.1016/j.sna.2007.06.021.

[10] S. Cova, A. Longoni, and A. Andreoni, "Towards picosecond resolution with single-photon avalanche diodes," *Review of Scientific Instruments*, vol. 52, no. 3, pp. 408–412, Mar. 1981, doi: 10.1063/1.1136594.

[11] S. Cova, A. Lacaita, M. Ghioni, G. Ripamonti, and T. A. Louis, "20-ps timing resolution with single-photon avalanche diodes," *Review of Scientific Instruments*, vol. 60, no. 6, pp. 1104–1110, Jun. 1989, doi: 10.1063/1.1140324.

[12] S. Rink *et al.*, "A self-sustaining Single Photon Avalanche Diode Model," in *ESSDERC 2022 - IEEE 52nd European Solid-State Device Research Conference (ESSDERC)*, Sep. 2022, pp. 281–284. doi: 10.1109/ESSDERC55479.2022.9947120.

[13] A. Gallivanoni, I. Rech, and M. Ghioni, "Progress in Quenching Circuits for Single Photon Avalanche Diodes," *IEEE Transactions on Nuclear Science*, vol. 57, no. 6, pp. 3815–3826, Dec. 2010, doi: 10.1109/TNS.2010.2074213.

[14] F. Zappa, A. Lotito, A. C. Giudice, S. Cova, and M. Ghioni, "Monolithic active-quenching and active-reset circuit for single-photon avalanche detectors," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 7, pp. 1298–1301, Jul. 2003, doi: 10.1109/JSSC.2003.813291.

[15] V. Savuskan, M. Javitt, G. Visokolov, I. Brouk, and Y. Nemirovsky, "Selecting Single Photon Avalanche Diode (SPAD) Passive-Quenching Resistance: An Approach," *IEEE Sensors Journal*, vol. 13, no. 6, pp. 2322–2328, Jun. 2013, doi: 10.1109/JSEN.2013.2253603.

[16] G. Giustolisi, A. D. Grasso, and G. Palumbo, "Integrated Quenching-and-Reset Circuit for Single-Photon Avalanche Diodes," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 1, pp. 271–277, Jan. 2015, doi: 10.1109/TIM.2014.2338652.

[17] A. Katz, T. Blank, A. Fenigstein, T. Leitner, and Y. Nemirovsky, "Active-Reset for the N+P Single-Ended SPAD Used in the NIR LiDAR Receivers," *IEEE Transactions on Electron Devices*, vol. 66, no. 12, pp. 5191–5195, Dec. 2019, doi: 10.1109/TED.2019.2946015.

[18] C. Bruschini, H. Homulle, I. M. Antolovic, S. Burri, and E. Charbon, “Single-photon avalanche diode imagers in biophotonics: review and outlook,” *Light Sci Appl*, vol. 8, no. 1, Art. no. 1, Sep. 2019, doi: 10.1038/s41377-019-0191-5.

[19] M. Perenzoni, N. Massari, D. Perenzoni, L. Gasparini, and D. Stoppa, “A 160 × 120 Pixel Analog-Counting Single-Photon Imager With Time-Gating and Self-Referenced Column-Parallel A/D Conversion for Fluorescence Lifetime Imaging,” *IEEE Journal of Solid-State Circuits*, vol. 51, no. 1, pp. 155–167, Jan. 2016, doi: 10.1109/JSSC.2015.2482497.

[20] M. Vitali *et al.*, “A Single-Photon Avalanche Camera for Fluorescence Lifetime Imaging Microscopy and Correlation Spectroscopy,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 20, no. 6, pp. 344–353, Nov. 2014, doi: 10.1109/JSTQE.2014.2333238.

[21] R. M. Field, S. Realov, and K. L. Shepard, “A 100 fps, Time-Correlated Single-Photon-Counting-Based Fluorescence-Lifetime Imager in 130 nm CMOS,” *IEEE Journal of Solid-State Circuits*, vol. 49, no. 4, pp. 867–880, Apr. 2014, doi: 10.1109/JSSC.2013.2293777.

[22] L. Pancheri, N. Massari, and D. Stoppa, “SPAD Image Sensor With Analog Counting Pixel for Time-Resolved Fluorescence Detection,” *IEEE Transactions on Electron Devices*, vol. 60, no. 10, pp. 3442–3449, Oct. 2013, doi: 10.1109/TED.2013.2276752.

[23] Y. Maruyama and E. Charbon, “An all-digital, time-gated 128X128 spad array for on-chip, filter-less fluorescence detection,” in *2011 16th International Solid-State Sensors, Actuators and Microsystems Conference*, Jun. 2011, pp. 1180–1183. doi: 10.1109/TRANSDUCERS.2011.5969324.

[24] D.-U. Li *et al.*, “Real-time fluorescence lifetime imaging system with a 32 × 32 0.13µm CMOS low dark-count single-photon avalanche diode array,” *Opt. Express, OE*, vol. 18, no. 10, pp. 10257–10269, May 2010, doi: 10.1364/OE.18.010257.

[25] M. Wayne, A. Ulku, A. Ardelean, P. Mos, C. Bruschini, and E. Charbon, “A 500 × 500 Dual-Gate SPAD Imager With 100% Temporal Aperture and 1 ns Minimum Gate Length for FLIM and Phasor Imaging Applications,” *IEEE Transactions on Electron Devices*, vol. 69, no. 6, pp. 2865–2872, Jun. 2022, doi: 10.1109/TED.2022.3168249.

[26] A. C. Ulku *et al.*, “A 512 × 512 SPAD Image Sensor With Integrated Gating for Widefield FLIM,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 25, no. 1, pp. 1–12, Jan. 2019, doi: 10.1109/JSTQE.2018.2867439.

[27] A. Diéguez, J. Canals, N. Franch, J. Diéguez, O. Alonso, and A. Vilà, “A Compact Analog Histogramming SPAD-Based CMOS Chip for Time-Resolved Fluorescence,” *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, no. 2, pp. 343–351, Apr. 2019, doi: 10.1109/TBCAS.2019.2892825.

[28] S. Jahromi, J.-P. Jansson, P. Keränen, and J. Kostamovaara, “A 32 × 128 SPAD-257 TDC Receiver IC for Pulsed TOF Solid-State 3-D Imaging,” *IEEE Journal of Solid-State Circuits*, vol. 55, no. 7, pp. 1960–1970, Jul. 2020, doi: 10.1109/JSSC.2020.2970704.

[29] M.-A. Tétrault *et al.*, “Real-Time Discrete SPAD Array Readout Architecture for Time of Flight PET,” *IEEE Transactions on Nuclear Science*, vol. 62, no. 3, pp. 1077–1082, Jun. 2015, doi: 10.1109/TNS.2015.2409783.

[30] A. Ronchini Ximenes, P. Padmanabhan, M.-J. Lee, Y. Yamashita, D.-N. Yaung, and E. Charbon, “A Modular, Direct Time-of-Flight Depth Sensor in 45/65-nm 3-D-Stacked CMOS Technology,” *IEEE Journal of Solid-State Circuits*, vol. 54, no. 11, pp. 3203–3214, Nov. 2019, doi: 10.1109/JSSC.2019.2938412.

[31] F. Villa *et al.*, “CMOS Imager With 1024 SPADs and TDCs for Single-Photon Timing and 3-D Time-of-Flight,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 20, no. 6, pp. 364–373, Nov. 2014, doi: 10.1109/JSTQE.2014.2342197.

[32] C. Niclass, M. Soga, H. Matsubara, S. Kato, and M. Kagami, “A 100-m Range 10-Frame/s 340 × 96-Pixel Time-of-Flight Depth Sensor in 0.18-\mu\hboxm CMOS,” *IEEE Journal of Solid-State Circuits*, vol. 48, no. 2, pp. 559–572, Feb. 2013, doi: 10.1109/JSSC.2012.2227607.
[33] H. Xu, D. Perenzoni, A. Tomasi, and N. Massari, “A 16 × 16 Pixel Post-Processing Free Quantum Random Number Generator Based on SPADs,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 5, pp. 627–631, May 2018, doi: 10.1109/TCSII.2018.2821904.
[34] N. Massari *et al.*, “16.3 A 16×16 pixels SPAD-based 128-Mb/s quantum random number generator with −74dB light rejection ratio and −6.7ppm/°C bias sensitivity on temperature,” in *2016 IEEE International Solid-State Circuits Conference (ISSCC)*, Jan. 2016, pp. 292–293. doi: 10.1109/ISSCC.2016.7418022.
[35] A. Tontini, L. Gasparini, N. Massari, and R. Passerone, “SPAD-Based Quantum Random Number Generator With an N^{th} -Order Rank Algorithm on FPGA,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 66, no. 12, pp. 2067–2071, Dec. 2019, doi: 10.1109/TCSII.2019.2909013.
[36] X. Jiang *et al.*, “InP-Based Single-Photon Detectors and Geiger-Mode APD Arrays for Quantum Communications Applications,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 21, no. 3, pp. 5–16, May 2015, doi: 10.1109/JSTQE.2014.2358685.
[37] M. Zarghami *et al.*, “A 32 × 32-Pixel CMOS Imager for Quantum Optics With Per-SPAD TDC, 19.48% Fill-Factor in a 44.64-µm Pitch Reaching 1-MHz Observation Rate,” *IEEE Journal of Solid-State Circuits*, vol. 55, no. 10, pp. 2819–2830, Oct. 2020, doi: 10.1109/JSSC.2020.3005756.
[38] I. Takai, H. Matsubara, M. Soga, M. Ohta, M. Ogawa, and T. Yamashita, “Single-Photon Avalanche Diode with Enhanced NIR-Sensitivity for Automotive LIDAR Systems,” *Sensors*, vol. 16, no. 4, Art. no. 4, Apr. 2016, doi: 10.3390/s16040459.
[39] D. Bronzi, Y. Zou, F. Villa, S. Tisa, A. Tosi, and F. Zappa, “Automotive Three-Dimensional Vision Through a Single-Photon Counting SPAD Camera,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 3, pp. 782–795, Mar. 2016, doi: 10.1109/TITS.2015.2482601.
[40] I. Vornicu, J. M. López-Martínez, F. N. Bandi, R. C. Galán, and Á. Rodríguez-Vázquez, “Design of High-Efficiency SPADs for LiDAR Applications in 110nm CIS Technology,” *IEEE Sensors Journal*, vol. 21, no. 4, pp. 4776–4785, Feb. 2021, doi: 10.1109/JSEN.2020.3032106.
[41] S. Zhuo *et al.*, “Solid-State dToF LiDAR System Using an Eight-Channel Addressable, 20-W/Ch Transmitter, and a 128 × 128 SPAD Receiver With SNR-Based Pixel Binning and Resolution Upscaling,” *IEEE Journal of Solid-State Circuits*, vol. 58, no. 3, pp. 757–770, Mar. 2023, doi: 10.1109/JSSC.2022.3227078.
[42] C. Zhang, S. Lindner, I. M. Antolović, J. Mata Pavia, M. Wolf, and E. Charbon, “A 30-frames/s, 252×144 SPAD Flash LiDAR With 1728 Dual-Clock 48.8-ps TDCs, and Pixel-Wise Integrated Histogramming,” *IEEE Journal of Solid-State Circuits*, vol. 54, no. 4, pp. 1137–1151, Apr. 2019, doi: 10.1109/JSSC.2018.2883720.
[43] J. Zhao *et al.*, “On Analog Silicon Photomultipliers in Standard 55-nm BCD Technology for LiDAR Applications,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 28, no. 5: Lidars and Photonic Radars, pp. 1–10, Sep. 2022, doi: 10.1109/JSTQE.2022.3161089.
[44] E. Sarbazi, M. Safari, and H. Haas, “Statistical Modeling of Single-Photon Avalanche Diode Receivers for Optical Wireless Communications,” *IEEE Transactions on Communications*, vol. 66, no. 9, pp. 4043–4058, Sep. 2018, doi: 10.1109/TCOMM.2018.2822815.
[45] E. Sarbazi, M. Safari, and H. Haas, “The Bit Error Performance and Information Transfer Rate of SPAD Array Optical Receivers,” *IEEE Transactions on Communications*, vol. 68, no. 9, pp. 5689–5705, Sep. 2020, doi: 10.1109/TCOMM.2020.2993374.
[46] S. Huang and M. Safari, “SPAD-Based Optical Wireless Communication With Signal Pre-Distortion and Noise Normalization,” *IEEE Transactions on Communications*, vol. 70, no. 4, pp. 2593–2605, Apr. 2022, doi: 10.1109/TCOMM.2022.3151888.

[47] S. Gnecchi *et al.*, "Analysis of Photon Detection Efficiency and Dynamic Range in SPAD-Based Visible Light Receivers," *Journal of Lightwave Technology*, vol. 34, no. 11, pp. 2774–2781, Jun. 2016, doi: 10.1109/JLT.2016.2550497.
[48] C. Wang, H.-Y. Yu, and Y.-J. Zhu, "A Long Distance Underwater Visible Light Communication System With Single Photon Avalanche Diode," *IEEE Photonics Journal*, vol. 8, no. 5, pp. 1–11, Oct. 2016, doi: 10.1109/JPHOT.2016.2602330.
[49] M. M. Vignetti *et al.*, "Development of a 3D silicon coincidence avalanche detector for charged particle tracking in medical applications," in *2016 IEEE Nuclear Science Symposium, Medical Imaging Conference and Room-Temperature Semiconductor Detector Workshop (NSS/MIC/RTSD)*, Oct. 2016, pp. 1–7. doi: 10.1109/NSSMIC.2016.8069553.
[50] A. Ficorella *et al.*, "Crosstalk Characterization of a Two-Tier Pixelated Avalanche Sensor for Charged Particle Detection," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, no. 2, pp. 1–8, Mar. 2018, doi: 10.1109/JSTQE.2017.2755119.
[51] A. Incoronato, F. Severini, F. Villa, and F. Zappa, "Multi-Channel SPAD Chip for Silicon Photonics With Multi-Photon CoIncidence Detection," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 28, no. 3: Hybrid Integration for Silicon Photonics, pp. 1–7, May 2022, doi: 10.1109/JSTQE.2021.3134071.
[52] W. Shockley and W. T. Read, "Statistics of the Recombinations of Holes and Electrons," *Phys. Rev.*, vol. 87, no. 5, pp. 835–842, Sep. 1952, doi: 10.1103/PhysRev.87.835.
[53] R. N. Hall, "Electron-Hole Recombination in Germanium," *Phys. Rev.*, vol. 87, no. 2, pp. 387–387, Jul. 1952, doi: 10.1103/PhysRev.87.387.
[54] R. J. McIntyre, "On the avalanche initiation probability of avalanche diodes above the breakdown voltage," *IEEE Transactions on Electron Devices*, vol. 20, no. 7, pp. 637–641, Jul. 1973, doi: 10.1109/T-ED.1973.17715.
[55] W. G. Oldham, R. R. Samuelson, and P. Antognetti, "Triggering phenomena in avalanche diodes," *IEEE Transactions on Electron Devices*, vol. 19, no. 9, pp. 1056–1060, Sep. 1972, doi: 10.1109/T-ED.1972.17544.
[56] F. Gramuglia *et al.*, "Engineering Breakdown Probability Profile for PDP and DCR Optimization in a SPAD Fabricated in a Standard 55 nm BCD Process," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 28, no. 2: Optical Detectors, pp. 1–10, Mar. 2022, doi: 10.1109/JSTQE.2021.3114346.
[57] C.-H. Liu, C.-A. Hsien, and S.-D. Lin, "2-D Photon-Detection-Probability Simulation and a Novel Guard-Ring Design for Small CMOS Single-Photon Avalanche Diodes," *IEEE Transactions on Electron Devices*, vol. 69, no. 6, pp. 2873–2878, Jun. 2022, doi: 10.1109/TED.2021.3119264.
[58] Hiwa Mahmoudi, S. S. K. Poushi, B. Steindl, M. Hofbauer, and H. Zimmermann, "Optical and Electrical Characterization and Modeling of Photon Detection Probability in CMOS Single-Photon Avalanche Diodes," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 7572–7580, Mar. 2021, doi: 10.1109/JSEN.2021.3051365.
[59] J. Y. Wu and C.-H. Liu, "Design and Characterization of n/p-well CMOS SPAD With Low Dark Count Rate and High Photon Detection Efficiency," *IEEE Transactions on Electron Devices*, vol. 70, no. 2, pp. 582–587, Feb. 2023, doi: 10.1109/TED.2022.3231575.
[60] D. Shin, B. Park, Y. Chae, and I. Yun, "The Effect of a Deep Virtual Guard Ring on the Device Characteristics of Silicon Single Photon Avalanche Diodes," *IEEE Transactions on Electron Devices*, vol. 66, no. 7, pp. 2986–2991, Jul. 2019, doi: 10.1109/TED.2019.2913714.
[61] E. A. G. Webster, L. A. Grant, and R. K. Henderson, "A High-Performance Single-Photon Avalanche Diode in 130-nm CMOS Imaging Technology," *IEEE Electron Device Letters*, vol. 33, no. 11, pp. 1589–1591, Nov. 2012, doi: 10.1109/LED.2012.2214760.

[62] E. Kizilkan *et al.*, “Guard-Ring-Free InGaAs/InP Single-Photon Avalanche Diode Based on a Novel One-Step Zn-Diffusion Technique,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 28, no. 5: Lidars and Photonic Radars, pp. 1–9, Sep. 2022, doi: 10.1109/JSTQE.2022.3162527.
[63] C. Veerappan and E. Charbon, “CMOS SPAD Based on Photo-Carrier Diffusion Achieving PDP >40% From 440 to 580 nm at 4 V Excess Bias,” *IEEE Photonics Technology Letters*, vol. 27, no. 23, pp. 2445–2448, Dec. 2015, doi: 10.1109/LPT.2015.2468067.
[64] S. Lindner, S. Pellegrini, Y. Henrion, B. Rae, M. Wolf, and E. Charbon, “A High-PDE, Backside-Illuminated SPAD in 65/40-nm 3D IC CMOS Pixel With Cascoded Passive Quenching and Active Recharge,” *IEEE Electron Device Letters*, vol. 38, no. 11, pp. 1547–1550, Nov. 2017, doi: 10.1109/LED.2017.2755989.
[65] S. S. K. Poushi, H. Mahmoudi, M. Hofbauer, B. Steindl, and H. Zimmermann, “Photon detection probability enhancement using an anti-reflection coating in CMOS-based SPADs,” *Appl. Opt., AO*, vol. 60, no. 25, pp. 7815–7820, Sep. 2021, doi: 10.1364/AO.432219.
[66] M. M. Vignetti, F. Calmon, P. Lesieur, and A. Savoy-Navarro, “Simulation study of a novel 3D SPAD pixel in an advanced FD-SOI technology,” *Solid-State Electronics*, vol. 128, pp. 163–171, Feb. 2017, doi: 10.1016/j.sse.2016.10.014.
[67] Y. Liu *et al.*, “Accurate Prediction of Photon Detection Probability Based on the 2-D Dead-Space Model for SPADs,” *IEEE Sensors Journal*, vol. 23, no. 2, pp. 1115–1123, Jan. 2023, doi: 10.1109/JSEN.2022.3226237.
[68] M.-J. Lee, P. Sun, G. Pandraud, C. Bruschini, and E. Charbon, “First Near-Ultraviolet- and Blue-Enhanced Backside-Illuminated Single-Photon Avalanche Diode Based on Standard SOI CMOS Technology,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 25, no. 5, pp. 1–6, Sep. 2019, doi: 10.1109/JSTQE.2019.2918930.
[69] J. Hu *et al.*, “4H-SiC Visible-Blind Single-Photon Avalanche Diode for Ultraviolet Detection at 280 and 350 nm,” *IEEE Transactions on Electron Devices*, vol. 55, no. 8, pp. 1977–1983, Aug. 2008, doi: 10.1109/TED.2008.926669.
[70] I. M. Antolovic, S. Burri, C. Bruschini, R. A. Hoebe, and E. Charbon, “SPAD imagers for super resolution localization microscopy enable analysis of fast fluorophore blinking,” *Sci Rep*, vol. 7, no. 1, Art. no. 1, Mar. 2017, doi: 10.1038/srep44108.
[71] Y. Li, S. Videv, M. Abdallah, K. Qaraqe, M. Uysal, and H. Haas, “Single photon avalanche diode (SPAD) VLC system and application to downhole monitoring,” in *2014 IEEE Global Communications Conference*, Dec. 2014, pp. 2108–2113. doi: 10.1109/GLOCOM.2014.7037119.
[72] E. Fisher, I. Underwood, and R. Henderson, “A reconfigurable 14-bit 60GPhoton/s Single-Photon receiver for visible light communications,” in *2012 Proceedings of the ESSCIRC (ESSCIRC)*, Sep. 2012, pp. 85–88. doi: 10.1109/ESSCIRC.2012.6341262.
[73] Y. Li, M. Safari, R. Henderson, and H. Haas, “Optical OFDM With Single-Photon Avalanche Diode,” *IEEE Photonics Technology Letters*, vol. 27, no. 9, pp. 943–946, May 2015, doi: 10.1109/LPT.2015.2402151.
[74] F. Villa *et al.*, “CMOS SPADs with up to 500 µm diameter and 55% detection efficiency at 420 nm,” *Journal of Modern Optics*, vol. 61, no. 2, pp. 102–115, Jan. 2014, doi: 10.1080/09500340.2013.864425.
[75] S. Pellegrini *et al.*, “Industrialised SPAD in 40 nm technology,” in *2017 IEEE International Electron Devices Meeting (IEDM)*, Dec. 2017, p. 16.5.1-16.5.4. doi: 10.1109/IEDM.2017.8268404.
[76] J. M. Pavia, M. Scandini, S. Lindner, M. Wolf, and E. Charbon, “A 1 × 400 Backside-Illuminated SPAD Sensor With 49.7 ps Resolution, 30 pJ/Sample TDCs Fabricated in 3D CMOS Technology for Near-Infrared Optical Tomography,” *IEEE Journal of Solid-State Circuits*, vol. 50, no. 10, pp. 2406–2418, Oct. 2015, doi: 10.1109/JSSC.2015.2467170.

[77] J. Mata Pavia, M. Wolf, and E. Charbon, “Single-Photon Avalanche Diode Imagers Applied to Near-Infrared Imaging,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 20, no. 6, pp. 291–298, Nov. 2014, doi: 10.1109/JSTQE.2014.2313983.
[78] A. McCarthy *et al.*, “Kilometer-range depth imaging at 1550 nm wavelength using an InGaAs/InP single-photon avalanche diode detector,” *Opt. Express, OE*, vol. 21, no. 19, pp. 22098–22113, Sep. 2013, doi: 10.1364/OE.21.022098.
[79] A. Gulinatti *et al.*, “New silicon SPAD technology for enhanced red-sensitivity, high-resolution timing and system integration,” *Journal of Modern Optics*, vol. 59, no. 17, pp. 1489–1499, Oct. 2012, doi: 10.1080/09500340.2012.701340.
[80] A. Sammak, M. Aminian, L. Qi, W. B. de Boer, E. Charbon, and L. K. Nanver, “A CMOS compatible Ge-on-Si APD operating in proportional and Geiger modes at infrared wavelengths,” in *2011 International Electron Devices Meeting*, Dec. 2011, p. 8.5.1-8.5.4. doi: 10.1109/IEDM.2011.6131515.
[81] A. Dalla Mora *et al.*, “Fast-Gated Single-Photon Avalanche Diode for Wide Dynamic Range Near Infrared Spectroscopy,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 4, pp. 1023–1030, Jul. 2010, doi: 10.1109/JSTQE.2009.2035823.
[82] D. A. Ramirez *et al.*, “Detection efficiencies and generalized breakdown probabilities for nanosecond-gated near infrared single-photon avalanche photodiodes,” *IEEE Journal of Quantum Electronics*, vol. 42, no. 2, pp. 137–145, Feb. 2006, doi: 10.1109/JQE.2005.861627.
[83] M. Liu *et al.*, “High-Performance InGaAs/InP Single-Photon Avalanche Photodiode,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 13, no. 4, pp. 887–894, Jul. 2007, doi: 10.1109/JSTQE.2007.903855.
[84] J. Zhang, R. Thew, J.-D. Gautier, N. Gisin, and H. Zbinden, “Comprehensive Characterization of InGaAs–InP Avalanche Photodiodes at 1550 nm With an Active Quenching ASIC,” *IEEE Journal of Quantum Electronics*, vol. 45, no. 7, pp. 792–799, Jul. 2009, doi: 10.1109/JQE.2009.2013210.
[85] Q. Chen, S. Wu, L. Zhang, W. Fan, and C. S. Tan, “Simulation of High-Efficiency Resonant-Cavity-Enhanced GeSn Single-Photon Avalanche Photodiodes for Sensing and Optical Quantum Applications,” *IEEE Sensors Journal*, vol. 21, no. 13, pp. 14789–14798, Jul. 2021, doi: 10.1109/JSEN.2021.3074407.
[86] S. C. Liew Tat Mun *et al.*, “A Theoretical Comparison of the Breakdown Behavior of $In_{0.52}Al_{0.48}As$ and InP Near-Infrared Single-Photon Avalanche Photodiodes,” *IEEE Journal of Quantum Electronics*, vol. 45, no. 5, pp. 566–571, May 2009, doi: 10.1109/JQE.2009.2013094.
[87] A. Tosi, N. Calandri, M. Sanzaro, and F. Acerbi, “Low-Noise, Low-Jitter, High Detection Efficiency InGaAs/InP Single-Photon Avalanche Diode,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 20, no. 6, pp. 192–197, Nov. 2014, doi: 10.1109/JSTQE.2014.2328440.
[88] M.-J. Lee, P. Sun, and E. Charbon, “A first single-photon avalanche diode fabricated in standard SOI CMOS technology with a full characterization of the device,” *Opt. Express, OE*, vol. 23, no. 10, pp. 13200–13209, May 2015, doi: 10.1364/OE.23.013200.
[89] M.-J. Lee and E. Charbon, “Progress in single-photon avalanche diode image sensors in standard CMOS: From two-dimensional monolithic to three-dimensional-stacked technology,” *Jpn. J. Appl. Phys.*, vol. 57, no. 10, p. 1002A3, Sep. 2018, doi: 10.7567/JJAP.57.1002A3.
[90] D. Rideau *et al.*, “Direct Measurements and Modeling of Avalanche Dynamics and Quenching in SPADs,” in *Essderc 2023, Lisbon, Portugal*, Lisbonne, Portugal, 2023. Accessed: Aug. 17, 2023. [Online]. Available: https://hal.science/hal-04118005
[91] J. D. Petticrew, S. J. Dimler, X. Zhou, A. P. Morrison, C. H. Tan, and J. S. Ng, “Avalanche Breakdown Timing Statistics for Silicon Single Photon Avalanche Diodes,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, no. 2, pp. 1–6, Mar. 2018, doi: 10.1109/JSTQE.2017.2779834.

[92] C. H. Tan, J. S. Ng, G. J. Rees, and J. P. R. David, "Statistics of Avalanche Current Buildup Time in Single-Photon Avalanche Diodes," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 13, no. 4, pp. 906–910, Jul. 2007, doi: 10.1109/JSTQE.2007.903843.
[93] A. Ingargiola, M. Assanelli, A. Gallivanoni, I. Rech, M. Ghioni, and S. Cova, "Avalanche buildup and propagation effects on photon-timing jitter in Si-SPAD with non-uniform electric field," in *Advanced Photon Counting Techniques III*, SPIE, Apr. 2009, pp. 103–114. doi: 10.1117/12.818521.
[94] F. Sun, Y. Xu, Z. Wu, and J. Zhang, "A Simple Analytic Modeling Method for SPAD Timing Jitter Prediction," *IEEE Journal of the Electron Devices Society*, vol. 7, pp. 261–267, 2019, doi: 10.1109/JEDS.2019.2895151.
[95] M. Assanelli, A. Ingargiola, I. Rech, A. Gulinatti, and M. Ghioni, "Photon-Timing Jitter Dependence on Injection Position in Single-Photon Avalanche Diodes," *IEEE Journal of Quantum Electronics*, vol. 47, no. 2, pp. 151–159, Feb. 2011, doi: 10.1109/JQE.2010.2068038.
[96] F. Nolet *et al.*, "Quenching Circuit and SPAD Integrated in CMOS 65 nm with 7.8 ps FWHM Single Photon Timing Resolution," *Instruments*, vol. 2, no. 4, Art. no. 4, Dec. 2018, doi: 10.3390/instruments2040019.
[97] F. Gramuglia, M.-L. Wu, C. Bruschini, M.-J. Lee, and E. Charbon, "A Low-Noise CMOS SPAD Pixel With 12.1 Ps SPTR and 3 Ns Dead Time," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 28, no. 2: Optical Detectors, pp. 1–9, Mar. 2022, doi: 10.1109/JSTQE.2021.3088216.
[98] F. Ceccarelli, G. Acconcia, A. Gulinatti, M. Ghioni, I. Rech, and R. Osellame, "Recent Advances and Future Perspectives of Single-Photon Avalanche Diodes for Quantum Photonics Applications," *Advanced Quantum Technologies*, vol. 4, no. 2, p. 2000102, 2021, doi: 10.1002/qute.202000102.
[99] M. J. Hsu, H. Finkelstein, and S. C. Esener, "A CMOS STI-Bound Single-Photon Avalanche Diode With 27-ps Timing Resolution and a Reduced Diffusion Tail," *IEEE Electron Device Letters*, vol. 30, no. 6, pp. 641–643, Jun. 2009, doi: 10.1109/LED.2009.2019974.
[100] G. Acconcia, M. Ghioni, and I. Rech, "37ps-Precision Time-Resolving Active Quenching Circuit for High-Performance Single Photon Avalanche Diodes," *IEEE Photonics Journal*, vol. 10, no. 6, pp. 1–13, Dec. 2018, doi: 10.1109/JPHOT.2018.2884258.
[101] F. Ceccarelli, G. Acconcia, A. Gulinatti, M. Ghioni, and I. Rech, "83-ps Timing Jitter With a Red-Enhanced SPAD and a Fully Integrated Front End Circuit," *IEEE Photonics Technology Letters*, vol. 30, no. 19, pp. 1727–1730, Oct. 2018, doi: 10.1109/LPT.2018.2867805.
[102] A. Tosi, F. Acerbi, M. Anti, and F. Zappa, "InGaAs/InP Single-Photon Avalanche Diode With Reduced Afterpulsing and Sharp Timing Response With 30 ps Tail," *IEEE Journal of Quantum Electronics*, vol. 48, no. 9, pp. 1227–1232, Sep. 2012, doi: 10.1109/JQE.2012.2208097.
[103] A. Gulinatti, F. Ceccarelli, M. Ghioni, and I. Rech, "Custom silicon technology for SPAD-arrays with red-enhanced sensitivity and low timing jitter," *Opt. Express, OE*, vol. 29, no. 3, pp. 4559–4581, Feb. 2021, doi: 10.1364/OE.413821.
[104] R. H. Haitz, "Mechanisms Contributing to the Noise Pulse Rate of Avalanche Diodes," *Journal of Applied Physics*, vol. 36, no. 10, pp. 3123–3131, Jul. 2004, doi: 10.1063/1.1702936.
[105] E. O. Kane, "Theory of Tunneling," *Journal of Applied Physics*, vol. 32, no. 1, pp. 83–91, Jun. 2004, doi: 10.1063/1.1735965.
[106] A. Panglosse, P. Martin-Gonthier, O. Marcelot, C. Virmontois, O. Saint-Pé, and P. Magnan, "Dark Count Rate Modeling in Single-Photon Avalanche Diodes," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 67, no. 5, pp. 1507–1515, May 2020, doi: 10.1109/TCSI.2020.2971108.
[107] E. A. G. Webster and R. K. Henderson, "A TCAD and Spectroscopy Study of Dark Count Mechanisms in Single-Photon Avalanche Diodes," *IEEE Transactions on Electron Devices*, vol. 60, no. 12, pp. 4014–4019, Dec. 2013, doi: 10.1109/TED.2013.2285163.
[108] E. A. G. Webster, J. A. Richardson, L. A. Grant, D. Renshaw, and R. K. Henderson, "A Single-Photon Avalanche Diode in 90-nm CMOS Imaging Technology With 44% Photon Detection Efficiency

at 690 nm," *IEEE Electron Device Letters*, vol. 33, no. 5, pp. 694–696, May 2012, doi: 10.1109/LED.2012.2187420.

[109] S. Cova, A. Lacaita, and G. Ripamonti, "Trapping phenomena in avalanche photodiodes on nanosecond scale," *IEEE Electron Device Letters*, vol. 12, no. 12, pp. 685–687, Dec. 1991, doi: 10.1109/55.116955.

[110] A. V. Losev *et al.*, "Dead Time Duration and Active Reset Influence on the Afterpulse Probability of InGaAs/InP Single-Photon Avalanche Diodes," *IEEE Journal of Quantum Electronics*, vol. 58, no. 3, pp. 1–11, Jun. 2022, doi: 10.1109/JQE.2022.3171671.

[111] D. P. Palubiak, Z. Li, and M. J. Deen, "Afterpulsing Characteristics of Free-Running and Time-Gated Single-Photon Avalanche Diodes in 130-nm CMOS," *IEEE Transactions on Electron Devices*, vol. 62, no. 11, pp. 3727–3733, Nov. 2015, doi: 10.1109/TED.2015.2475126.

[112] M. Moreno-García, L. Pancheri, M. Perenzoni, R. del Río, Ó. Guerra Vinuesa, and Á. Rodríguez-Vázquez, "Characterization-Based Modeling of Retriggering and Afterpulsing for Passively Quenched CMOS SPADs," *IEEE Sensors Journal*, vol. 19, no. 14, pp. 5700–5709, Jul. 2019, doi: 10.1109/JSEN.2019.2903937.

[113] R. Enne, B. Steindl, M. Hofbauer, and H. Zimmermann, "Fast Cascoded Quenching Circuit for Decreasing Afterpulsing Effects in 0.35- µ m CMOS," *IEEE Solid-State Circuits Letters*, vol. 1, no. 3, pp. 62–65, Mar. 2018, doi: 10.1109/LSSC.2018.2827881.

[114] D. Bronzi, F. Villa, S. Tisa, A. Tosi, and F. Zappa, "SPAD Figures of Merit for Photon-Counting, Photon-Timing, and Imaging Applications: A Review," *IEEE Sensors Journal*, vol. 16, no. 1, pp. 3–12, Jan. 2016, doi: 10.1109/JSEN.2015.2483565.

[115] M. A. Wayne, A. Restelli, J. C. Bienfang, and P. G. Kwiat, "Afterpulse Reduction Through Prompt Quenching in Silicon Reach-Through Single-Photon Avalanche Diodes," *Journal of Lightwave Technology*, vol. 32, no. 21, pp. 4097–4103, Nov. 2014, doi: 10.1109/JLT.2014.2346736.

[116] M. 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," *Sensors*, vol. 21, no. 12, Art. no. 12, Jan. 2021, doi: 10.3390/s21124014.

[117] K. Morimoto and E. Charbon, "A Scaling Law for SPAD Pixel Miniaturization," *Sensors*, vol. 21, no. 10, Art. no. 10, Jan. 2021, doi: 10.3390/s21103447.

[118] J. M. Pavia, M. Wolf, and E. Charbon, "Measurement and modeling of microlenses fabricated on single-photon avalanche diode arrays for fill factor recovery," *Opt. Express, OE*, vol. 22, no. 4, pp. 4202–4213, Feb. 2014, doi: 10.1364/OE.22.004202.

[119] S. Donati, G. Martini, and E. Randone, "Improving Photodetector Performance by Means of Microoptics Concentrators," *Journal of Lightwave Technology*, vol. 29, no. 5, pp. 661–665, Mar. 2011, doi: 10.1109/JLT.2010.2103302.

[120] I. M. Antolovic, S. Burri, C. Bruschini, R. Hoebe, and E. Charbon, "Nonuniformity Analysis of a 65-kpixel CMOS SPAD Imager," *IEEE Transactions on Electron Devices*, vol. 63, no. 1, pp. 57–64, Jan. 2016, doi: 10.1109/TED.2015.2458295.

[121] I. M. Antolovic *et al.*, "Optical-stack optimization for improved SPAD photon detection efficiency," in *Quantum Sensing and Nano Electronics and Photonics XVI*, SPIE, Feb. 2019, pp. 359–365. doi: 10.1117/12.2511301.

[122] I. Gyongy *et al.*, "Cylindrical microlensing for enhanced collection efficiency of small pixel SPAD arrays in single-molecule localisation microscopy," *Opt. Express, OE*, vol. 26, no. 3, pp. 2280–2291, Feb. 2018, doi: 10.1364/OE.26.002280.

[123] G. Intermite *et al.*, "Enhancing the fill-factor of CMOS SPAD arrays using microlens integration," in *Photon Counting Applications 2015*, SPIE, May 2015, pp. 64–75. doi: 10.1117/12.2178950.

[124] A. R. Ximenes, P. Padmanabhan, M.-J. Lee, Y. Yamashita, D. N. Yaung, and E. Charbon, "A 256×256 45/65nm 3D-stacked SPAD-based direct TOF image sensor for LiDAR applications with

optical polar modulation for up to 18.6dB interference suppression," in *2018 IEEE International Solid - State Circuits Conference - (ISSCC)*, Feb. 2018, pp. 96–98. doi: 10.1109/ISSCC.2018.8310201.

[125] S. Burri, Y. Maruyama, X. Michalet, F. Regazzoni, C. Bruschini, and E. Charbon, "Architecture and applications of a high resolution gated SPAD image sensor," *Opt. Express, OE*, vol. 22, no. 14, pp. 17573–17589, Jul. 2014, doi: 10.1364/OE.22.017573.

[126] M.-J. Lee *et al.*, "High-Performance Back-Illuminated Three-Dimensional Stacked Single-Photon Avalanche Diode Implemented in 45-nm CMOS Technology," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, no. 6, pp. 1–9, Nov. 2018, doi: 10.1109/JSTQE.2018.2827669.

[127] R. K. Henderson *et al.*, "5.7 A 256×256 40nm/90nm CMOS 3D-Stacked 120dB Dynamic-Range Reconfigurable Time-Resolved SPAD Imager," in *2019 IEEE International Solid- State Circuits Conference - (ISSCC)*, Feb. 2019, pp. 106–108. doi: 10.1109/ISSCC.2019.8662355.

[128] K. Morimoto *et al.*, "3.2 Megapixel 3D-Stacked Charge Focusing SPAD for Low-Light Imaging and Depth Sensing," in *2021 IEEE International Electron Devices Meeting (IEDM)*, Dec. 2021, p. 20.2.1-20.2.4. doi: 10.1109/IEDM19574.2021.9720605.

[129] F. Gramuglia *et al.*, "CMOS 3D-Stacked FSI Multi-Channel Digital SiPM for Time-of-Flight PET Applications," in *2020 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, Oct. 2020, pp. 1–3. doi: 10.1109/NSS/MIC42677.2020.9507833.

[130] E. Charbon, C. Bruschini, and M.-J. Lee, "3D-Stacked CMOS SPAD Image Sensors: Technology and Applications," in *2018 25th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, Dec. 2018, pp. 1–4. doi: 10.1109/ICECS.2018.8617983.

[131] P. Padmanabhan *et al.*, "7.4 A 256×128 3D-Stacked (45nm) SPAD FLASH LiDAR with 7-Level Coincidence Detection and Progressive Gating for 100m Range and 10klux Background Light," in *2021 IEEE International Solid- State Circuits Conference (ISSCC)*, Feb. 2021, pp. 111–113. doi: 10.1109/ISSCC42613.2021.9366010.

[132] T. Al Abbas, N. A. W. Dutton, O. Almer, S. Pellegrini, Y. Henrion, and R. K. Henderson, "Backside illuminated SPAD image sensor with 7.83µm pitch in 3D-stacked CMOS technology," in *2016 IEEE International Electron Devices Meeting (IEDM)*, Dec. 2016, p. 8.1.1-8.1.4. doi: 10.1109/IEDM.2016.7838372.

[133] S. Zhang *et al.*, "Challenges and recent prospectives of 3D heterogeneous integration," *e-Prime - Advances in Electrical Engineering, Electronics and Energy*, vol. 2, p. 100052, Jan. 2022, doi: 10.1016/j.prime.2022.100052.

[134] W. J. Kindt, H. W. van Zeijl, and S. Middelhoek, "Optical Cross Talk in Geiger Mode Avalanche Photodiode Arrays: Modeling, Prevention and Measurement," in *28th European Solid-State Device Research Conference*, Sep. 1998, pp. 192–195.

[135] I. Rech *et al.*, "Optical crosstalk in single photon avalanche diode arrays: a new complete model," *Opt. Express, OE*, vol. 16, no. 12, pp. 8381–8394, Jun. 2008, doi: 10.1364/OE.16.008381.

[136] C. Veerappan and E. Charbon, "A Substrate Isolated CMOS SPAD Enabling Wide Spectral Response and Low Electrical Crosstalk," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 20, no. 6, pp. 299–305, Nov. 2014, doi: 10.1109/JSTQE.2014.2318436.

[137] D.-R. Wu, C.-M. Tsai, Y.-H. Huang, and S.-D. Lin, "Crosstalk Between Single-Photon Avalanche Diodes in a 0.18-µm High-Voltage CMOS Process," *Journal of Lightwave Technology*, vol. 36, no. 3, pp. 833–837, Feb. 2018, doi: 10.1109/JLT.2017.2779490.

[138] A. T. Erdogan *et al.*, "A High Dynamic Range 128 × 120 3-D Stacked CMOS SPAD Image Sensor SoC for Fluorescence Microendoscopy," *IEEE Journal of Solid-State Circuits*, vol. 57, no. 6, pp. 1649–1660, Jun. 2022, doi: 10.1109/JSSC.2022.3150721.

[139] A. Katz *et al.*, "Passive CMOS Single Photon Avalanche Diode Imager for a Gun Muzzle Flash Detection System," *IEEE Sensors Journal*, vol. 19, no. 14, pp. 5851–5858, Jul. 2019, doi: 10.1109/JSEN.2019.2904312.

[140] M. Liu, C. Hu, J. C. Campbell, Z. Pan, and M. M. Tashima, "Reduce Afterpulsing of Single Photon Avalanche Diodes Using Passive Quenching With Active Reset," *IEEE Journal of Quantum Electronics*, vol. 44, no. 5, pp. 430–434, May 2008, doi: 10.1109/JQE.2007.916688.

[141] D. Palubiak, M. M. El-Desouki, O. Marinov, M. J. Deen, and Q. Fang, "High-Speed, Single-Photon Avalanche-Photodiode Imager for Biomedical Applications," *IEEE Sensors Journal*, vol. 11, no. 10, pp. 2401–2412, Oct. 2011, doi: 10.1109/JSEN.2011.2123090.

[142] F. Acerbi, A. D. Frera, A. Tosi, and F. Zappa, "Fast Active Quenching Circuit for Reducing Avalanche Charge and Afterpulsing in InGaAs/InP Single-Photon Avalanche Diode," *IEEE Journal of Quantum Electronics*, vol. 49, no. 7, pp. 563–569, Jul. 2013, doi: 10.1109/JQE.2013.2260726.

[143] A. Ruggeri, P. Ciccarella, F. Villa, F. Zappa, and A. Tosi, "Integrated Circuit for Subnanosecond Gating of InGaAs/InP SPAD," *IEEE Journal of Quantum Electronics*, vol. 51, no. 7, pp. 1–7, Jul. 2015, doi: 10.1109/JQE.2015.2438436.

[144] M. Renna, A. Ruggeri, M. Sanzaro, F. Villa, F. Zappa, and A. Tosi, "High Detection Rate Fast-Gated CMOS Single-Photon Avalanche Diode Module," *IEEE Photonics Journal*, vol. 12, no. 5, pp. 1–12, Oct. 2020, doi: 10.1109/JPHOT.2020.3017092.

[145] D. Bronzi, S. Tisa, F. Villa, S. Bellisai, A. Tosi, and F. Zappa, "Fast Sensing and Quenching of CMOS SPADs for Minimal Afterpulsing Effects," *IEEE Photonics Technology Letters*, vol. 25, no. 8, pp. 776–779, Apr. 2013, doi: 10.1109/LPT.2013.2251621.

[146] Y. Xu, J. Lu, and Z. Wu, "A Compact High-Speed Active Quenching and Recharging Circuit for SPAD Detectors," *IEEE Photonics Journal*, vol. 12, no. 5, pp. 1–8, Oct. 2020, doi: 10.1109/JPHOT.2020.3015872.

[147] F. Severini, I. Cusini, D. Berretta, K. Pasquinelli, A. Incoronato, and F. Villa, "SPAD Pixel With Sub-NS Dead-Time for High-Count Rate Applications," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 28, no. 2: Optical Detectors, pp. 1–8, Mar. 2022, doi: 10.1109/JSTQE.2021.3124825.

[148] A. Dervić, M. Hofbauer, B. Goll, and H. Zimmermann, "Integrated Fast-Sensing Triple-Voltage SPAD Quenching/Resetting Circuit for Increasing PDP," *IEEE Photonics Technology Letters*, vol. 33, no. 3, pp. 139–142, Feb. 2021, doi: 10.1109/LPT.2020.3044484.

[149] A. Dervić, M. Hofbauer, B. Goll, and H. Zimmermann, "High Slew-Rate Quadruple-Voltage Mixed-Quenching Active-Resetting Circuit for SPADs in 0.35-µm CMOS for Increasing PDP," *IEEE Solid-State Circuits Letters*, vol. 4, pp. 18–21, 2021, doi: 10.1109/LSSC.2020.3044866.

[150] B. Goll, B. Steindl, and H. Zimmermann, "Cascoded Active Quencher for SPADs With Bipolar Differential Amplifier in 0.35 µm BiCMOS," *IEEE Photonics Journal*, vol. 14, no. 2, pp. 1–8, Apr. 2022, doi: 10.1109/JPHOT.2022.3149719.

[151] A. Cathelin, "Fully Depleted Silicon on Insulator Devices CMOS: The 28-nm Node Is the Perfect Technology for Analog, RF, mmW, and Mixed-Signal System-on-Chip Integration," *IEEE Solid-State Circuits Magazine*, vol. 9, no. 4, pp. 18–26, Fall 2017, doi: 10.1109/MSSC.2017.2745738.

[152] J.-P. Noel *et al.*, "Multi V_T UTBB FDSOI Device Architectures for Low-Power CMOS Circuit," *IEEE Transactions on Electron Devices*, vol. 58, no. 8, pp. 2473–2482, Aug. 2011, doi: 10.1109/TED.2011.2155658.

[153] T. Poiroux *et al.*, "Leti-UTSOI2.1: A Compact Model for UTBB-FDSOI Technologies—Part I: Interface Potentials Analytical Model," *IEEE Transactions on Electron Devices*, vol. 62, no. 9, pp. 2751–2759, Sep. 2015, doi: 10.1109/TED.2015.2458339.

[154] T. C. de Albuquerque *et al.*, "Integration of SPAD in 28nm FDSOI CMOS technology," in *2018 48th European Solid-State Device Research Conference (ESSDERC)*, Sep. 2018, pp. 82–85. doi: 10.1109/ESSDERC.2018.8486852.

[155] D. Issartel *et al.*, "Architecture optimization of SPAD integrated in 28 nm FD-SOI CMOS technology to reduce the DCR," *Solid-State Electronics*, vol. 191, p. 108297, May 2022, doi: 10.1016/j.sse.2022.108297.

[156] D. Issartel *et al.*, “SPAD FDSOI cell optimization for lower dark count rate achievement,” in *2020 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS)*, Sep. 2020, pp. 1–5. doi: 10.1109/EUROSOI-ULIS49407.2020.9365292.

[157] T. C. de Albuquerque, D. Issartel, R. Clerc, P. Pittet, R. Cellier, and F. Calmon, “Lowering the Dark Count Rate of SPAD Implemented in CMOS FDSOI Technology,” in *2019 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS)*, Apr. 2019, pp. 1–4. doi: 10.1109/EUROSOI-ULIS45800.2019.9041916.

[158] S. Gao *et al.*, “Improving the Photon Detection Probability of SPAD implemented in FD-SOI CMOS Technology with light-trapping concept,” in *2021 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EuroSOI-ULIS)*, Sep. 2021, pp. 1–4. doi: 10.1109/EuroSOI-ULIS53016.2021.9560684.

[159] S. Gao *et al.*, “3D Electro-optical Simulations for Improving the Photon Detection Probability of SPAD Implemented in FD-SOI CMOS Technology,” in *2021 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, Sep. 2021, pp. 301–304. doi: 10.1109/SISPAD54002.2021.9592555.

[160] T. C. de Albuquerque *et al.*, “Body-biasing considerations with SPAD FDSOI: advantages and drawbacks,” in *ESSDERC 2019 - 49th European Solid-State Device Research Conference (ESSDERC)*, Sep. 2019, pp. 210–213. doi: 10.1109/ESSDERC.2019.8901825.

[161] T. Chaves de Albuquerque *et al.*, “Indirect avalanche event detection of Single Photon Avalanche Diode implemented in CMOS FDSOI technology,” *Solid-State Electronics*, vol. 163, p. 107636, Jan. 2020, doi: 10.1016/j.sse.2019.107636.

[162] W. Bae, “CMOS Inverter as Analog Circuit: An Overview,” *Journal of Low Power Electronics and Applications*, vol. 9, no. 3, Art. no. 3, Sep. 2019, doi: 10.3390/jlpea9030026.

[163] M. Liu *et al.*, “A 16-Channel Analog CMOS SiPM With On-Chip Front-End for D-ToF LiDAR,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 69, no. 6, pp. 2376–2386, Jun. 2022, doi: 10.1109/TCSI.2022.3160265.

[164] X. Li, D. Li, C. Tang, and S. Liu, “Fully-Integrated SPAD-Based Receiver With Nanosecond Dead Time for Optical Wireless Communication,” *Journal of Lightwave Technology*, vol. 41, no. 2, pp. 653–661, Jan. 2023, doi: 10.1109/JLT.2022.3217075.

[165] J. M. Rabaey, A. P. Chandrakasan, and B. Nikolić, *Digital Integrated Circuits: A Design Perspective*. Pearson Education, 2003.

[166] J.-P. Colinge, *Silicon-on-Insulator Technology: Materials to VLSI*. Boston, MA: Springer US, 2004. doi: 10.1007/978-1-4419-9106-5.

[167] A. Dalla Mora, A. Tosi, S. Tisa, and F. Zappa, “Single-Photon Avalanche Diode Model for Circuit Simulations,” *IEEE Photonics Technology Letters*, vol. 19, no. 23, pp. 1922–1924, Dec. 2007, doi: 10.1109/LPT.2007.908768.

[168] A. Cathelin, “Fully Depleted Silicon on Insulator Devices CMOS: The 28-nm Node Is the Perfect Technology for Analog, RF, mmW, and Mixed-Signal System-on-Chip Integration,” *IEEE Solid-State Circuits Magazine*, vol. 9, no. 4, pp. 18–26, 2017, doi: 10.1109/MSSC.2017.2745738.

References

Introduction

L'objectif de ce travail est de concevoir et de mettre en œuvre des circuits de détection d'avalanche et d'extinction active (AQ) pour la première diode à avalanche à photon unique (SPAD) réalisée dans une technologie CMOS de silicium sur isolant entièrement appauvri (FD-SOI) qui améliore les paramètres de performance de la SPAD tels que l'effet d'impulsion et l'efficacité de la détection de photons (PDE). Premier du genre, le SPAD utilisé dans ce travail est un dispositif non optimisé. Un circuit de détection d'avalanche sensible peut également aider à caractériser le SPAD afin d'identifier les optimisations nécessaires.

Chapitre 1 : Circuits SPAD et Quenching

Le SPAD est une jonction pn, polarisée inversement au-delà de sa tension de claquage (dans la région de claquage). Il peut convertir un photon unique en une avalanche mesurable de porteurs de charge (courant d'avalanche) dans sa région de charge d'espace (SCR). Le processus d'avalanche est illustré à la figure 1. Le processus d'arrêt du courant d'avalanche est appelé extinction, qui peut être réalisée par une résistance (extinction passive (PQ)) ou par des dispositifs actifs (extinction active (AQ)). La figure 2 montre la caractéristique I-V et la dynamique d'avalanche et d'extinction d'un SPAD. Le SPAD souffre d'une source de bruit appelée afterpulsing. Au cours d'une avalanche, certains porteurs peuvent être piégés par les centres de piégeage du semi-conducteur. Si ces porteurs sont libérés lorsque le SPAD est réinitialisé dans la région de claquage, ils peuvent déclencher une avalanche parasite appelée "afterpulse". Ainsi, pour réduire ce bruit, le temps

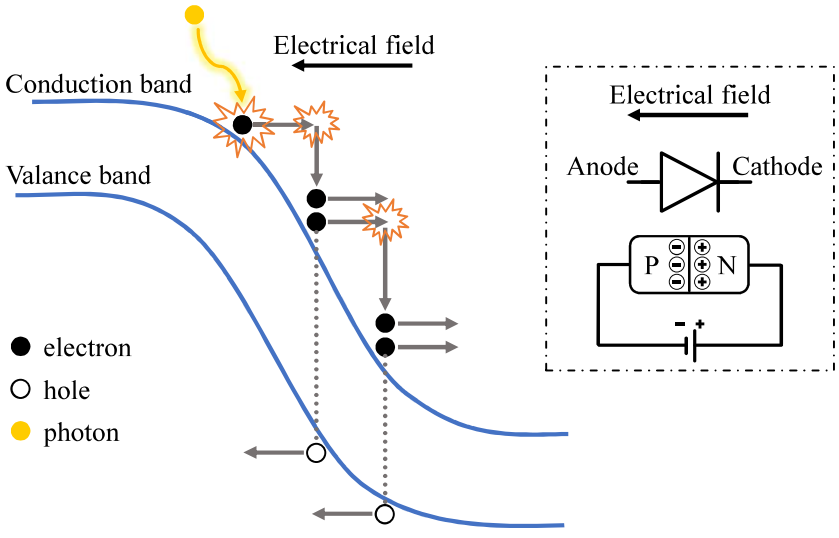


Fig. 1 Processus d'avalanche dans une jonction PN polarisée en sens inverse, dû au phénomène photoélectrique. Encadré : une diode en polarisation inverse avec son symbole électrique.

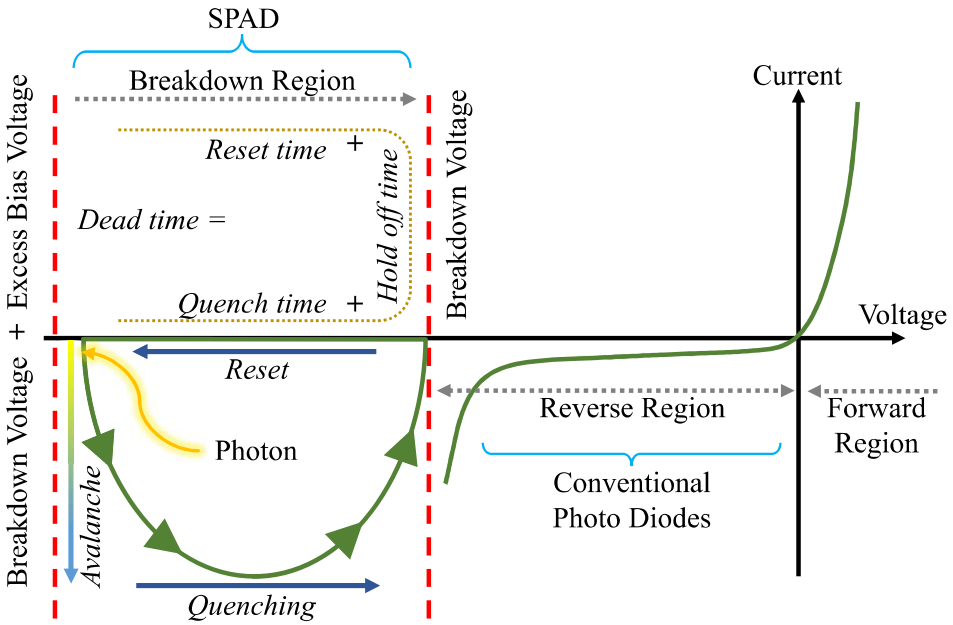


Fig. 2 Caractéristique I-V symbolique d'un SPAD avec dynamique quench-reset.

pendant lequel le SPAD est hors de sa zone de claquage (temps mort) doit être prolongé. Un temps mort long limite le taux de comptage de photons maximum réalisable d'un pixel SPAD. Les circuits AQ sont capables de réduire l'effet d'impulsion en éteignant le SPAD plus rapidement, ce qui réduit la charge d'avalanche sans détériorer le taux de comptage maximal du pixel. Toutefois, les circuits nécessaires pour mettre en œuvre une telle stratégie sont gourmands en surface et dégradent donc le facteur de remplissage (le rapport entre la surface active du pixel et sa surface totale) et la PDE (probabilité de détection de photons (PDP) multipliée par le facteur de remplissage). Il existe donc des compromis entre l'effet d'afterpulsing, le taux de comptage maximal de photons, le facteur de remplissage et la PDE. Les pixels 3D empilés sont une solution pour atténuer ces compromis, mais ils sont coûteux car ils nécessitent deux puces séparées. Par conséquent, l'un des objectifs de cette thèse est de trouver une solution ultime pour réduire le bruit SPAD tout en maintenant le taux de comptage maximal, le facteur de remplissage et l'EDP aussi élevés que possible d'une manière plus rentable. Récemment, l'idée d'un pixel SPAD 3D monolithique dans la technologie CMOS FD-SOI a été présentée dans la littérature. Dans cette technologie, une couche d'oxyde enfouie (BOX) isole les dispositifs de leurs puits correspondants, ce qui permet une intégration monolithique du SPAD et des dispositifs placés au-dessus du SPAD sur la couche BOX. Les auteurs ont montré qu'il est possible de détecter l'avalanche à travers un diviseur de tension au-dessus du SPAD grâce à l'effet de polarisation du body. Cependant, aucun circuit d'extinction n'a encore été intégré dans cette structure 3D monolithique. La figure 3 présente une coupe transversale de la technologie CMOS

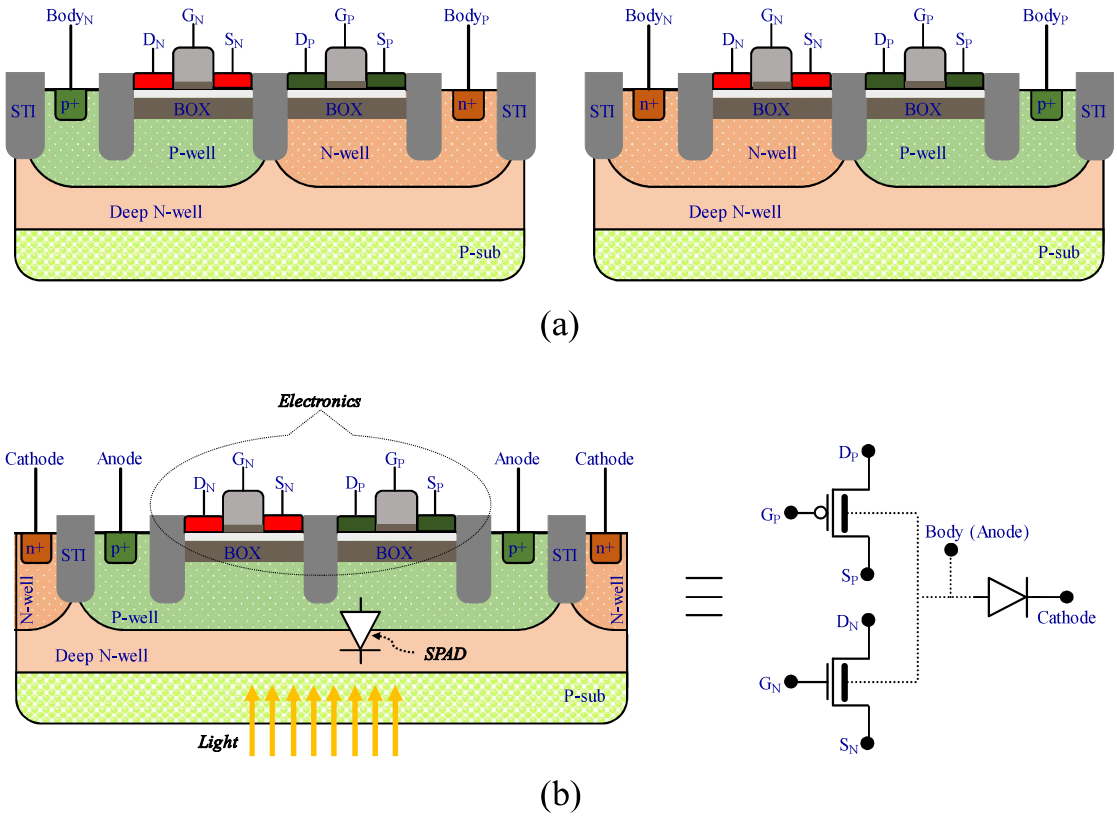


Fig. 3 Coupe transversale symbolique de : (a) la technologie CMOS FD-SOI, puits régulier (à gauche) et puits retourné (à droite), (b) mélange de puits retourné et de puits régulier dans la technologie CMOS FD-SOI et formation du SPAD à la jonction du puits P et du puits N profond, résultant en un pixel 3D intrinsèque (à gauche) et son équivalent schématique (à droite). Le pixel est censé être rétroéclairé.

FD-SOI et du pixel 3D monolithique. L'intégration d'un circuit d'extinction dans cette structure 3D monolithique est un autre objectif de ce travail qui s'aligne sur le premier objectif (atténuer les compromis mentionnés). Le fait de placer un circuit AQ fonctionnel et efficace au-dessus de la zone active du SPAD dans une configuration d'illumination par l'arrière (BSI) peut fortement réduire la dépendance du facteur de remplissage par rapport à l'électronique du pixel et donc atténuer les compromis mentionnés.

Chapitre 2 : Circuit de réinitialisation active à extinction basé sur un inverseur polarisé par son substrat

Plus l'avalanche est détectée tôt, plus vite elle peut être éteinte, ce qui réduit la charge d'avalanche. Ainsi, l'efficacité d'un circuit AQ dans la réduction de la post-impulsion dépend fortement du temps de détection de l'avalanche. C'est pourquoi nous proposons d'utiliser un

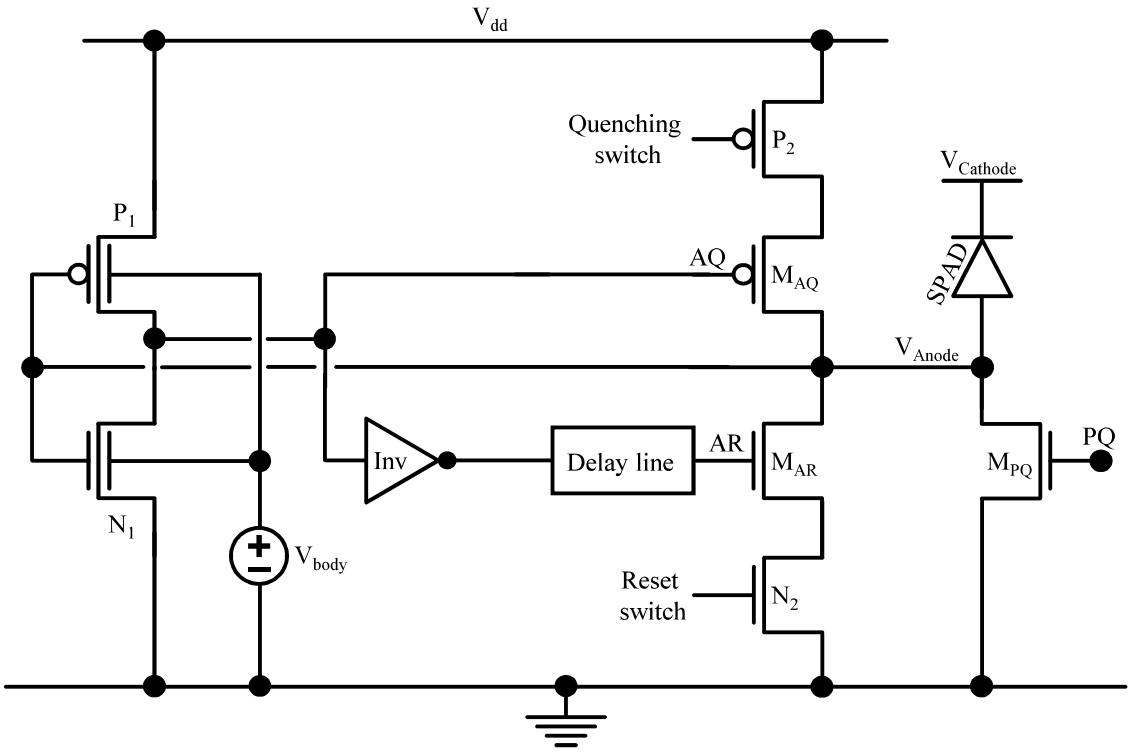


Fig. 4 Schema du circuit BBI AQAR (Body Biased Inverter-based Active Quenching Active Reset) proposé.

inverseur à bas seuil comme circuit de détection. La sensibilité d'un inverseur est augmentée en abaissant son seuil de commutation par un dimensionnement approprié. Nous proposons de combiner la méthode de dimensionnement avec la technique de polarisation du substrat pour augmenter encore la sensibilité de l'inverseur et raccourcir le temps de détection. Dans la technologie CMOS FD-SOI, la couche BOX empêche la formation de diodes parasites entre le drain/source et le substrat des MOSFET. Cela permet d'appliquer une tension de substrat jusqu'à la tension maximale supportable de l'oxyde de grille arrière (BOX), qui est d'environ 3V dans la technologie utilisée. La tension de substrat permet de moduler la tension de seuil des MOSFET. Le taux de cette variation est défini comme le facteur de substrat qui est d'environ 85 mV/V dans cette technologie. Cela signifie qu'en appliquant une tension de substrat de 3 V, la tension de seuil varie de 255 mV. Dans un inverseur, la diminution et l'augmentation des tensions de seuil du NMOS et du PMOS respectivement se traduisent par une diminution du seuil de commutation d'une valeur presque identique. En combinaison avec un dimensionnement approprié, l'inverseur à polarisation par le substrat proposé atteint un seuil de commutation très bas. La figure 4 présente le schéma du circuit AQAR proposé basé sur un inverseur à polarisation corporelle (BBI). Le chronogramme de fonctionnement du circuit est illustré à la figure 5. À l'arrivée du photon, le processus d'avalanche commence. Le courant d'avalanche généré décharge la capacité de jonction du SPAD et augmente

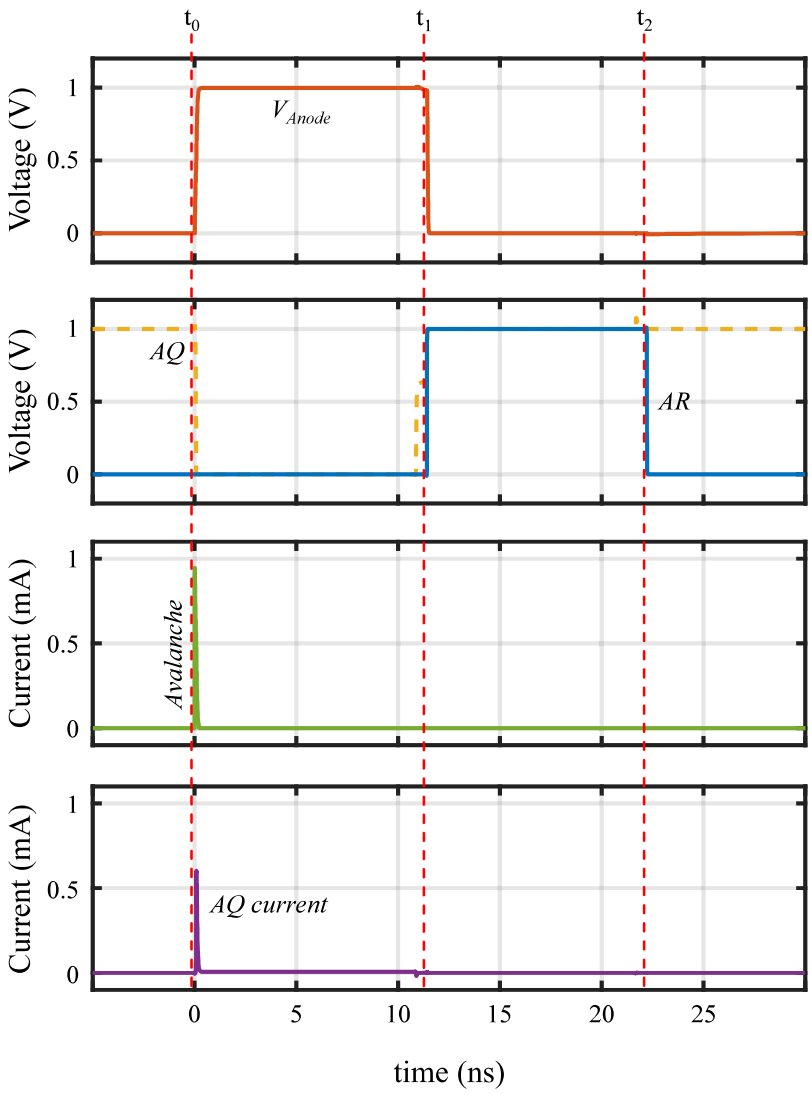


Fig. 5 Diagramme temporel du fonctionnement du circuit AQAR du BBI. De haut en bas : tension d'anode, signaux d'extinction active et de réinitialisation active, courant d'avalanche et courant d'extinction active. L'avalanche démarre à t_0 . À t_1 , le temps de maintien est terminé et la phase de réinitialisation commence. À t_2 , la phase de réinitialisation est terminée et le pixel est prêt pour un nouvel événement.

ainsi la tension de l'anode (t_0). Peu après, cette augmentation de la tension anodique est détectée par l'inverseur polarisé par le substrat. Pendant ce temps, le signal AQ chute et le transistor AQ M_{AQ} commence à conduire et à générer un courant élevé et rapide pour contribuer au processus d'extinction. Après l'extinction, la tension de l'anode reste élevée jusqu'à la fin du temps de maintien déterminé (t_1). À ce moment, le signal AR augmente et donc M_{AR} commence la phase de réinitialisation en chargeant la capacité de jonction du SPAD. Par conséquent, la tension de l'anode diminue. Le signal AR reste élevé jusqu'au temps de réinitialisation déterminé (t_2). Le signal AR tombe à zéro pour éteindre M_{AR}. À partir de maintenant, le pixel est prêt pour une nouvelle

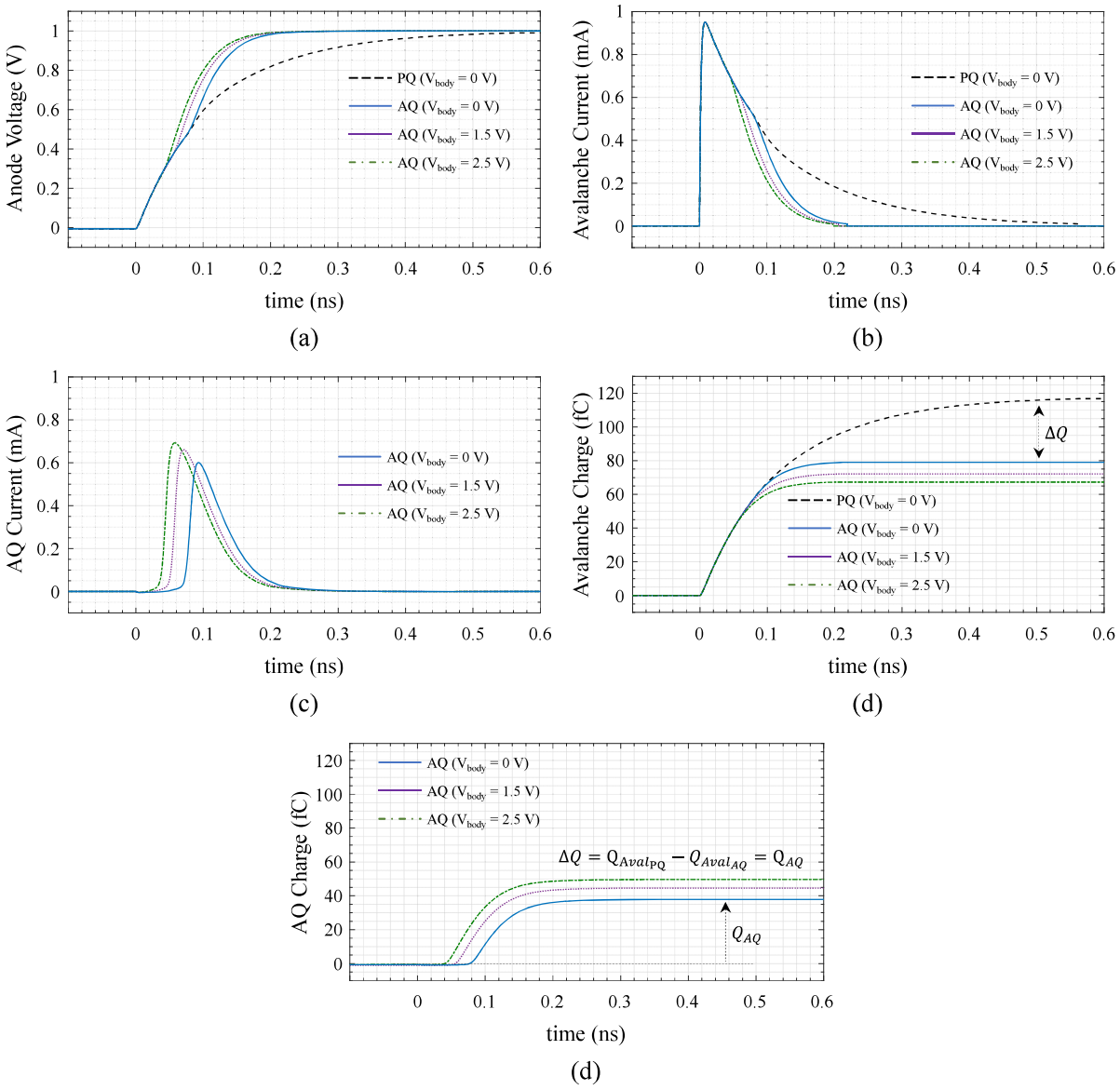


Fig. 6 Résultats de simulation lors d'une avalanche en modes PQ et AQ pour différentes tensions de substrats. (a) tension anodique. (b) courant d'avalanche. (c) courant d'extinction actif. (d) charge d'avalanche. (d) charge contribuant à l'extinction active.

détection de photon. D'après les résultats de la simulation après layout (voir la figure 6), avec une tension de substrat de 2,5 V, l'avalanche est détectée en moins de 40 ps et éteinte activement en seulement 150 ps, ce qui se traduit par une réduction relative de la charge d'avalanche d'environ 42 % en mode AQ par rapport au mode PQ pur. Le tableau 1 présente tous les résultats de la simulation post-layout pour le circuit BBI AQAR. La figure 7 montre les variations mesurées de la tension anodique pendant l'avalanche. Le circuit AQAR proposé réduit le temps d'extinction de 500 ps (la moitié du temps d'extinction en mode PQ). Les mesures de afterpulsing montrent une

TABLEAU 1

RESULTATS DE SIMULATION POUR LE CIRCUIT BBI AQAR DANS LES MODES PQ ET AQ A DIFFERENTES TENSIONS DU SUBSTRATS.

	V_{body} = 0 V	V_{body} = 1,5 V	V_{body} = 2,5 V
Temps de détection de l'avalanche (ps)	< 80	< 60	< 40
Temps de trempe en PQ (ps)	380	380	380
Temps de trempe en AQ (ps)	170	160	150
Charge d'avalanche dans PQ (fC)	117	117	117
Charge d'avalanche en AQ (fC)	79	72	67
Contribution à la charge de QA	38	45	50

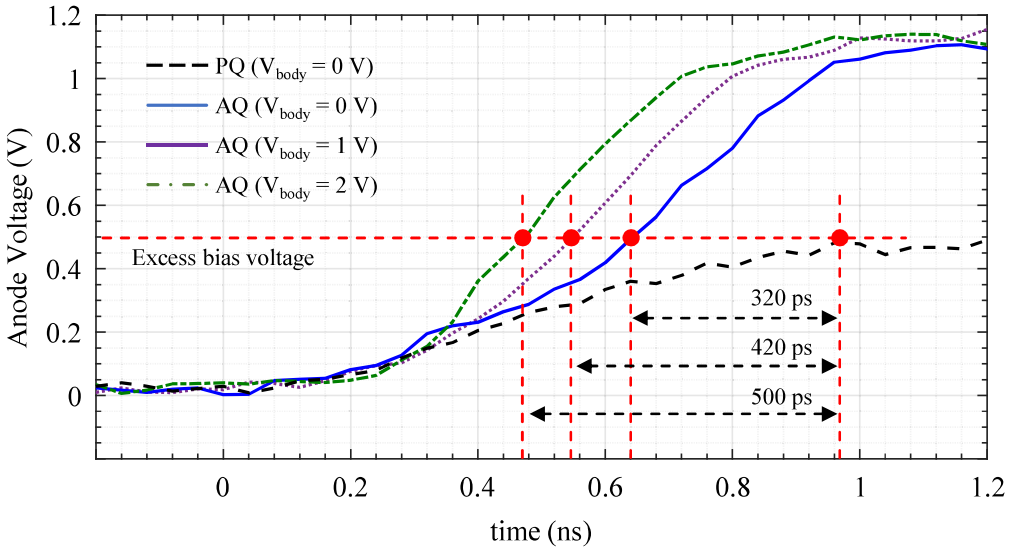


Fig. 7 Variations mesurées de la tension anodique pendant une avalanche en modes PQ et AQ à 40 °C avec 500 mV de tension de polarisation excédentaire à différentes tensions de substrats.

réduction de plus de 50 % du pourcentage de probabilité de afterpulsing (APP) en mode AQ avec une tension de substrat de 2 V par rapport au mode PQ, avec un temps mort aussi court que 5 ns. Ainsi, le circuit AQAR proposé, basé sur un inverseur polarisé par le substrat, est capable d'atténuer le compromis entre le bruit de rémanence et le taux de comptage maximal. La figure 8 présente les valeurs APP pour des temps d'arrêt de 5 ns, 10 ns, 20 ns et 50 ns, alors que la tension du substrat est balayée de 0 V à 2 V, afin d'étudier l'efficacité du circuit proposé dans la réduction du bruit de rémanence par rapport au mode PQ.

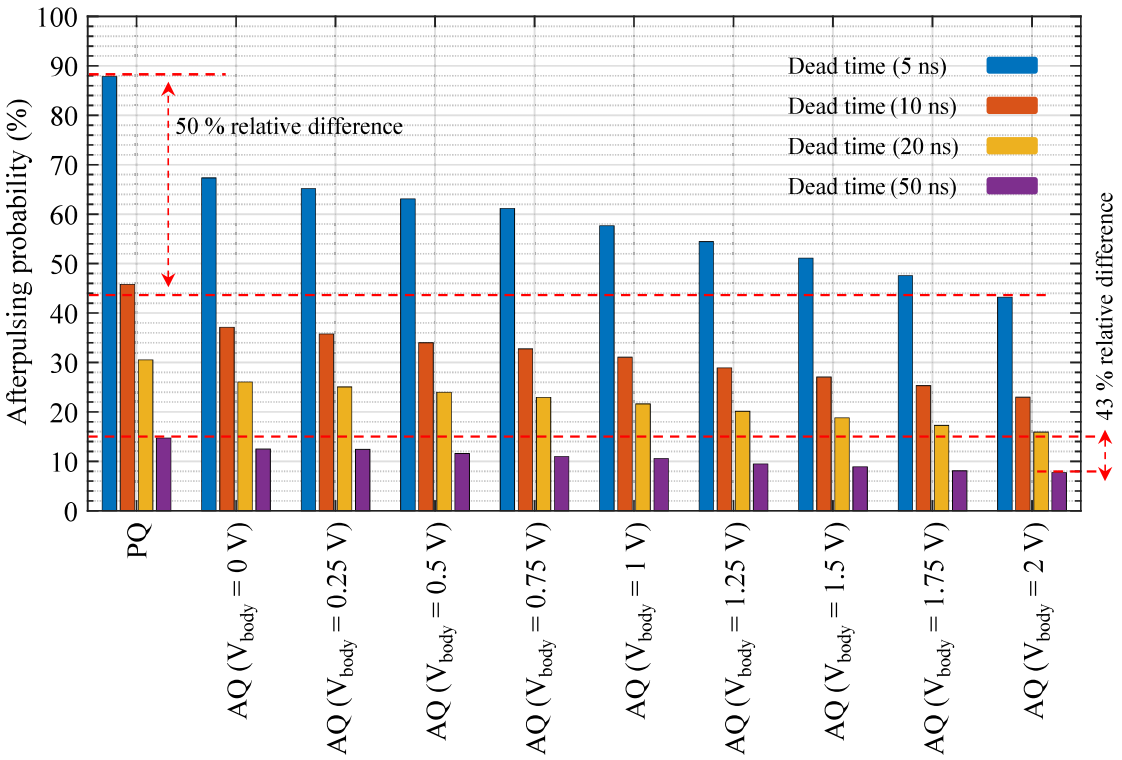


Fig. 8 Probabilité d'impulsions secondaires dans les modes PQ et AQ à différentes tensions de substrat pour l'arrêt. de 5, 10, 20 et 50 ns. La tension de polarisation est de 400 mV et la température est de 40°C.

Chapitre 3 : Inverseur “chatouilleux” et pixels 3D monolithiques

Dans ce qui suit, la thèse se concentre sur la poursuite de deux objectifs en même temps : Concevoir un circuit AQ encore plus rapide qui puisse être intégré à la structure 3D monolithique. Tant qu'un inverseur fonctionne comme une porte numérique, il est limité par son temps de propagation. L'inverseur polarisé par le substrat diminue fortement ce délai, mais ne peut pas éliminer ce défaut complétement. Nous proposons de polariser l'inverseur à son seuil de commutation pour rendre ce délai presque nul. Dans ce cas, les variations minimales de tension à l'entrée de l'inverseur multipliées par le gain maximal de l'inverseur modifient sa sortie. Pour obtenir un pixel 3D monolithique, ce pré- polarisation doit être réalisée de manière à être immunisé contre les variations de tension du substrat des transistors sur le SPAD. En fait, le substrat des MOSFETs dans cette structure est le nœud dynamique du SPAD (anode). Une avalanche augmente la tension de l'anode du SPAD ou, en d'autres termes, la tension du substrat des MOSFETs. Par conséquent, une avalanche impose une polarisation de substrat non désirée aux transistors situés

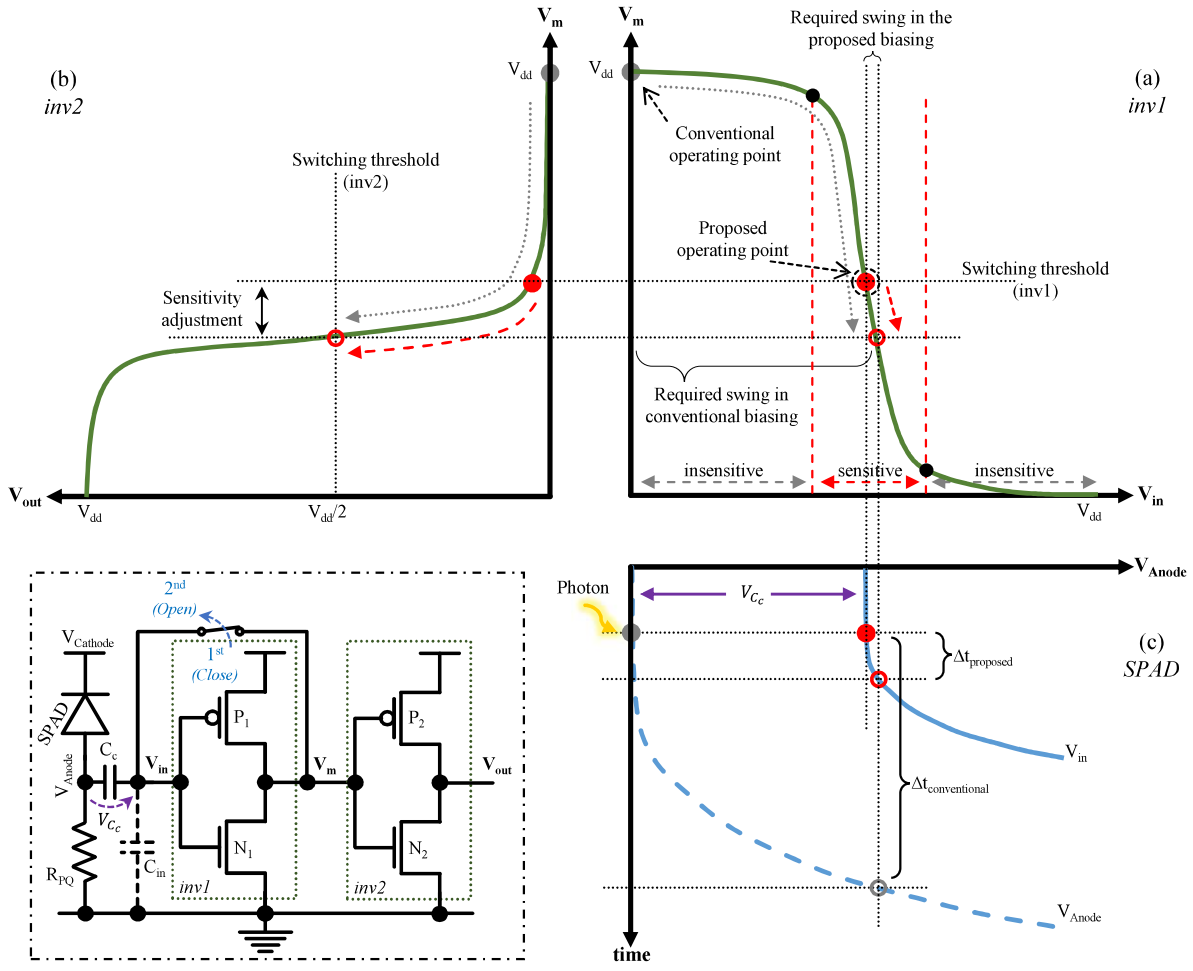


Fig. 9 Principes de fonctionnement du détecteur d'avalanche proposé. (a) VTC d'un inverseur typique (inv1) qui est polarisé à son seuil de commutation (b) VTC de l'inverseur suivant (inv2) avec un seuil de commutation plus bas que son prédécesseur (c) Tension d'entrée du circuit de détection et tension transitoire de l'anode pendant une avalanche. Encadré : Schéma du détecteur proposé. L'interrupteur parallèle avec inv1 est d'abord fermé (1st), puis il devient ouvert (2nd).

au-dessus du SPAD. Pour résoudre ce problème, une porte de transmission (T-gate) est placée en parallèle avec l'inverseur de détection. La porte de transmission (T-gate) polarise l'inverseur à son seuil de commutation en court-circuitant son entrée et sa sortie à n'importe quelle tension de substrat de son transistor. La figure 9 illustre le concept mentionné. Le choix d'un inverseur comme circuit de détection est également un avantage pour le circuit AQ dans la structure 3D monolithique. En effet, comme nous l'avons expliqué, l'application de la tension de substrat à un inverseur module son seuil de commutation. Dans ce cas, l'augmentation de la tension anodique diminue le seuil de commutation de l'inverseur de détection et entraîne une détection encore plus rapide et donc une extinction plus rapide. Sur la base de ce circuit de détection, trois circuits de détection d'avalanche AQAR (Active Quenching-Active Reset) sont présentés à la figure 10 : dans le circuit Direct

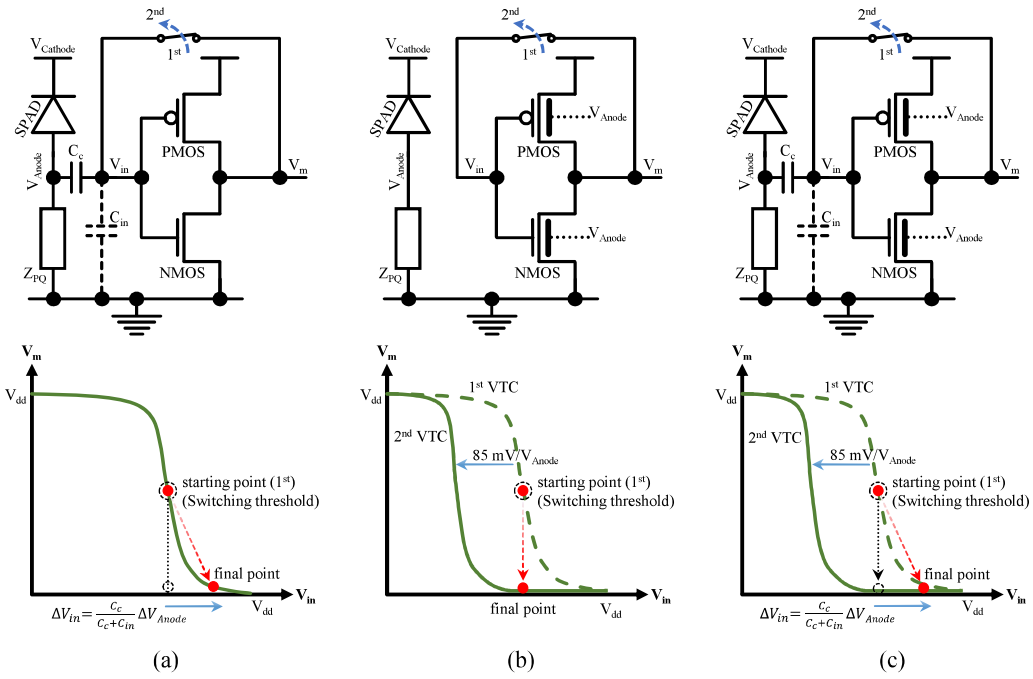


Fig. 10 Schéma (en haut) et VTC (en bas) de : (a) Circuit de détection directe (b) Circuit de détection indirecte (c) Circuit de détection mixte.

Sensing (DS), l'électronique n'est pas au-dessus du SPAD et il est conçu pour prouver la fonctionnalité du circuit AQ proposé dans les pixels standard. Il est appelé DS car l'avalanche est détectée à travers la grille des MOSFET de l'inverseur de détection. Les deux autres circuits peuvent être intégrés dans la structure 3D monolithique. Le premier est le Mixed Sensing (MS) dans lequel l'avalanche est détectée simultanément par la grille et le substrat des MOSFET de l'inverseur. Un autre circuit de détection ne détecte l'avalanche qu'à travers le substrat de ses transistors, appelé pixel à détection indirecte (InDS). Dans ce pixel, il est possible d'appliquer une tension de polarisation excédentaire trois fois supérieure à la marge de tension technologique, ce qui améliore le PDP et la gigue du pixel.

Nous proposons maintenant 3 AQAR utilisant les 3 circuits de détection. Leurs schémas sont présentés à la figure 11. Le circuit DS AQAR doit être placé à côté du SPAD, tandis que les deux autres doivent être placés au-dessus de la zone active du SPAD. La figure 11 (a) montre le schéma du circuit DS AQAR. L'inverseur de détection est composé de P_1, N_1 et est polarisé par la porte T. Les inverseurs 2 et 3 (*inv2* et *inv3)* sont des inverseurs conçus sur mesure, qui conditionnent le signal de détection pour piloter le PMOS large AQ, M_{AQ}. Il convient de noter que l'*inv2* doit avoir un seuil de commutation inférieur au seuil de l'*inv1*. Une ligne à retard accordable de l'extérieur contrôle les temps de maintien et de réinitialisation du SPAD en pilotant la porte T et le NMOS de réinitialisation active (AR) M_{AR}. M_{PQ} met en œuvre une résistance d'extinction passive (PQ)

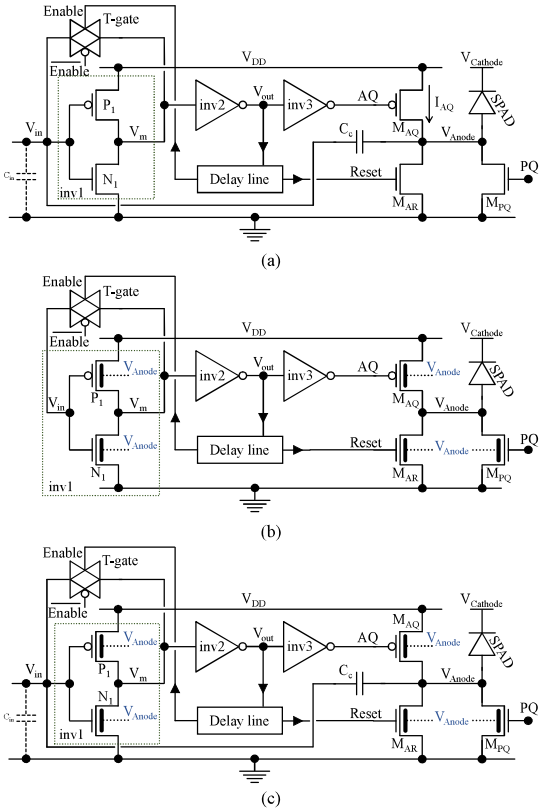


Fig. 11 Schémas des circuits AQAR proposés : (a) détection directe (b) détection indirecte (c) détection mixte.

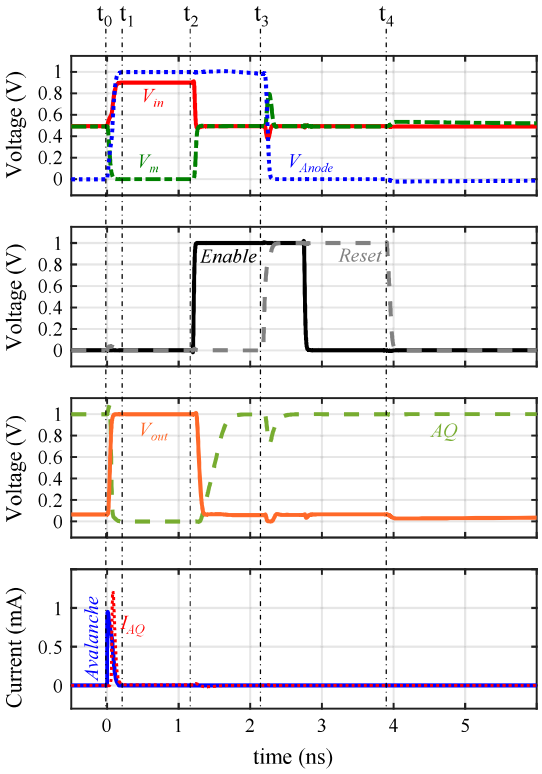


Fig. 12 Diagramme de fonctionnement simulé du circuit DS AQAR ( ). La phase d'extinction commence à t0 et le temps mort se termine à t4.

accordable à des fins de mesure, et C_c est le condensateur de couplage. Le diagramme temporel du circuit est présenté à la figure 12. Lorsque la tension de l'anode commence à augmenter lorsqu'une avalanche se produit (t_0), le détecteur le détecte par l'intermédiaire de C_c. Grâce à sa grande sensibilité, il détecte l'avalanche dans ses premiers instants. Ce signal de détection commande M_{AQ} et lui fait charger l'anode pour sortir le SPAD de sa zone de claquage afin d'arrêter l'avalanche (t_1) plus rapidement que ne le ferait la résistance PQ. Après un certain délai (t_2), la porte T est fermée pour préparer le détecteur à l'événement suivant en polarisant l'inverseur à son seuil de commutation. La fin du temps de maintien souhaité (t_3), M_{AR} réinitialise le SPAD dans sa zone de claquage. Après la phase de réinitialisation (t_4), alors que la porte T est ouverte, le pixel est à nouveau prêt pour la détection de photons. Comme nous l'avons vu, on peut remarquer que toutes les injections de charge se font dans une direction qui ne risque pas de compromettre la fiabilité du circuit. La figure 11 (b) montre le schéma du circuit AQAR de l'InDS. Il est similaire à celui du circuit DS AQAR, mais le circuit est placé au-dessus du SPAD et que tous les substrats sont donc égaux à la tension de l'anode. Il n'y a pas non plus de condensateur de couplage. Dans le circuit MS

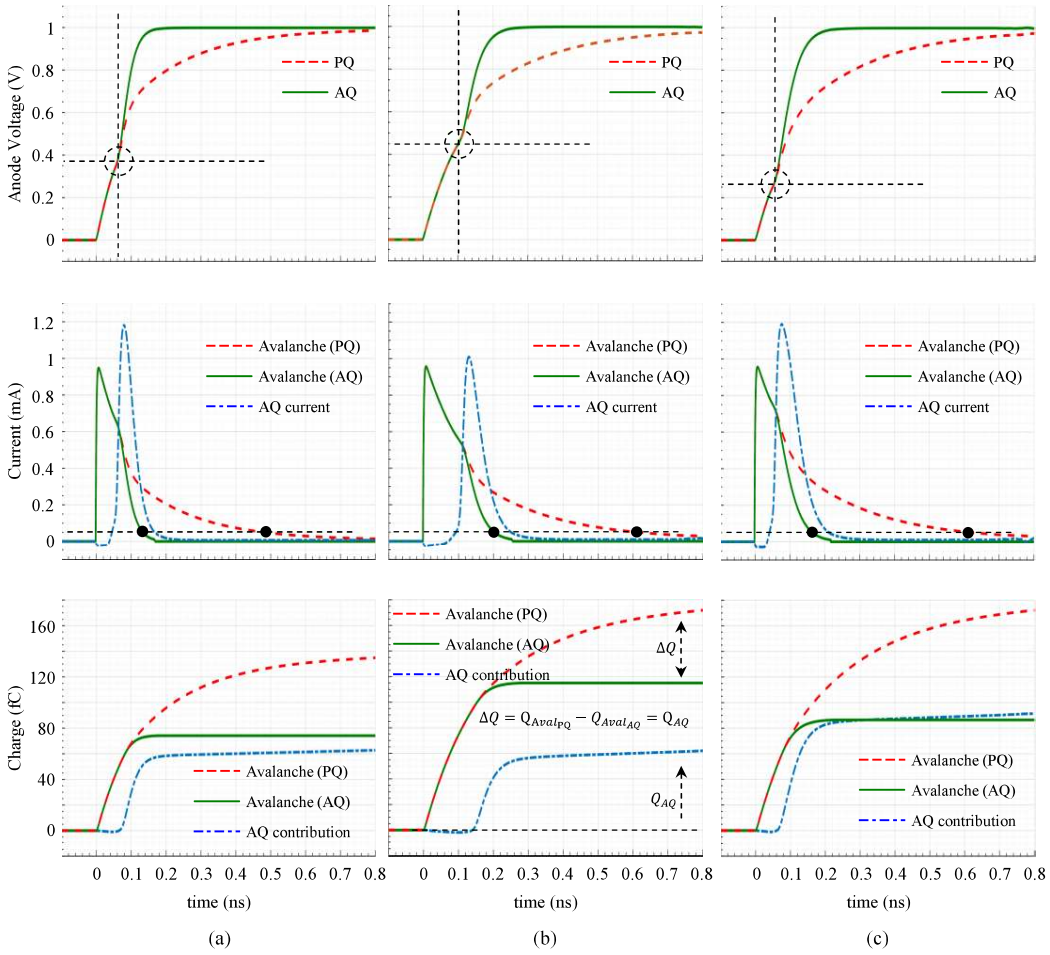


Fig. 13 Résultats de la simulation de post-layout pour : (a) pixel à détection directe (b) pixel à détection indirecte (c) pixel à détection mixte en modes PQ et AQ. De haut en bas : Tension transitoire de l'anode, courant d'avalanche et d'extinction active (I_{AQ}), charges d'avalanche et d'AQ.

AQAR, Fig. 11 (c), le condensateur de couplage C_c est de retour, et le reste du circuit est exactement comme le circuit InDS AQAR. Le fonctionnement de ces deux circuits est similaire à celui du circuit DS AQAR, à l'exception de leurs mécanismes de détection, qui sont expliqués dans leurs propres sections.

D'après les résultats de la simulation post-layout de la figure 13 et du tableau 2, le pixel MS, qui est le pixel le plus rapide et le plus efficace, détecte une avalanche avant 50 ps et l'éteint activement en 160 ps. Il en résulte une réduction de plus de 50 % de la charge d'avalanche par rapport au fonctionnement du pixel en mode PQ. Dans le cas du DS, la valeur de cette réduction est d'environ 44 %. La consommation d'énergie statique est d'environ 100 µW, principalement consommée par les lignes de retard contrôlant le temps mort. Le circuit d'extinction lui-même consomme environ 43 µW en raison de la polarisation de son inverseur de détection.

TABLEAU 2
RÉSULTATS DE LA SIMULATION DE LA POST LAYOUT POUR LES PIXELS DS, MS ET INDS, DANS LES MODES DE TREMPE ACTIVE ET PASSIVE.

	Direct Sensing	Mixed Sensing	InDS Sensing
Avalanche detection time (ps)	< 60	< 50	< 100
Quenching time in PQ (ps)	480	620	600
Quenching time in AQ (ps)	130	160	200
Avalanche charge in PQ (fC)	135	175	175
Avalanche charge in AQ (fC)	75	86	115
AQ charge contribution (fC)	60	89	60
Avalanche charge reduction (%)	44	51	34

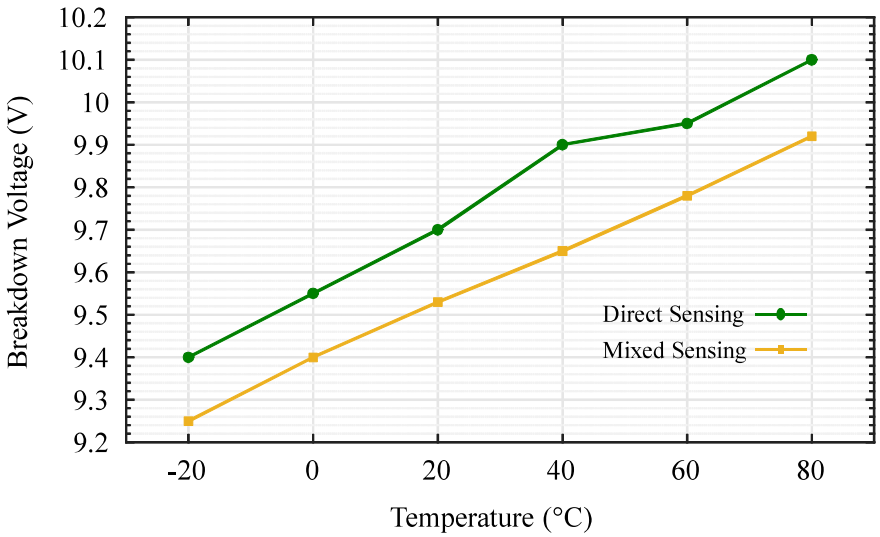


Fig. 14 Tension de claquage en fonction de la température pour les pixels à détection directe (non 3D) et à détection mixte (3D monolithique).

Dans les mesures, les effets du placement de l'électronique sur le SPAD sont étudiés pour la première fois. On observe que les tranchées isolantes (STI) ajoutés en raison du placement des circuits sur le SPAD augmentent à la fois le taux de comptage de l'obscurité (DCR : les avalanches lorsque le SPAD est dans l'obscurité) et les bruits d'impulsions ultérieures. En outre, elle entraîne une augmentation du claquage prématuré à travers le SPAD, ce qui fait que la tension de claquage du pixel MS est inférieure d'au moins 150 mV à celle du pixel DS (voir la figure 14). Par conséquent, la tension de polarisation excessive du pixel MS est limitée à environ 3 fois moins que celle du pixel DS en mode PQ afin d'éviter la saturation du pixel par les sources de bruit. Toutefois, en utilisant le circuit AQAR proposé, le niveau de bruit est abaissé de telle sorte que le pixel MS

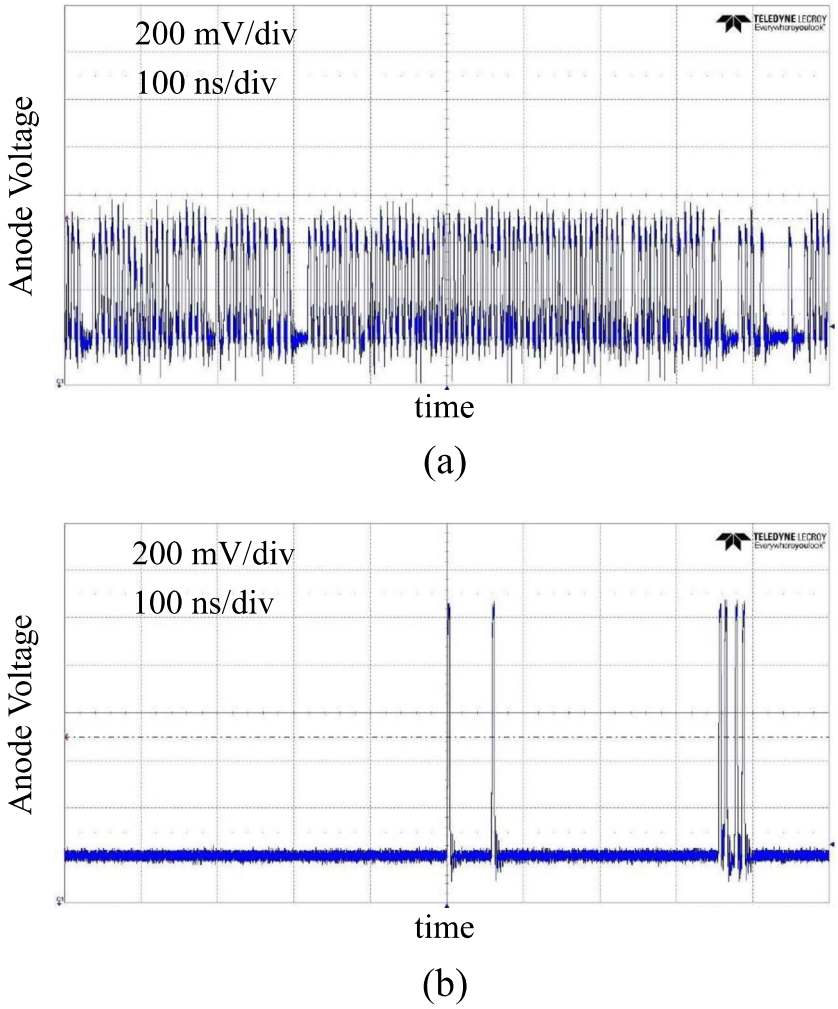


Fig. 15 Avalanches dans le pixel de détection mixte sous 500 mV de tension de polarisation excessive à 60 °C avec un temps mort de 8 ns : (en haut) mode PQ (en bas) mode AQ .

peut fonctionner normalement avec des tensions d'excès de polarisation plus élevées dans son mode de fonctionnement AQ, comme le montre la figure 15. En d'autres termes, l'électronique ajoutée au-dessus du SPAD augmente le niveau de bruit de telle sorte que le SPAD ne fonctionne pas bien avec des tensions d'excès élevées. Cependant, le circuit AQ proposé est très efficace dans la réduction du bruit qui compense le bruit ajouté, tout en ne consommant pas de surface de pixel dans une configuration BSI. Dans le pixel DS, on observe une réduction de 85 % du nombre d'impulsions secondaires par rapport au mode PQ à 800 mV de tension de polarisation excédentaire. Dans le pixel MS, pour avoir une comparaison entre les modes AQ et PQ, l'excès de tension de polarisation est limité à 250 mV. Dans ce cas, une réduction de 94 % du nombre d'impulsions a été mesurée. Ces résultats sont présentés à la figure 16.

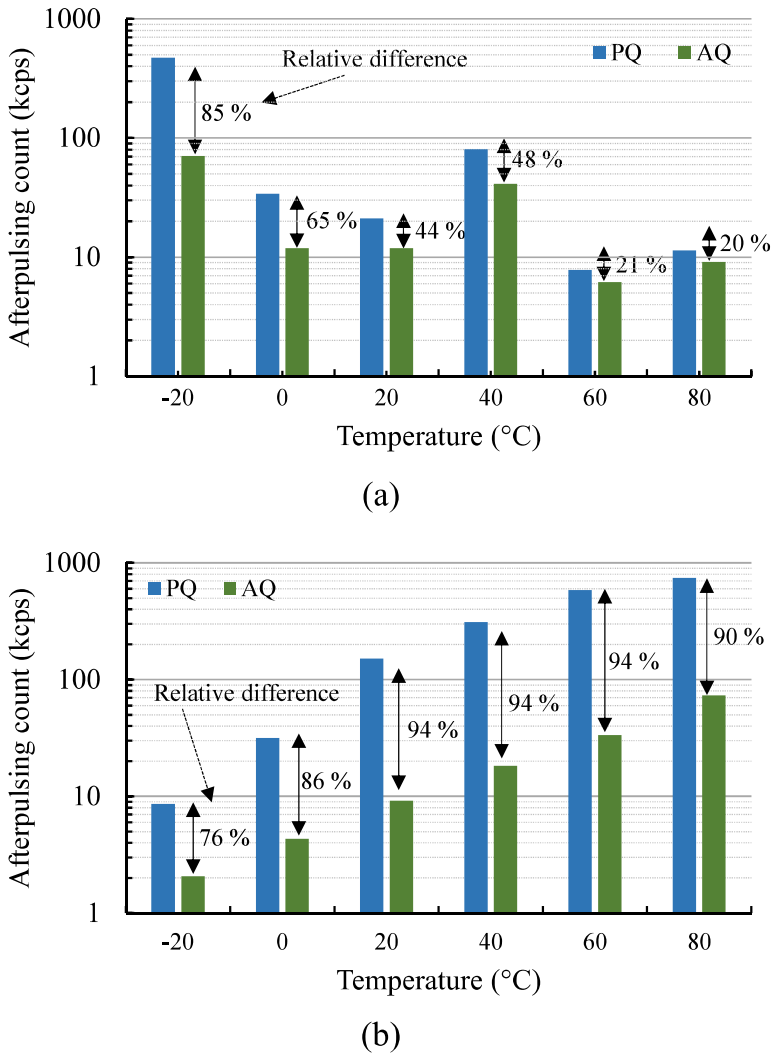


Fig. 16 Comptages après impulsion en modes PQ et AQ à différentes températures avec un temps mort de 8 ns. (en haut) Pixel à détection directe sous 800 mV de tension de polarisation excessive (en bas) Pixel à détection mixte sous 250 mV de tension de polarisation excessive.

Chapitre 4 : Circuit de polarisation de l'inverseur indépendant de la tension de substrat

Après avoir réalisé la première intégration d'un circuit AQAR dans un pixel SPAD 3D monolithique, un circuit AQAR économe en énergie, presque aussi sensible que le circuit précédent et également compatible avec la structure 3D monolithique, est introduit. Au lieu de polariser l'inverseur de détection à son seuil de commutation, où il consomme la puissance statique maximale, la polarisation de l'inverseur à une tension légèrement inférieure à ce point peut réduire fortement la consommation d'énergie statique tout en conservant une sensibilité suffisamment

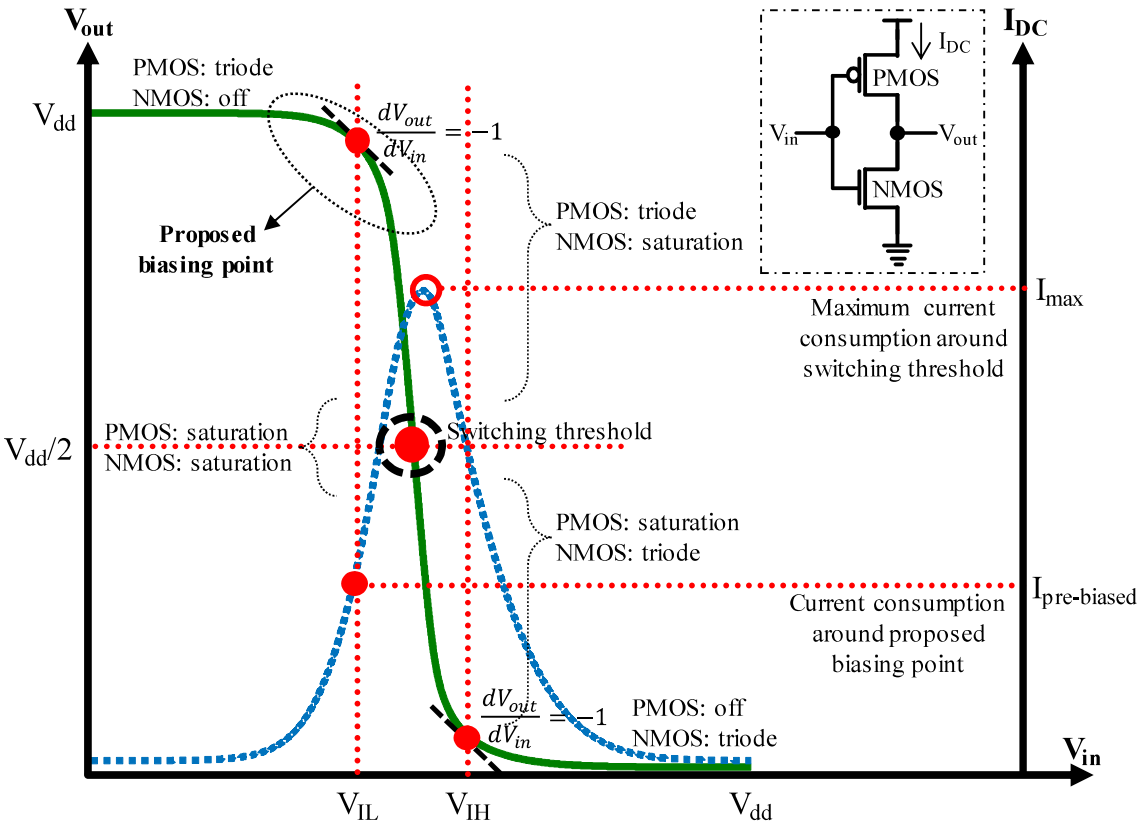


Fig. 17 Caractéristique de transfert de tension (VTC) - courbe verte - et caractéristique de transfert de courant (CTC) - courbe bleue en pointillés - d'un onduleur typique. Les points de fonctionnement des transistors sont indiqués pour chaque partie de la VTC. La consommation de courant pour le point de polarisation proposé est indiquée sur la CTC.

élevée. La figure 17 illustre cette idée. Le circuit de polarisation proposé (figure 18) présente trois caractéristiques essentielles : pendant une avalanche, il ne lutte pas contre l'augmentation de la tension de l'anode, pendant la phase de réinitialisation, il ramène le point de polarisation à sa valeur initiale dès que possible et, pour être compatible avec la structure 3D monolithique, il est indépendant de la tension du substrat. Par conséquent, nous présentons pour la première fois une source de courant indépendante de la tension du substrat pour polariser les circuits AQAR proposés. Le schéma de ce circuit est présenté à la figure 19. Le circuit AQAR final proposé, basé sur un inverseur pré-polarisé, est conçu sous trois formes différentes : DS, MS et pixels InDS, comme dans le chapitre précédent (voir figure 20).

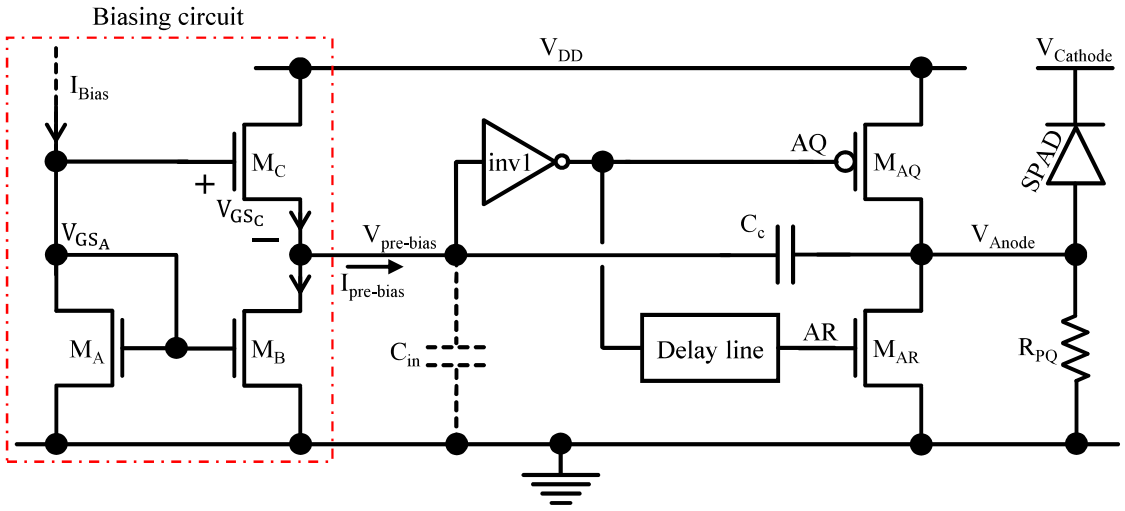


Fig. 18 Schéma du circuit de détection et d'extinction d'avalanche pré-polarisé. Le circuit de polarisation est inscrit dans le rectangle rouge. Le condensateur en pointillé représente la capacité totale observée à l'entrée de inv1 (capacités parasite et d'entrée).

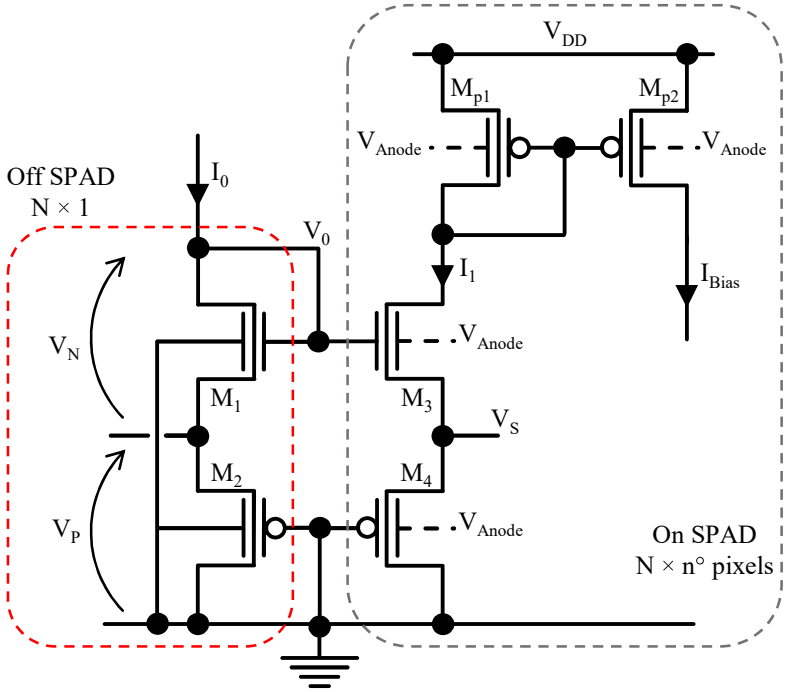


Fig. 19 Schéma de la source de courant indépendante de la tension du substrat. Les transistors à l'intérieur du rectangle rouge sont implémentés une fois en dehors de la zone active du SPAD. V_0 est partagé par tous les pixels du réseau et les transistors dans le rectangle gris sont placés sur le SPAD de chaque pixel.

Dans les résultats expérimentaux, un temps d'extinction d'environ 500 ps est mesuré pour le pixel DS avec un excès de tension de polarisation de 700 mV, ce qui est 600 ps plus rapide que le mode PQ. Presque le même temps est mesuré pour le pixel MS à 500 mV de tension de polarisation

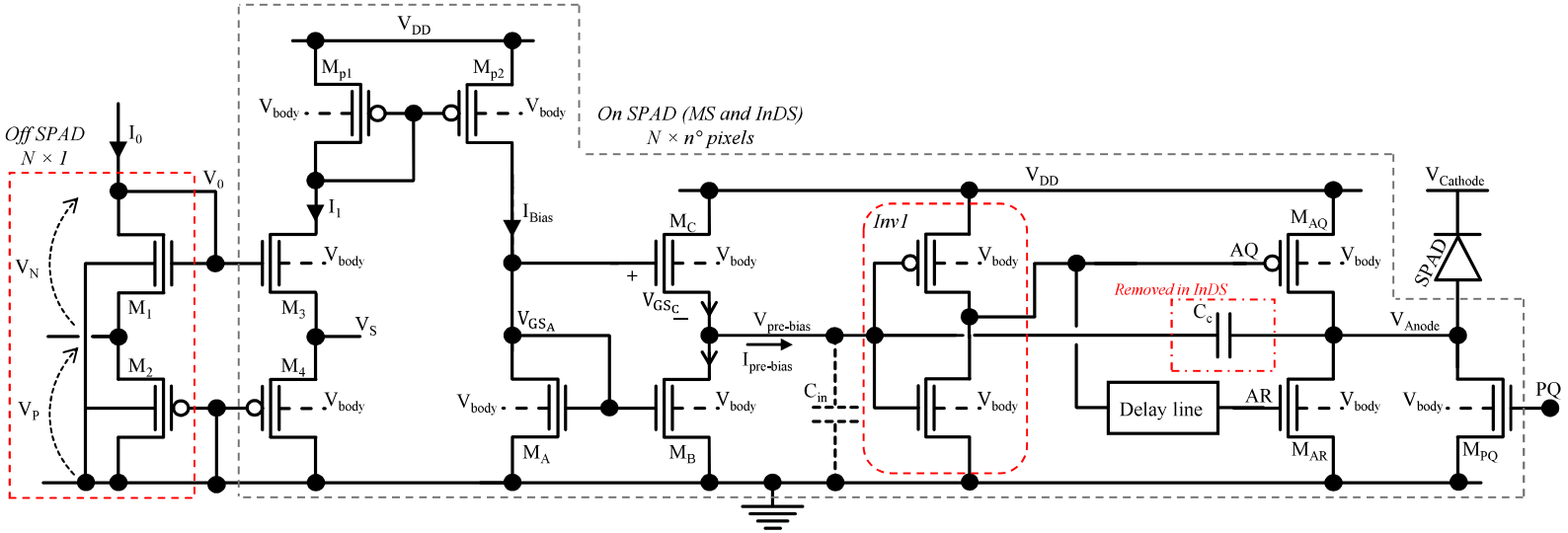


Fig. 20 Schéma complet du circuit AQAR pré-biaisé proposé. Dans le pixel DS, toutes les tensions de substrat (V_{body}) sont connectées à la masse. Dans les pixels MS et InDS, elles sont toutes égales à la tension de l'anode. Pour le pixel InDS, le condensateur de couplage doit être retiré du schéma. M_1 et M_2 sont implémentés une seule fois en dehors du SPAD et sont utilisés pour tous les pixels.

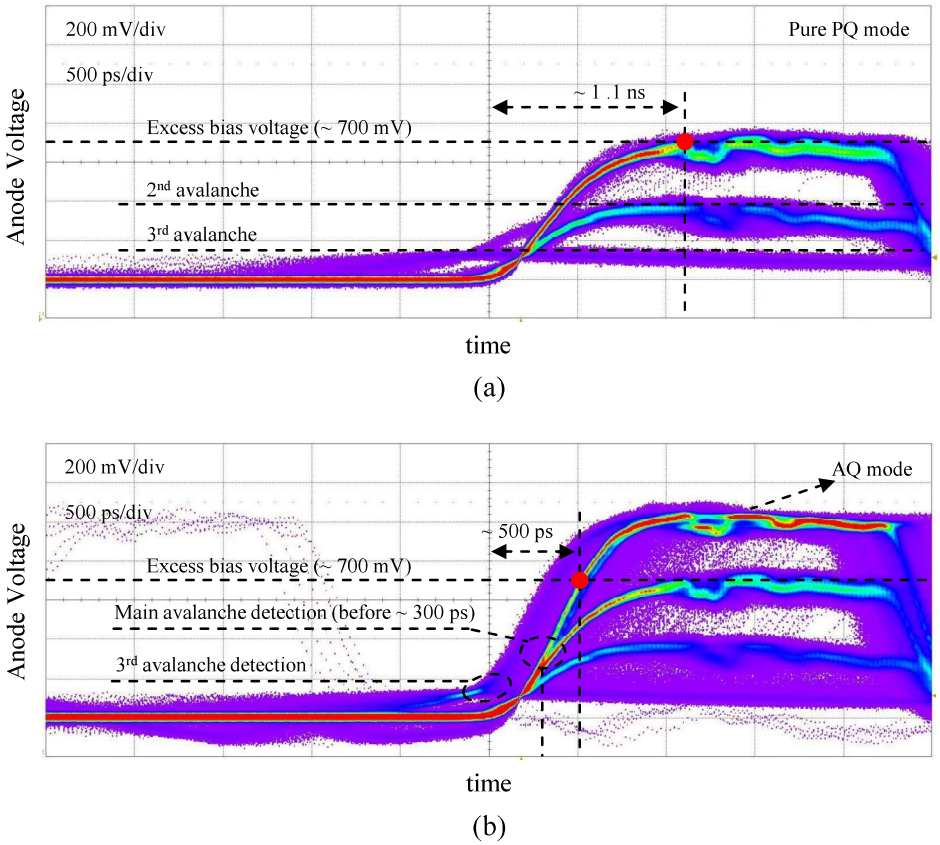


Fig. 21 Transition de la tension anodique dans le pixel à détection directe sous 700 mV de tension de polarisation excessive et $R_{Sensitivity}$ = 533 kΩ, à 60 °C, lorsque l'oscilloscope est en mode persistance : (a) mode PQ (b) lorsque AQ est activé.

excédentaire, comme le montre la figure 21. Ici aussi, on observe les mêmes effets du placement de l'électronique sur le SPAD que pour les pixels précédents : l'augmentation du niveau de bruit dans la structure 3D monolithique et la compensation de ce bruit grâce au circuit AQ. Les mesures montrent une réduction de 62 % et 96 % du nombre d'afterpulsing pour les pixels DS et MS, respectivement. Dans le pixel MS, grâce au circuit AQ proposé, un APP de plus de 40 % en mode PQ est réduit à environ 3 % en mode AQ. La réduction du nombre d'impulsions et la réduction de l'APP sont présentées dans les figures 22 et 23, respectivement. On peut conclure que les circuits AQAR proposés sont essentiels pour les pixels 3D monolithiques et qu'ils peuvent transformer le SPAD bruyant en un SPAD qui se comporte presque bien. Tous ces résultats sont obtenus avec une consommation statique inférieure à 15 µW, soit 3 fois moins que le circuit proposé au chapitre 3.

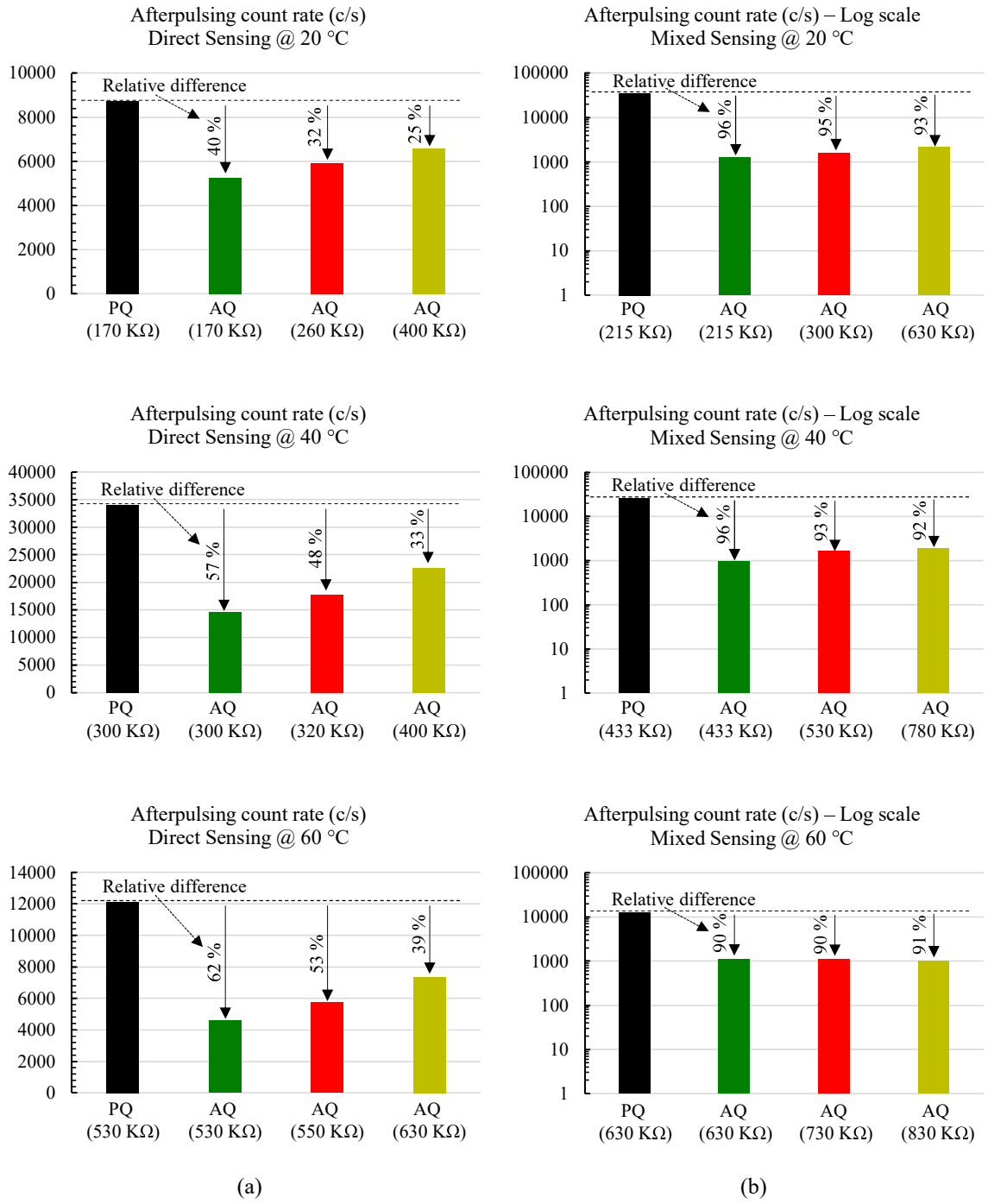


Fig. 22 Taux de comptage afterpulsing en modes PQ et AQ à différents niveaux de sensibilité avec un temps mort de 6 ns pour différentes températures : 20, 40 et 60 °C (de haut en bas). (a) pixel à détection directe sous 800 mV de tension de polarisation excessive (b) pixel à détection mixte sous 250 mV de tension de polarisation excessive. La différence relative des taux de comptage afterpulsing entre les modes PQ et AQ est indiquée sur les figures. Les valeurs de résistance sur l'axe des x correspondent à différentes valeurs de sensibilité.

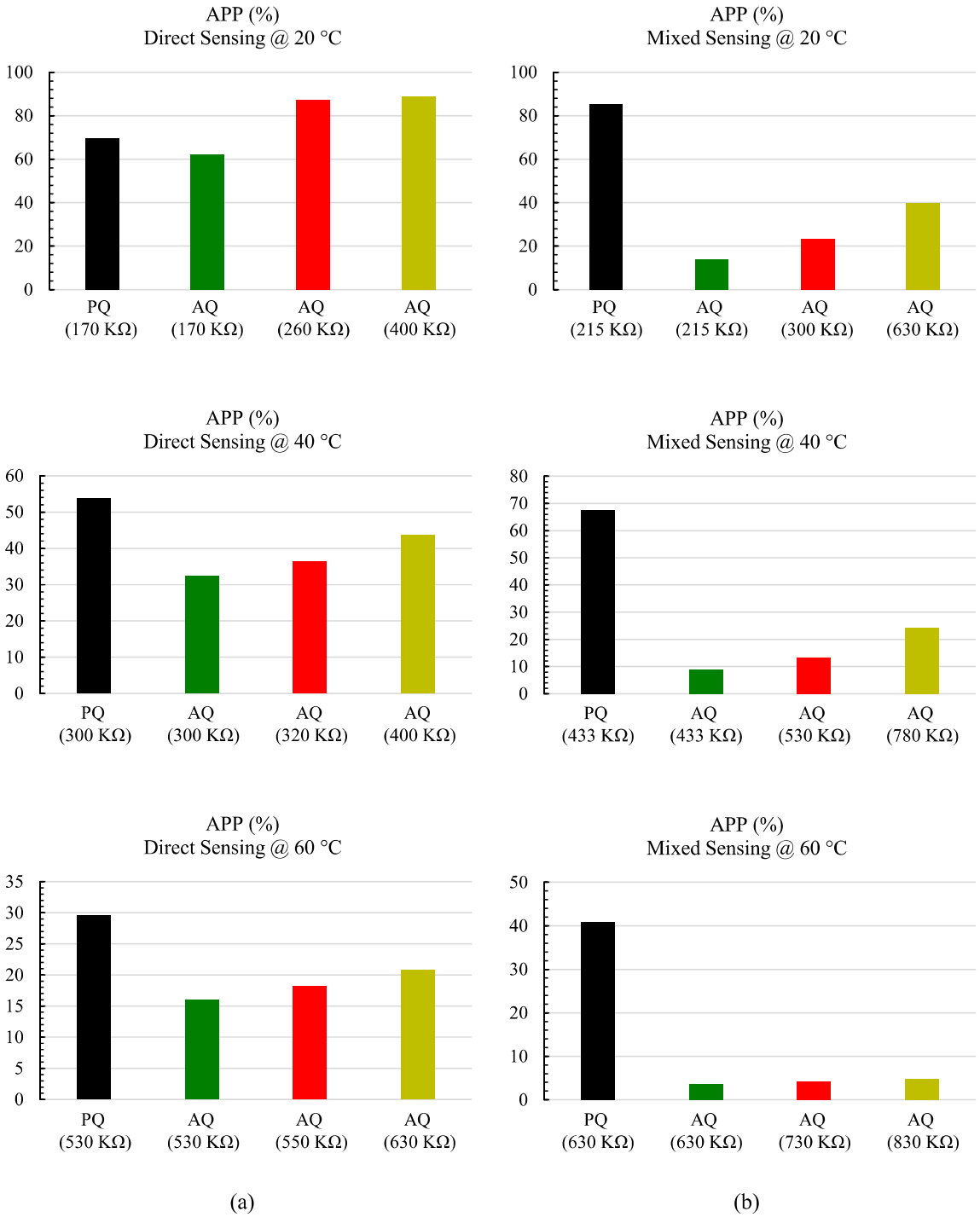


Fig. 23 Pourcentage de probabilité d'afterpulsing (APP) en modes PQ et AQ à différents niveaux de sensibilité avec un temps mort de 6 ns pour différentes températures : 20, 40 et 60 °C (de haut en bas). (a) pixel à détection directe sous 800 mV de tension de polarisation excessive (b) pixel à détection mixte sous 250 mV de tension de polarisation excessive. Les valeurs de résistance sur l'axe des x correspondent à différentes valeurs de sensibilité.

Chapitre 5 : Conclusions

Nous avons présenté trois circuits AQAR différents et prouvé leur fonctionnalité et leur efficacité dans la réduction de l'afterpulsing. Nous avons également conçu le tout premier pixel actif SPAD 3D monolithique et des mesures justifient son fonctionnement. Une nouvelle méthode de détection des avalanches, appelée détection mixte, a également été introduite pour améliorer le temps de détection des avalanches. Une source de courant innovante indépendante de la tension du substrat est également introduite à des fins de polarisation dans le pixel 3D monolithique en technologie CMOS FD-SOI. Malgré le bruit élevé inhérent au dispositif SPAD dans la structure 3D monolithique, les circuits AQAR proposés l'atténuent efficacement, ce qui permet au dispositif de fonctionner de manière fiable même à des tensions de polarisation excédentaires plus élevées. En utilisant le pixel 3D monolithique dans une configuration éclairée par l'arrière, avec les circuits AQAR proposés, il est possible de faciliter la recherche d'un compromis entre le taux de comptage maximal, l'afterpulsing, le facteur de remplissage et l'efficacité de la détection des photons. Grâce aux circuits proposés, l'afterpulsing et le DCR du pixel SPAD 3D monolithique sont mesurés pour la première fois. Ces résultats sont prometteurs pour la réalisation de systèmes de détection denses et efficaces. Ces résultats sont prometteurs pour la réalisation de réseaux SPAD denses et à faible bruit.

Mohammadreza DOLATPOOR LAKEH

Design of a Single Photon Sensor on a 28 nm FD-SOI CMOS technology

Résumé

Les photodétecteurs à avalanche de photons uniques (SPAD) suscitent un intérêt croissant dans les applications de comptage de photons pour lesquels un taux de comptage élevé est nécessaire. Dans les pixels à SPAD conventionnels, un taux de comptage élevé est généralement source de bruit dit « afterpulsing ». Les circuits d'extinction active (AQ) sont une solution pour limiter cet effet au prix d'une consommation de surface supplémentaire réduisant ainsi la sensibilité du pixel. L'efficacité d'un circuit AQ dépend du temps de détection de l'avalanche, de sa contribution en charge au processus d'extinction, ainsi que de ses consommations d'espace et de puissance. Cette thèse vise à concevoir des circuits AQ capables de réduire efficacement l'effet d' « afterpulsing » avec une faible consommation d'énergie et en occupant une surface minimale pour permettre ainsi le fonctionnement de pixels SPAD dans les applications à taux de comptage élevé. À cet égard, trois circuits AQ sont conçus avec différentes variations dans une technologie 28 nm CMOS FD-SOI. Les mesures expérimentales démontrent que ces circuits sont capables de réduire fortement l'effet d' « after pusling » tout en respectant les compromis mentionnés. De plus, la première intégration 3D monolithique de circuits AQ au-dessus d'une SPAD est présentée dans ce travail.

Résumé en anglais

The Single Photon Avalanche Diode (SPAD) is gaining increasing interest in photon counting applications requiring a high photon count rate. However, conventional pixels face challenges with elevated afterpulsing noise as the count rate rises. Active Quenching (AQ) circuits present a solution to navigate this trade-off, albeit with increased area consumption, diminishing pixel sensitivity. The efficacy of an AQ circuit hinges on factors such as avalanche detection time, its charge contribution to the quenching process, and associated area and power consumptions. This thesis endeavors to design AQ circuits capable of effectively mitigating afterpulsing noise with minimal power consumption, while occupying the smallest possible area. This design objective aims to enable the operation of SPAD pixels in high photon count rate applications. To achieve this goal, three AQ circuits have been meticulously designed in various configurations using a 28 nm FD-SOI CMOS technology. Each circuit demonstrates a robust reduction in afterpulsing effects while carefully balancing the specified trade-offs, as verified through measurements. Furthermore, this work introduces the pioneering concept of the first monolithic 3D integration of SPAD with AQ circuits.