UCLouvain

epl

École polytechnique de Louvain

Simulation and Characterization of an Ultraviolet enhanced Backside-Illuminated Single-Photon Avalanche Diode

Author: **Diego Haya Enriquez**
Supervisors: **Denis Flandre, Nicolas André**
Reader: **Dimitri Lederer**
Academic year 2021–2022
Master [120] in Physical Engineering

Abstract

In this master thesis is simulated and characterized a novel SOI backside-illuminated SPAD designed especially to target the UV region in terms of photodetective capabilities. Backside-illumination has the potential to increase the fill factor in SPAD arrays as the micro-cells and front-end electronics can be optimized separately. The use of SOI technology allows the SPAD to limits absorption losses and improve photodetection efficiency thanks to the thin transparent BOX layer. The thin body and depletion region is optimized to absorb UV photonx, which has great potential in applications where this wavelength range is in use such as positron emission tomography. The device is simulated with Silvaco's process simulator (Athena) and device simulator (Atlas) that are used to model the doping profile, breakdown voltage, triggering probability and quantum efficiencies. The result for the breakdown voltage (8.3 V) is similar to earlier measurement made before this work. The simulated peak quantum efficiency in the case of backside-illumination is 67% at 450 nm which is a sensible value considering other state of the art devices, however it is quite far from the measured value which is more than 90%. Dark counts and afterpulse characterizations were attempted on another device using a Gaussian-shaping amplifier and post-processing methods for the signal analysis. The setup and methods are well suited for DCR characterization (136 cps/μm^2) was measured at room temperature for an excess bias of 0.8 V and 94 cps/μm^2 for sub-zero temperatures. However the amplifier output pulse did not allow to resolve afterpulses despite pulse pile-up correction and the pulse height histograms resulted in exponential decays which is unexpected and should be investigated.

Acknowledgement

This work is the culmination of years of education and training . As any scientific project, it could not have been achieved on my own and I gladly wish to express my gratitude to the people that helped me throughout this journey.

I first would like to thank my supervisor Prof. Denis Flandre for his insights and making himself available as well as useful suggestions when I was lost or needed to overcome a hurdle.

I am also grateful for the help provided by Valeriya Kilchytska who guided me when I had to make use of the high-tech equipments of the Welcome lab of which I would not have any idea how to operate them without her. She always answered my questions as well when needed and her extensive knowledge in diverse areas proved to be very helpful when I had to process data.

I would like to express a heartfelt thanks to Pierre Gérard without whom it would not have been possible to set up the amplifier circuit which turned out to be a critical piece of this work as it allowed me to obtain important informations to analyze.

Finally, a massive thanks goes to Nicolas André who accompanied me during the year and was always there from start to finish. His enthusiastic personality was always refreshing and made working on this thesis very pleasant. He is always very dedicated to students and the university truly cannot afford to lose him.

Contents

1 Introduction **7**

2 State of the art **8**
2.1 SPAD and SiPM sensors . . . 8
2.1.1 The SPAD cell . . . 8
2.1.2 SiPMs . . . 10
2.2 Properties and performances . . . 12
2.2.1 Photodetection probability . . . 12
2.2.2 Dark count rate . . . 14
2.2.3 Afterpulsing . . . 14
2.2.4 Crosstalk . . . 14
2.2.5 Breakdown voltage . . . 15
2.2.6 Timing jitter . . . 16
2.2.7 Dead time . . . 17
2.3 Design features and types of sensors . . . 17
2.3.1 Custom technologies . . . 17
2.3.2 Standard CMOS technologies . . . 19
2.4 Applications and FOM . . . 20
2.4.1 Performances and Figure of merits . . . 21
2.5 SPAD comparisons . . . 23

3 Theory of the SPAD cell operation **24**
3.1 Impact ionization and breakdown . . . 24
3.1.1 Avalanche multiplication . . . 24
3.1.2 Ionization coefficients . . . 26
3.1.3 Edge breakdown . . . 26
3.2 Photodetection probability . . . 28
3.2.1 Photon absorption . . . 28
3.2.2 Avalanche triggering probability . . . 29
3.2.3 Photodetection probability . . . 31
3.3 Dark count rate and afterpulse . . . 32
3.3.1 Trap assisted carrier generation . . . 33
3.3.2 Band to band tunnelling generation . . . 36
3.3.3 Temperature and Voltage bias dependence of DCR . . . 37
3.3.4 Afterpulsing . . . 38
3.4 Quenching and equivalent electrical model . . . 41
3.4.1 Equivalent electrical model for a single SPAD . . . 41
3.4.2 Output pulse . . . 44

4 The BSI UV enhanced SOI SPAD **46**
4.1 Design and structure . . . 46
4.2 Device simulation . . . 47
4.2.1 Optimized doping profile with Athena . . . 48
4.2.2 Device simulation with Atlas . . . 48
4.3 Characterization: First results . . . 54

5 Characterization of dark count on another device **57**
5.1 I-V curves . . . 57
5.2 PCB setup and A225 shaping amplifier . . . 60
5.3 Pulse analysis and signal processing . . . 63
5.3.1 Peak detection . . . 63
5.3.2 Pulse pile-up correction . . . 64
5.4 DCR characterization VS temperature . . . 66
5.4.1 Inter-arrival times and pulse amplitude histograms . . . 66
5.4.2 Characterization of the DCR with respect to the temperature . . . 68

6 Conclusion and future prospects **69**

List of Figures

1 Three-dimensional cross-section of a planar SPAD . . . 9
2 I–V characteristic of a SPAD . . . 10
3 Depiction of the micro-cell structure of the SiPM surface . . . 11
4 Pulse produced by a micro-cell in response to photon absorption . . . 12
5 Example of a PDP curve for a SPAD at different applied bias . . . 13
6 Representation of direct and indirect optical crosstalk, electrical and delayed crosstalk . . . 15
7 Typical reversed I-V curve for a single SPAD device . . . 16
8 Jitter measured as the delay time with respect to a precise and punctual light source . . . 17
9 Custom SPAD designs . . . 18
10 CMOS SPAD designs . . . 19
11 SOI SPAD design . . . 20
12 Impact ionization causing avalanche current in the presence of a high electric field . . . 24
13 Illustration of a secondary e-h pair generated by impact ionization . . . 25
14 Depiction the doping profile with curved edges generated by diffusion . . . 27
15 SPAD design structure exhibiting a diffused n-well guard ring . . . 27
16 Silicon optical properties . . . 29
17 Basic structure of a SPAD . . . 30
18 Photondetection efficiency examples . . . 32
19 Schematic of several sources of dark counts in a SPAD device . . . 33
20 Generation and recombination in an indirect bandgap semiconductor . . . 34
21 Trap assisted tunelling . . . 35
22 Band-to-band tunelling mechanism . . . 36
23 Temperature and Voltage bias dependence of DCR . . . 37
24 Arhennius plot illlustration . . . 38
25 Afterpulsing probability with respect to the dead time and excess bias voltage . . . 41
26 Traditional equivalent electrical model of a single SPAD micro-cell . . . 42
27 Voltage and current output waveform across the SPAD cell . . . 43
28 Pulse waveforms of a SPAD of that operates with a passive quenching circuit . . . 44
29 Pulse output modes . . . 45
30 Design features of the BSI SOI device . . . 46
31 Design variations of the proposed device shape and dimensions . . . 47
32 Simulated doping profile . . . 48
33 Simulated I-V curve under dark conditions . . . 50
34 Simulated electric field profile at breakdown . . . 51
35 Simulated avalanche triggering probabilties . . . 51
36 Illumination modes of the simulated device . . . 52
37 Simulated I-V curves under illumination . . . 53
38 Simulated quantum efficiency and photodetection probability . . . 54
39 Measured I-V curves under dark and low light power . . . 55
40 Measured quantum efficiencies . . . 56
41 Measured DCR with respect to temperature and overvoltage . . . 56
42 Arrhenius plot computed with DCR measurement results . . . 57
43 I-V curves of the devices measured on SOITEC and Ultrasil SOI wafer . . . 58
44 Breakdown voltage characterization with respect to temperature . . . 60
45 PCB board layout . . . 61
46 Circuit schematic of the experimental setup . . . 61
47 A225 amplifier output pulses . . . 62
48 Results of the filtering and peak-finding methods . . . 64
49 Illustration of the pile-up correction method . . . 65
50 Results of the pile-up correction method application . . . 66

51 Histograms of inter-arrival times and pulse amplitudes 67
52 Scatter plot of inter-arrival time VS pulse amplitude 68
53 DCR as a function of temperature at the excess bias voltage of 0.8 V 69

List of Tables

1 State of the art 23
2 SPAD design variations 47
3 Physical models of the TCAD simulation 49
4 Measured breakdown voltages for six types of devices for two different SOI wafer suppliers . . 58
5 Comparison of the SPAD performances for timing and counting applications targetting smaller wavelength ranges 69

1 Introduction

The idea behind the Single photon avalanche photodiodes (SPAD for short) emanates from the study of avalanche breakdown and microplasma fluctuations in silicon initiated by McKay [1], [2], Champlin [3], McIntyre [4], [5] et al. It was clear from the beginning that the avalanche breakdown process could be exploited for photodetection applications and the first SPAD detector prototypes can be traced back in the 1960's [6], [7]. These large devices and subsequent developments required high reverse bias voltages and were stand-alone structures incompatible with other circuit elements but still promising for single photon detection.

Thus far the photomultiplier tubes that exploits the photoelectric effect were the default choice for this purpose. Although providing very good sensitivity, noise and timing characteristics, these detectors have severe handicaps that limit their application range: they are, bulky, fragile, require high operating voltages, are very sensitive to magnetic fields and their price is high because the complicated mechanical structure inside the vacuum container is mostly handmade [8].

With intense research and development over the years and the advent of semiconductor technology, the SPAD detectors became smaller, faster, more precise and with the improvements of the front-end electronics it became possible to connect individual cells with each other into dense arrays [9] replacing progressively photomultipliers when these proved to be inconvenients. Eventually, an important milestone was reached in 2003 with the demonstration of the first SPAD-based pixel arrays in CMOS technology [10] which opened new perspectives for imaging applications as SPAD based cameras were shown to be feasible.

Nowadays, the SPAD is a mature technology for a wide range of applications. It is widespread in the automotive industry as LiDAR detectors [11], used in quantum computing for photon counting purposes [12] and in biophotonics for imaging applications [13].

With positron emission tomography (PET), SPAD array imagers are used to reveal the metabolic or biochemical function of tissues and organs [14]. As such, these imagers help diagnose some conditions including cancer. It can also help to find out where and whether cancer has spread potentially saving lives. Therefore research is still ongoing to improve the SPAD imagers efficiency, reliability and resolution [15]. For instance, these PET devices use scintillator crystals that emit UV photons that must be detected by the SPAD cell. This is one of the reason that motivates the development of SPADs that are designed and optimized to detect photons in this wavelength range .

This work intend to describe, simulate and characterize such a SPAD cell developed at UCLouvain that is optimized to detect UV, sub-UV photons.

With the aim of doing so, the SPAD working principle is first presented as well as its important parameters that must be characterized to properly assess the device performances. A review of the different structure and types of SPAD cells is then adressed to give a sense of the current state of the art and interplay between the parameters to be measured and the design. Useful figure of merits for targetted applications are introduced so that the overall performance can be evaluated which allows for a comparison between different devices. Based on these figure of merits, some selected state-of-the-art SPADs targetting UV photon detection are compared with one another to give an idea of the current benchmark and performances that are to aim for.

In the second part, the theory of the SPAD cell operation is layed out more formally to get a deeper insights into the device physics and provide the necessary tools for simulating the device, establish measurement protocols and discussing the results.

In the third chapter, the device under study is presented with its design choices and dimensions and is

simulated using TCAD tools to provide a first assessment of its main characteristics and performances. In the same section, the results of a first measurement campaign performed on a single device done outside this work are presented and discussed based on the theoretical background presented and compared with the simulation.

The section coming after further charatcterize other samples and a method to capture the SPAD output signal is presented using the tools available in our laboratory.

Finally, the device performances that were measured are compared to the current state of the art and future perspective are suggested so that the missing parameters could be characterized and to point out some potential directions.

2 State of the art

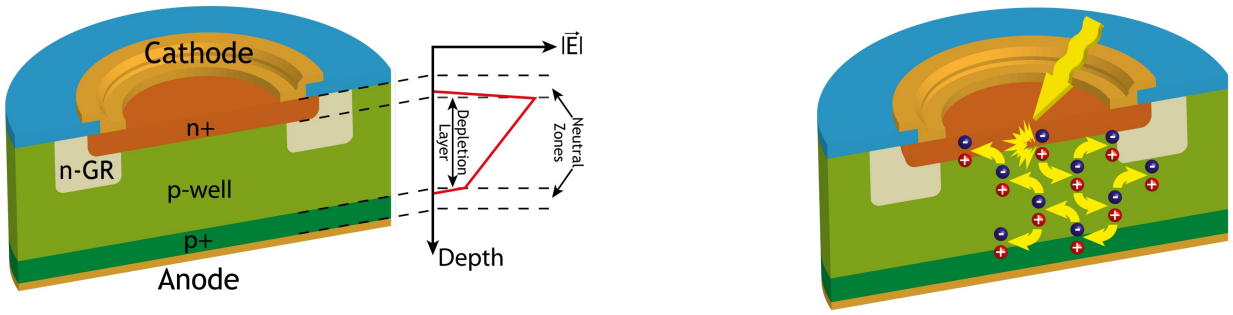
In this section is presented first the device working principle as well as the characteristics that are necessary to assess the performances of the SPAD/SiPM, so a presentation of the important parameters is introduced along with how they can be combined to make useful and relevant figure of merits. Also, an overview of the different main types of sensors with their most important design features is outlined to emphasize the relationships with performances and relate them with typical applications where these are in use nowadays. Finally, the figure of merits are developed based on the parameters introduced in order to assess the performances of state-of-the art devices that have similar goals to the novel apparatus developed at UCLouvain.

2.1 SPAD and SiPM sensors

2.1.1 The SPAD cell

A SPAD is a p-n junction (Figure 1a) reverse biased at a voltage well above its breakdown point V_{BD}. Above V_{BD}, the electric field in the p-n junction depletion layer is high enough so that upon the absorption of a photon, a single photogenerated carrier is able to trigger a macroscopic self-sustaining avalanche current by impact ionization [16] called avalanche breakdown, Geiger breakdown or simply breakdown. The breakdown process depicted on Figure 1b is triggered by the creation of an electron-hole pair (e-h pair) following the absorption of a photon. In the presence of a high electric field, the pair induces the multiplication of the number of carriers inside the depletion layer: due to the high energy acquired by the field, a given pair has a high probability to ionize another atom of the lattice. Once they collide a new e-h pair is created that is itself accelerated and causes other impact ionization events giving birth to carrier multiplication.

The noteworthy features on Figure 1a that are common in all SPAD types are the p-n junction (n+ shallow implant on top of a p-well on the figure), the n+ implant also acts a low resistivity sinker for electrons flowing to the cathode while the p+ implant at the bottom serves the same purpose for hole collection. The guard ring (n-GR) at the periphery lowers the electric field at the edge of the junction, thus preventing premature edge breakdown (see subsubsection 3.1.3).



(a) (b)

Figure 1: **(a)**Three-dimensional cross-section of a planar SPAD and a simplified representation of the electric field magnitude across the device. The n+ shallow implant and the p-doped well define the depletion layer and avalanche region. **(b)** Avalanche triggered by the photogenerated electron-hole pairs in the depletion layer. From [16]

The SPAD working principle from photon absorption to the device recharge can be subdivided into three main step [13] depicted on Figure 2 :

1. The SPAD is reverse biased above its breakdown voltage V_{BD} such that a single photon incident on the active (i.e. photosensitive) device area can trigger an avalanche of secondary carriers and the current rises swiftly across the junction. The avalanche build-up time is typically on the order of picoseconds, so that the associated change in voltage can be used to precisely measure the time of the photon arrival.

2. Once started, the self-sustaining avalanche in the SPAD needs to be stopped as soon as possible to prevent the destruction of the device itself due to the high current. The corresponding quenching occurs by lowering the SPAD bias breakdown plus overvoltage ($V_{bd} + V_{ov}$) below the breakdown voltage by using a resistor in series with the SPAD.

3. The voltage across the SPAD then needs to be restored to its initial value above the breakdown voltage before the next photon can trigger another avalanche. During this interval, which is typically on the order of tens of nanoseconds and known as dead time, the SPAD will be almost insensitive, depending on the exact quenching and recharge mechanism and SPAD front-end implementation (see subsection 3.4). The sensitivity of the SPAD will then gradually increase, until it reaches its nominal value when the recharge is complete. The time it takes for the SPAD to recharge to the full operating voltage is called the recovery time.

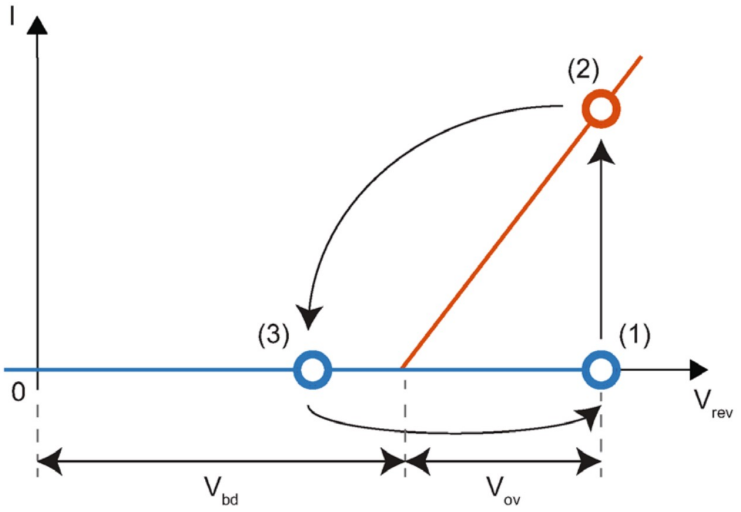
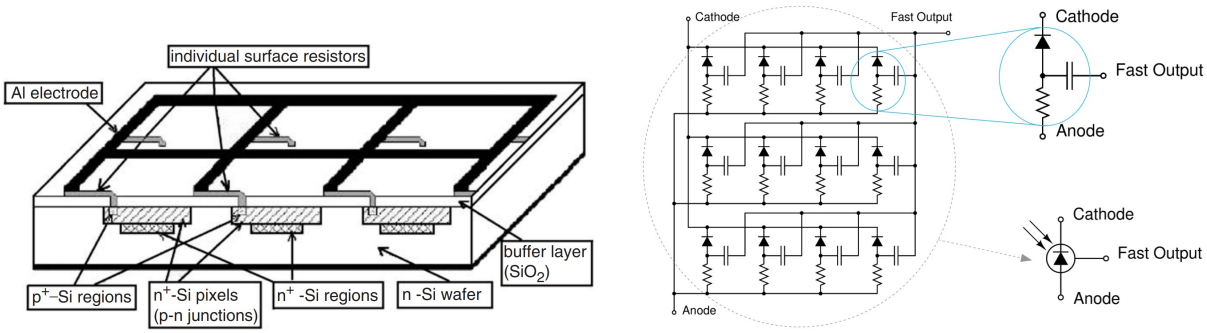


Figure 2: I–V characteristic of a SPAD; V_{rev} is the reverse bias, V_{bd} the breakdown voltage, and V_{ov} is the overvoltage, also known as excess bias. During **(1)** Geiger breakdown takes place and thec current rises almost instantaneously, in **(2)** the SPAD is quenched and the voltage bias fall below the breakdown voltage which stops the self-sustaining avalanche process. Finally during **(3)** the voltage across the junction gradually recovers to its original value until the SPAD is triggered again. From [12]

2.1.2 SiPMs

A single SPAD sensor biased above its breakdown voltage functions as a switch, in either an 'on' or 'off' state, this consequently results in a binary output. SPAD sensors are designed to capture very dim light sources and as such the output signal is not meant to be proportional to the incident source: regardless of the number of photons absorbed within a diode at the same time, it will produce a signal no different to that of a single photon and proportional information on the magnitude of an instantaneous photon flux is not available [17].

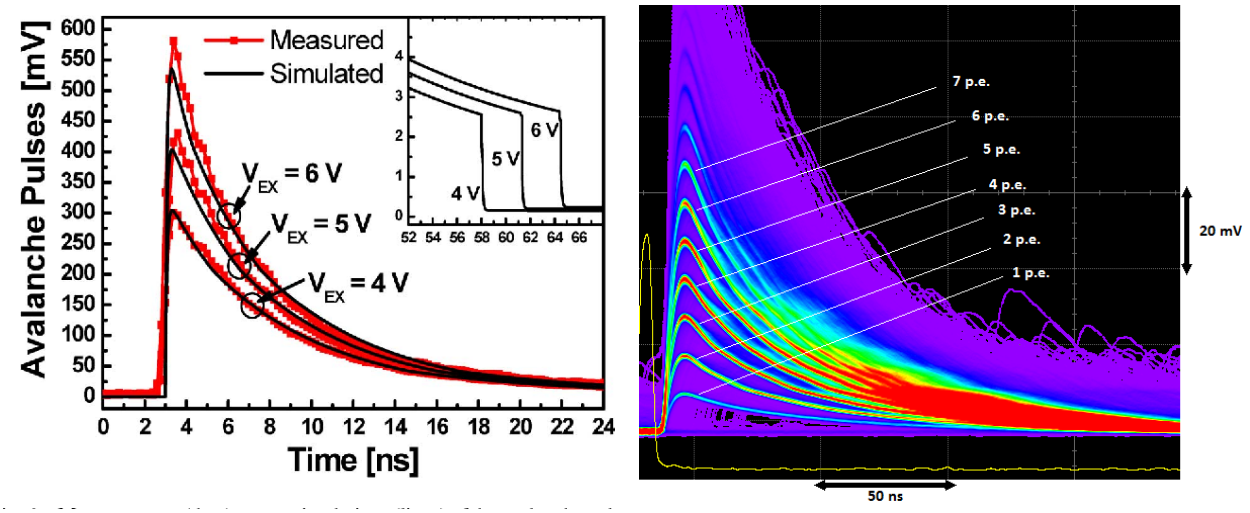
To overcome this lack of proportionality, the Silicon Photomultiplier (SiPM, also labeled Multi-Pixel Photon counter or MPPC) integrates a dense array of small, independent SPAD sensors, each with its own quenching system (Figure 3). Each operating unit of SPAD cell (or micro-cell) and quench resistor function independently from the others and when a cell in the SiPM fires in response to an absorbed photon, a Geiger avalanche is initiated causing a photocurrent to flow through the individual cell. Hence, the Geiger avalanche will be confined to the single cell it was initiated in. During the avalanche process, all other micro-cells will remain fully charged and ready to detect photons.



(a) (b)

Figure 3: **(a)** Depiction of the micro-cell structure of the SiPM surface. The SPAD micro-cells with their own resistors are connected in parallel to each other by a common electrode forming a SiPM. From [9] **(b)** Simplified circuit schematic of a SiPM showing each micro-cell which is composed of the SPAD, quench resistor and fast output capacitor in this case that allows the sensor to resolve fast pulses. All micro-cells forming the SiPM have a common output. From [17]

To understand this difference in proportionality, Figure 4 shows a typical avalanche pulse of the detector for a single SPAD and for SiPM. Figure 4a represents the time domain pulse waveform when a single cell is fired, the amplitude is proportional to the applied excess bias (overvoltage) but is always the same (defined as one photoelectron or p.e) regardless of the incident number of photons striking the detector at the same time. The inset shows the current vs voltage curve during quenching; the current drop suddenly below a certain value I_q when the avalanche is no longer self-sustaining (see subsection 3.4). On Figure 4b is shown the same waveform but with varying amplitude proportional to the number of photoelectron sensed by the SiPM. That way, the signal amplitude is proportional to the number of micro-cells fired simultaneously and allows to know the photon counts up to a certain limit given by the number of cells in the SiPM.



(a) (b)

Figure 4: **(a)** Pulse produced by a micro-cell in response to photon absorption (single cell signal), its amplitude is defined to be 1 p.e. The figures also shows the output pulse increase with excess bias as well as the avalanche stop when the current fall below a certain value (I-V cPulse produced by a micro-cell in response to photon absorptionurve in the inset). From [18] **(b)** Pulse produced by a SiPM when 1, 2, 3,... pixels are fired simultaneously in response to low incoming light (acuired by an oscilloscope). The amplitude is proportional to the numbers of micro-cells fired, thus equal to $N\times$ p.e .From [19]

All the micro-cells of the SiPM structure on Figure 3 have a common output and will produce signal as shown on Figure 4b, this type is called analog SiPMs where the number of counts is given by the amplitude of the sum of the single fired SPADs. More recently, digital SiPMs (dSiPMs) have been developed that allows to count each fired SPAD separately. The key difference is that each cell is connected to its own readout electronic block that also contains its own quenching and recharge circuit [20]. The dSiPM is arguably an improvement over its analog counterpart as not only it gives the number of micro-cells fired but also the possibility to know which ones. More than that, photons are detected and counted as digital signals, thus making the sensor less susceptible to temperature variations and electronic noise. On the other hand, the complex circuitry embedded inside the dSiPM means a reduced probability of detecting a photon since the photoactive area is reduced.

2.2 Properties and performances

When characterizing the SPAD or SIPM, it is important look for a given set of properties to get further insight into its operation and assess the device performances. The important parameters are reviewed here while in subsection 2.4 is summarized the way to combine these characteristics in order to get sensible figure of merits and section 3 goes further in depth for some of these properties.

2.2.1 Photodetection probability

As the purpose of the device is to detect incoming photons, it is essential to find a way to quantify this property. The photodetection probability (PDP) is the avalanche probability of the device in response to photon absorption at a given wavelength and can be defined as the ratio between the number of detected photons and the photons arriving at the detector [21]. Indeed, not all photons arriving on the photoactive area will result in the junction Geiger breakdown; it could happen that the photon is reflected at the surface of the sensor or on the contrary passing through the device without being absorbed. Even if it is absorbed within the substrate generating an e-h pair, at least one of the carrier has to drift to the multiplication region

before it recombines. Additionally, once in the depletion region where the field strength is strong enough to start an avalanche breakdown, this process itself is stochastic and has a finite probability of happening.

Due to the laws of light-matter interactions, photons with shorter wavelenghts are unlikely to travel very far into semiconducting materials while longer wavelengths are absorbed deeper. Hence, SPADs can be made from shallow junctions or deeper junctions: SPADs fabricated with shallow active regions favor the detection of photons absorbed near the silicon surface, so these devices are suitable for near UV-blue incident light. For a deep junction SPAD structure the multiplication region is located deeper into the substrate and is suited for detecting red to near infrared light. The PDP is expressed in % at different wavelengths, usually ranging from about 400 nm to 1000 nm as can be seen on Figure 5. From the graph it is possible to get the maximum PDP at a given wavelength as the peak of the curve. Although the PDP value itself depends on the applied bias, the peak wavelength is always the same for a given SPAD (.e.g. 15 % at 430 nm for an excess bias of 100 mV on the figure). It is also sometimes useful to quantify the spectral response range that could be losely define as the spectral width of which the PDP is above a given % value (.e.g. on the figure the spectral response range from about 380-600 nm above 5% PDP).

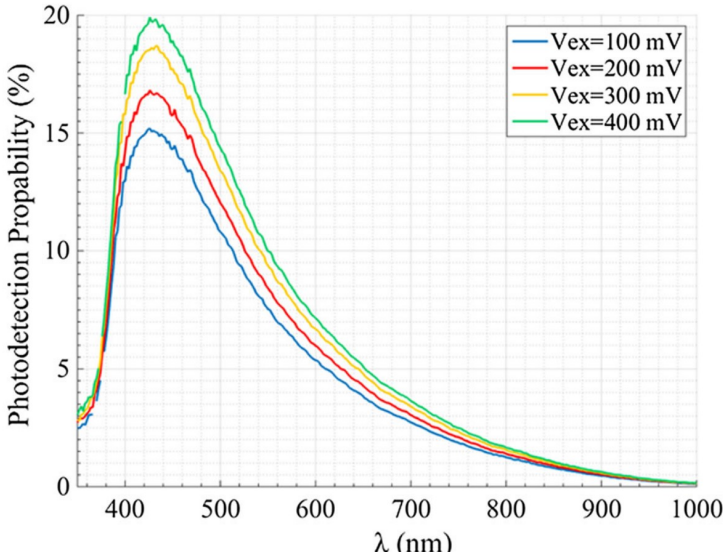


Figure 5: Example of a PDP curve for a SPAD at different applied bias with 430 nm peak PDP and a spectral range of 380-600 nm above 5% PDP. From [22]

Strictly speaking, the PDP is used only when characterizing a SPAD active area. For the dead space region and SIPMs, another important factor needs to be taken into account which is the fill factor (FF). The active area of a SPAD is the central photon-sensitive portion of the detector, however the impinging photon could land on the dead space (inactive part) of the SPADs array, in between two micro-cell for example. The electric field strength should be consistent across the photoactive part of the structure and especially not higher at the edges, so as to yield a homogenous breakdown probability. The fill factor is simply the the ratio of photo-sensitive area to total imaging or pixel area. The combination of PDP with FF gives the photodetection efficiency (PDE). Sometimes, it is found that the PDP is subdivided into the quantum efficiency (QE) and the avalanche triggering probability P_T [9], [21]. The quantum efficiency includes the probability of a photon to enter in the device (without being reflected) and to be absorbed in the part of the device where it has a chance to induce a Geiger breakdown. As such, the PDP and PDE can be expressed as:

$$
\begin{aligned}
PDP &= QE.P_T, \\
PDE &= QE.P_T.FF.
\end{aligned}
\tag{1}
$$

2.2.2 Dark count rate

If a photon striking the single SPAD photoactive area may not be detected, the other case could happen as well, .i.e. an avalanche breakdown might occur inside the device and a fake event could be detected. Even when there is no light such event could happen and this is referred as a dark count is measured by the dark count rate (DCR) in [cps] (counts per second) or [Hz].

Dark counts may happen for several reasons: the main contribution at room temperature is usually given by the Shockley–Read–Hall (SRH) mechanism, which basically consists in trap-assisted generation of electron–hole pairs caused by defects inside the device. Thus in order to avoid these fake counts, it is essential to have as much control as possible on the fabrication process as impurities are the cause of traps centers. The other cause of dark counts is a pure band-to-band tunneling process, which depends almost entirely on the electric field strength in the depletion region. The temperature at which the device operate has an important influence according to SRH theory [23]. When increased, the traps are more active and the DCR is enhanced but when lowered; the contribution of the SRH process drops quickly and the DCR may become dominated by tunnelling. This is quite undesirable because the DCR induced by tunnelling exhibits only a weak dependence on the temperature and therefore it limits the effectiveness of cooling the device. Fortunately, it is possible to reduce this tunnelling effect by tailoring the electric field accordingly, for example by avoiding high doping concentrations and in conjunction with smaller junction dimensions that implies higher electric field strength in the multiplication region and greater contribution to the DCR by tunneling mechanisms. Since the absolute DCR is proportional to the photoactive area [9], it is also common to normalize the DCR to the detective surface expressed in [cps/μm^2].

2.2.3 Afterpulsing

Afterpulsing is a particular type of dark count that identifies an avalanche pulse after a primary event (photodetection or primary dark count). During the avalanche process, the carriers flowing over the depletion region have a certain probability to be captured by trapping centers. These captured holes or electrons have a certain lifetime and when released during the SPAD recharge phase to excess bias voltage they may also initiate an avalanche event which is called afterpulsing. Once again, a clean device free of impurities will prevent this detrimental effect for happening. Also, the temperature is important since it regulates the trap activity but this time operation at higher temperature is better because the trapped carrier will be released quicker during the recharge phase when the device is inactive.

In some occasion, afterpulsing can also be "optically induced" (not related to traps). During the avalanche process, hot carriers produce secondary photons in the visible range and some of them might be reabsorbed in the same SPAD cell beneath the active region [21]. Then, the photogenerated carrier has the opportunity to diffuse to the multiplication region and trigger another avalanche.

2.2.4 Crosstalk

Crosstalk is another parasitic count detrimental to the diode operation that needs to be reduced as much as possible and is relevant for SIPMs only. Essentially, crosstalk happens in an array when an avalanche event taking place in a single SPAD induces other breakdowns in the neighbouring micro-cells. Crosstalk can be optically or electrically induced and the different mechanisms are represented on Figure 6. Optical crosstalk arise from the electroluminescence effect: it is estimated that during breakdown there are in average three photons emitted per 10^5 carriers which can subsequently land into the neighbouring cells and cause others breakdown [24]. As shown on the figure, a direct and indirect optical path exist. In the first case the

secondary photon travels in the shortest way possible to the neighbouring cell while in the second case, it is reflected first on some surface (e.g. a metallic protective layer). Crosstalk can also be electrically induced, in that case some of the carriers generated during the avalanche may exit the depletion region and diffuse laterally, eventually reaching the depletion region of another SPAD. Crosstalk could also be both optical and electrical: the secondary photon is absorbed in the substrate or neutral region generating a carrier that diffuses to the avalanche region. This process is slower than the direct and indirect path so it is often called "delayed crosstalk" while the other two are labelled "prompt crosstalk". Crosstalk can be atttenuated by ensuring that each cell is isolated from one another. Some methods use for this purpose is using trenches with a metal coated sidewalls [25] or V-shaped grooves in between pixels for direct optical or electrical crosstalk as well as adding a second p-n junction between the substrate and active layer preventing delayed crosstalk and adding light-absorbing layers created in the bulk and/or on the substrate backside for the indirect optical component [26].

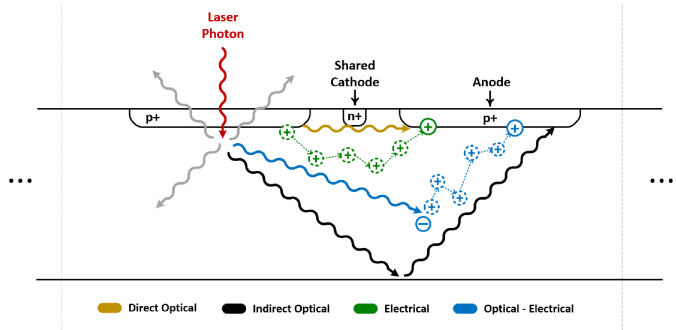


Figure 6: Representation of direct (1) and indirect (2) optical crosstalk, electrical (3) and delayed crosstalk (4). For simplicity, the SIPM is reduced to two neighbouring SPADs represented by a simple p-n junction. From [24]

2.2.5 Breakdown voltage

The breakdown voltage is the bias voltage where Geiger-breakdown occurs and can be more rigourously defined as the voltage where the multiplication factor M (see subsubsection 3.1.1) diverges [27]. In practice, the breakdown voltage can be identified on a reversed I-V curve where it starts to rise very steeply (Figure 7). This value depends on the internal structure of the diode and particularly on the doping profile at the p-n junction controlling the width of the depletion region. It is also dependent on the temperature, especially for wide depleted region [21]. The breakdown voltage is higher for a wider depleted region, this can be beneficial since the peak electric field in the avalanche multiplication zone is reduced which decreases dark counts induced by tunnelling but lower operational voltages are often preferred for CMOS SPADs and SIPMs.

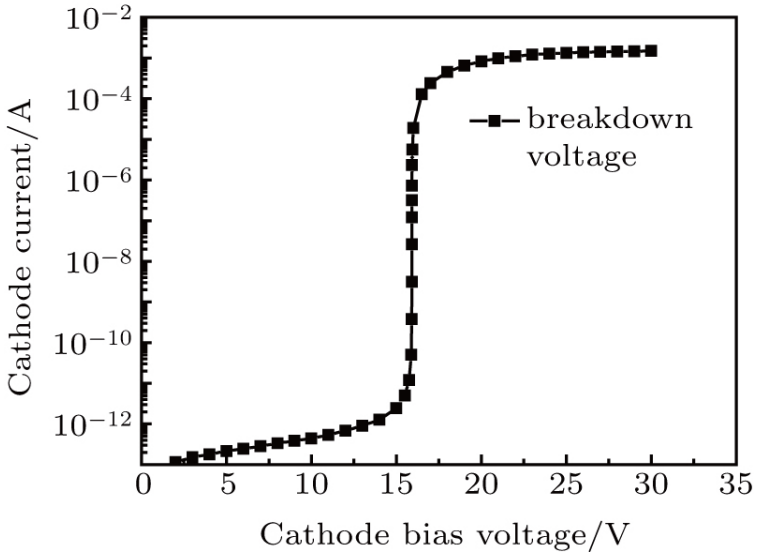


Figure 7: Typical reversed I-V curve for a single SPAD device, the breakdown voltage is approximately at 16.1 V of applied reverse bias. From [27]

2.2.6 Timing jitter

Timing jitter is an assessment of the device timing statistical fluctuation and is the temporal error interval of photon detection. Generally, the photon detection shows different timing performance when absorbed in different regions (depletion or neutral region). We can identify two main mechanisms contributing to the SPAD jitter: first build-up time and lateral propagation (multiplicated assisted diffusion) when photon absorption occurs in the depletion region and second, the transit of the carriers from the neutral to the depletion region. Some authors count absorption of secondary photon triggering an avalanche as part of the jitter characteristics [12] but this is arguably more an afterpulse phenomenon as presented in [21] since it is an unwanted event to be detected.

The multiplication assisted diffusion process contributes to the jitter for two reasons: while fast, the avalanche build-up process of impact ionization and diffusion inside the depletion region is probabilistic in nature and on the other hand, the geometry of propagation depends on the position of photon absorption (e.g., in the center or at the edge of the detector), which is random as well [28]. The build-up and lateral diffusion jitter decreases for higher overvoltages due to more efficient triggering and propagation mechanisms .The other cause of jitter is the diffusion of minority carriers created in the neutral regions that bring a random delay for avalanche triggering. This is accentuated for thick SPADs that have a wide diffusion zone because depending on its wavelength, the photon could be absorbed near the depletion region, at the end of the neutral region or anywhere in between.

The statistics of multiplication assisted diffusion induces a fast peak response (almost no timing jitter) while minority carriers diffusion in the neutral regions on bring more severe delay and worst jitter. Thus, the timing jitter modeling can be divided into two parts [29] shown on(Figure 8): a Gaussian peak for the avalanche build-up mechanism and an exponential tail for the pure diffusion mechanism allowing to measure these two components separately.

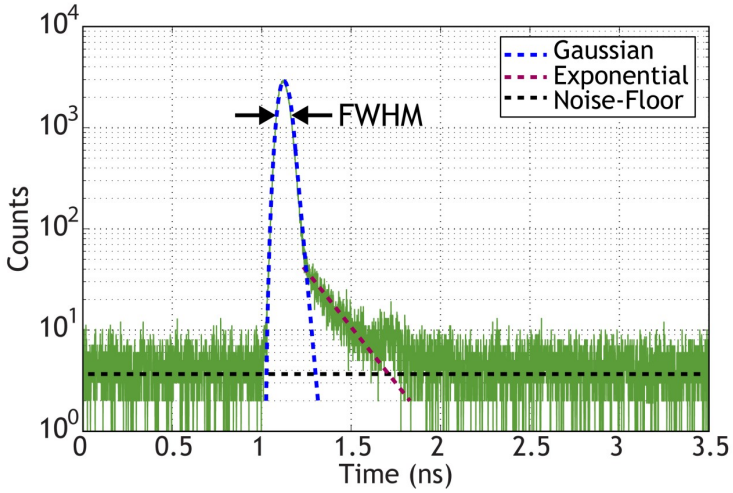


Figure 8: Jitter measured as the delay time with respect to a precise and punctual light source. The build-up and lateral diffusion part in the depletion region is characterized by a Gaussian curve with a certain FWHM and the exponential tail characterizes the jitter due to diffusion in the neutral region. From [16]

2.2.7 Dead time

The dead time is the duration after the start of the avalanche where the SPAD will be inactive and is to be minimized for enhancing the count rate. It is the sum of sensing (avalanche build-up), quenching (discharge phase) and hold-off or recovery (recharge phase) times highlighted on Figure 2. The two first phases are quite fast (pico and nanosecond order) but the SPAD recharge is longer. However, during this phase the cell is not exactly inactive but a photodetection event happening during this step could paralyze the pixel (see subsection 3.4). Thus a minimal hold-off time can be set with a more elaborated active quenching which at the same time accelerate the quenching discharge phase [30].

2.3 Design features and types of sensors

Overall, SPAD technologies can be classified into two main types: custom technologies and standard/CMOS compatible technologies with opposite benefits and downside. One one hand, custom technologies are purposely designed for the optimization of the detector performance but are difficult to integrate into a dense array, so they are not suited as elementary operation unit for a SIPM. On the other hand, standard technologies have the main advantages to be compatible with CMOS design and fabrication processes, making possible their integration into SIPM arrays with millions of pixels (essential for imaging applications) which increases count rate. Moreover, it can be combined with state-of the-art circuitry that improve performances and control [12]. However, the evolution of CMOS circuits demands not necessarily similar from the ones imposed by the optimization of the SPAD performance. Thus making the SPAD CMOS compatible add constraints on its design and reduces some important characteristics performances.

2.3.1 Custom technologies

Nowadays, custom SPADs can be divided into two main families: reach-through and planar thin SPADs shown on Figure 9. As for the main design features, the essential components of a SPAD are the guard ring, the overall shape (i.e. circular, square, others), active area diameter and the diode junction. Here, the junction itself is a n-type shallow implant with p-type enrichement layer beneath (n on p type).

For the planar thin type (Figure 9a), the p-n junction is made by shallow n+ phosphorous diffusion for the cathode, while the p-enrichment layer in this case is made by ion implantation. The p-thin quasi intrinsic

layer ($\sim 3\mu$m) extends the absorption region beneath the enrichement layer; minority photoelectrons created in this region will diffuse upwards and and reach the depletion layer where they might trigger an avalanche. For efficient carrier collection of holes, the p+ burried layer along with the p+ sinker, minimizes the series resistance of the SPAD toward the anode terminal [31]. The n+ substrate in this case plays a very important role for obtaining a fast temporal response with minimal jitter. As reported in [32], [33], as the reverse bias is increased, a depletion layer is also formed at the n+p+ junction and at sufficiently high overvoltage the two depletion layers merge which allows a jitter almost free of its diffusion component. Moreover with the n+ sinkers surrounding the active area, the SPAD is electrically isolated reducing crosstalk in the case of realization of detector arrays. However, the number of thin SPADs that can be integrated in an array structure is currently low and the design requires full customisation of the manufacturing process [8], hence why it is not considered a standard technology. For the guard ring, Instead of reducing the field in the outer region, the enrichement p-layer is actually highly doped for enhancing the electric field in the central region [31]. Thanks to its thinness and shallow p+-n+ junction, this design allows a high count rate they can achieve (up to 160 Mcps), a lowered DCR (a few cps) and afterpulsing probability ($\sim 5\%$) as well as a low timing jitter (down to 32 ps) [12]. However, the reduced thickness also limits the extension of the drift region, thus these thin SPADs exhibit decreased photodetection performances that is guaranteed by thicker structures and are suited for light detection on the blue and close U.V spectrum.

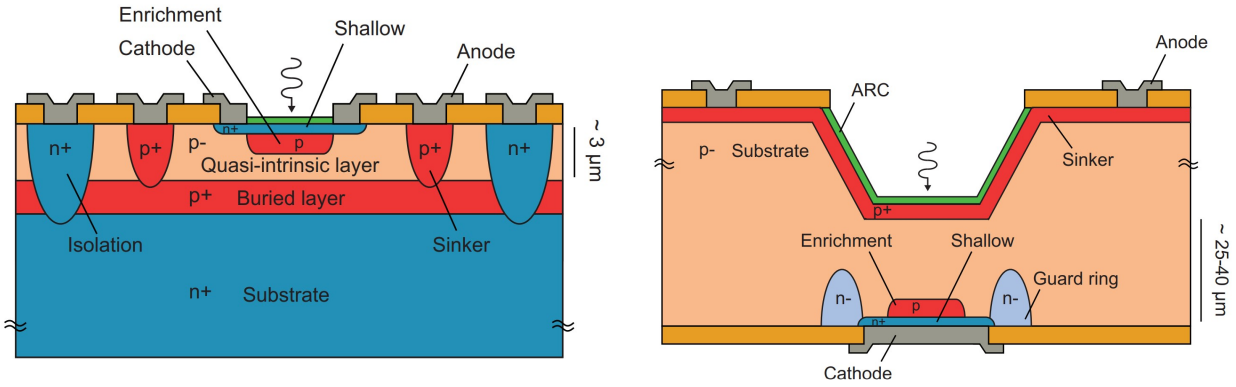


Figure 9: **(a)** Structure of the planar thin custom SPAD with n+ implant on p-enrichement layer shallow junction. **(b)** Structure of the custom backside illuminated reach-through SPAD sensor with n+ implant on p-enrichement layer, n-diffused guard ring and antireflection coating layer (ARC). From [12]

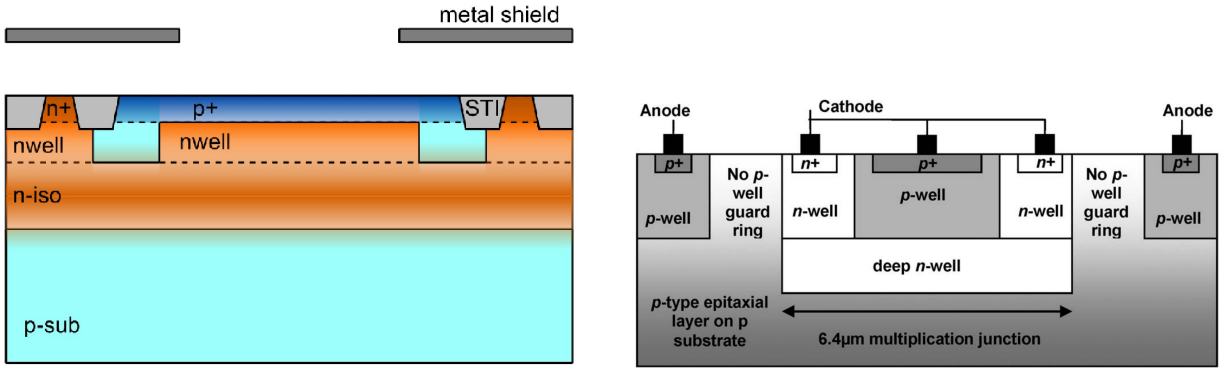
The reach-through structure on Figure 9b is special type that is illuminated on its backside, therefore called backside illuminated (BSI). The diode junction is also a n-type shallow implant on a p-type enrichement layer. This p-boron diffusion layer helps to tailors the electric field in the substrate and helps to prevent premature edge breakdown. The main feature stopping edge breakdown however is the lightly doped n-phosphorous guard ring that reduces the electric field on the edges. When a photon is absorbed, carrier multiplication will take place in the n+/p depletion zone but the absorption region extends into the thick p-substrate where the photoelectrons can drift to the avalanche region if they do not recombine. Once again, electrons are collected at the cathode and holes at the anode on the opposite side, hence why the p+ low resistivity sinker at the bottom of the substrate. An antireflection coating layer (ARC) is also present to prevent reflection and improve photodetection. This structure is tailored to maximize the PDE for a wide spectral response range (.e.g. 85% of max PDE ranging from 400 nm to 1100 nm for the commercially available SAP500 [34] with low dark counts and afterpulses if cooled down sufficiently. But its thick absorption region limits the detector performances in other areas: high timing jitter because of increased drift time uncertainty, a higher breakdown voltage means higher power dissipation too per detection event. High

power dissipation, along with the complexity of such structure, has prevented the fabrication of reach-trough SPAD arrays [12].

2.3.2 Standard CMOS technologies

As mentioned earlier, the use of CMOS technology does not completely allow to fully optimize the SPAD performances due to the fabrication constraints. Consequently, a higher DCR and lower PDE are expected compared to fully custom devices. However, the main advantage of CMOS fabrication is its accessibility at low cost for prototyping. The constraints generrally results in miniaturization of CMOS devices towards shallow junctions at few hundreds of nm from the surface, high doping concentrations of the well implantation, and low supply voltages [12]. Nowadays, fabrication processes use the 0.35 μm down to 65 nm CMOS technology node for SPADs and SIPMs.

A design in 0.15 μm CMOS technology with typical features is depicted on Figure 10a, similar structures can be found in [35]–[38]. The n-isolation layer separates the n-well from the p-substrate. The main difference with the other designs reviewed so far, is p+/n-well the junction. This is preferred if an integrated readout is to be fabricated, because low voltage circuits can be implemented [8], eventhough a n+/p-well structure is also feasible. The choice between the two types of junctions will also leads to different PDE, the p+/n-well has a narrower spectral range but is more efficient for detecting wavelengths around 400 nm [21]. Due to the CMOS process, the n-well has a large doping concentration which causes a reduction of the breakdown voltage and a higher DCR due to tunnelling because of the narrower depletion region [36]. The shallow trench isolation oxide combined with the low doped p-type ring surrounding the junction act as the guard-ring. STI guard ring is suitable with CMOS processes and allows to increase the fill factor, however the high amount of defects at the interface enhances the DCR and afterpulsing [8]. A solution to mitigate this effect is to employ a p-type passivation layer surrounding the STI [36].



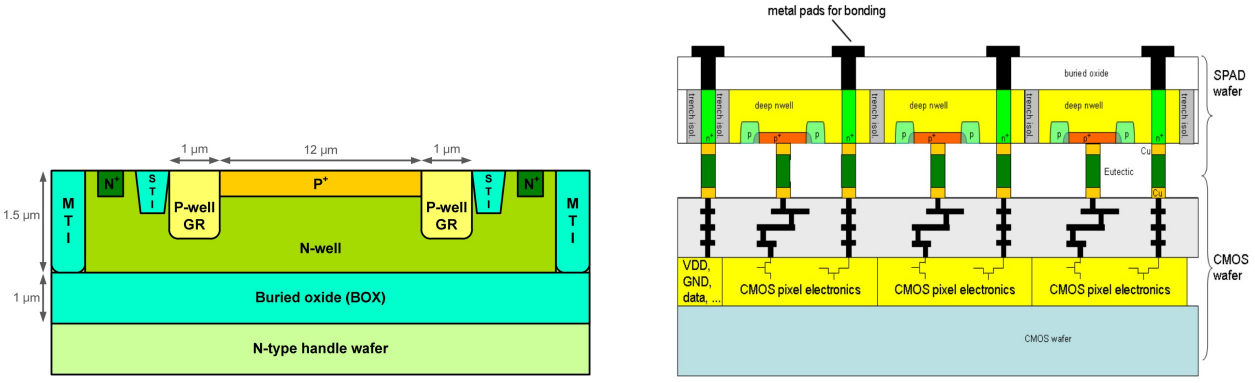
(a) (b)

Figure 10: **(a)** Cross section of a standard CMOS technology shallow junction structure. From [39] **(b)** Cross section of a CMOS BSI device with improved PDE in the near infrared region. From [40]

Increasing the PDE at the longer wavelengths can be achieved by burying the depletion layer of the detector deeper into the substrate. The main idea with this type of design is shown on Figure 10b and is also implemented in [41], [42]. Here, the high-field multiplication region is the junction between the deep n-well and a thin p-epitaxial layer with no isolation from the p-substrate. The shallow p-well is placed on top of the deep n-well so that short-wavelength photoelectrons generated in this part of the device have also the opportunity to reach the multiplication region. The guard ring is formed by using the retrograde doping profile of the deep n-well so that a sharp doping gradient at the boundary with the p-epitaxy on the side is

formed (retrograde deep n-well guard ring) [40]. That way, the peak photodetection probability is shifted around 600-700 nm but efficiency for the blue part of the spectrum is reduced.

The use of SOI technology was not really considered for SPADs until recently [43] and this type of device was first fully characterized in [44]. The use of a SOI wafer originates from the implementation of 3D-stacking schemes where the SPAD cell and circuitry are separated into two distinct tiers [45]. This allows firstly the optimization of the of the SPAD cell design and electronics independently but the separation also results in a lower fill factor and better sensitivity. However, this requires the SPAD to be illuminated from its backside meaning that similarly to the structure in Figure 9b, the photon has to strike the bottom of the substrate first while the effective diode junction is located at the opposite end. While this might be suitable for the red, near IR range, backside illumination done as such on a bulk silicon wafer suffers from poor PDP in the blue and violet region, even with backside thinning. This issue can be solved by adding a burried oxide layer below the SPAD structure using standard SOI technology. The bulk silicon layer can be removed by thinning the backside and subsequently performing isotropic etching where the burried oxide (BOX) is used as the etch stop layer [43]. This results in nearly transparent layer at the backside that is much thinner and the carriers generated close from the surface can reach the avalanche region more easily. This SPAD structure is shown on Figure 11a. Finally, all the readout circuitry is fabricated separately on another wafer and the SOI SPAD array is flipped down facing the circuitry and the two are bonded together. The final results is depicted on Figure 11b.



(a) (b)

Figure 11: **(a)** Cross section of the fabricated SOI CMOS SPAD with burried oxide layer. From [44] **(b)** Depiction of the SIPM design showing the SOI CMOS SPAD tier bonded to the CMOS electonics bottom tier. From [43]

2.4 Applications and FOM

SPADs and SIPMs are used for different purposes but the possible applications do not all share the same requirements and appropriate figure of merits (FOM) have to be established to quantify the device performance and optimize the design for its targeted goal. Overall, these applications can be classified in three main groups: photon counting, photon timing and photon imaging applications [46].

Typical applications requiring precise photon counting and optimized PDE are quantum key distribution (QKD) and quantum photonics [12]. Often, what is important is QKD and quantum photonics applications in general (boson sampling, quantum memory, applications related to quantum computation) is the optimization of photodetection efficiency. For instance, QKD relies on decoding a sequence of bits using photon polarization with two SPADs [47] and what is really critical for the detector is to catch the photonic bit-strings accurately and therefore a high PDE , low DCR and aterpulsing is needed. The SPAD system

does not necessary need to have a high count rate or low jitter since the light source can be controlled, although it is better for higher throughput. As such, custom SPADs with thick junctions as in Figure 9b fit these needs perfectly, because it is not essential to have large arrays and high counting rate.

An example of application with accurate photon timing would be Raman spectroscopy that can be used to provide data on the chemical composition and molecular structure of a compound [48]. This spectroscopy method relies upon inelastic scattering of photons from a monochromatic source (usually a laser), known as Raman scattering. The laser light interacts with molecular vibrations in the system, resulting in the energy of the laser photons being shifted up or down. It is typically used to determine vibrational modes of molecules and provide an IR spectrum by which molecules can be identified. One main problem is that it can be overshadowed by the sample autofluorescence itself that comes right after the useful but weaker Raman signal. One solution to avoid picking up this fluorescence signal is to operate the device in fast time-gating operation in which the photon detection is required only for a short time window after a pulse excitation, and the time window has to be precisely synchronized with the Raman excitation. Thus the SPAD or SIPM design must emphasize a high time resolution, low jitter and dead time in fast gated mode which is mainly part of the front end electronics [49].

It is possible and essential in some cases to use SPAD technology to make SiPM-based camera for acquiring 2D or 3D images [13]. One important example is positron emission tomography (PET).

PET is a nuclear medicine technique used to generate images of the body from the physiological point of view. In a typical setup PET relies on the use of weakly radioactive tracers, which concentrate in glucose-avid molecules, where they emit a positron when they decay. The latter annihilates at very short range and two gamma photons of 511 keV each are emitted in opposite direction (180 °apart) [14], [50]. This energy corresponds to a wavelength in the picometer range, so it is converted to visible photons with scintillating crystals arrays, placed in a ring surrounding the region to be imaged. The PET scanner needs to detect both emitted photons in coincidence to establish the line of response (LOR) along which the annihilation took place. After millions of LORs are acquired, a tomographic 3D image of the subject is formed, revealing the tracer concentration.

2.4.1 Performances and Figure of merits

Based on [46], [51], it is possible to establish a few useful FOM for the different type of applications that can be all solely expressed with the parameters discussed in subsection 2.2.

Photon counting: In photon counting, the detector is exploited to measure the intensity of slowly varying optical signals and the performance can be described using the signal to noise ratio (SNR) and the maximum detectable signal ($SMAX$).

The SPAD is able to detect one quantum of light so the SNR has to be defined in terms of the counting statistics related to the device, as some ratio of mean to standard deviation. For example, if the number of dark counts over a given interval is too high and the number of impinging photons too low, the output signal will be mostly made of unwanted detected events resulting in a low SNR. The signal itself is defined as the number of photons counted at the end of a measurement with integration time T_{int} and the noise is given by the fluctuation in counts of the source (assumed to follow Poisson statistics) in addition to dark counts (also Poissonian) [52], so the SNR can be written as:

$$SNR = \frac{PDP.\Phi_s.T_{int}}{\sqrt{PDE.\Phi_s.T_{int} + DCR.T_{int}}} \approx \frac{PDP.\Phi_s.T_{int}}{\sqrt{DCR.T_{int}}} \tag{2}$$

where Φ_s is the signal photon rate. Hence, for a given integration time the SNR is increased if the device as a high PDP and low DCR.

Moreover, in photon counting applications it is desirable to have a high maximum $SMAX$ detectable signal. $SMAX$ quantify the detector saturation or maximum photon counts achievable for a given integration time. It thus depends on the dead time T_{DEAD} but also on the afterpulsing probability P_{AP}. With these considerations, we get for $SMAX$:

$$SMAX = \frac{1 - P_{AP}}{T_{DEAD}} T_{int} \tag{3}$$

Furthermore, the SPAD active area should also be included in the FOM, mainly because the DCR is proportional to the photoactive area and it is common to find in tables and graphs units expressed in DCR/μm^2. The square root of the area is taken in order to not overestimate its contribution to the DCR.

Considering that both $SMAX$ and SNR increase with improved performance, we can compute a figure of merit for photon counting purposes:

$$FOM_C = PDP\sqrt{\frac{Area}{DCR}}\frac{1 - P_{AP}}{T_{Dead}} \propto SNR.SMAX.\sqrt{Area} \tag{4}$$

where an integration time $T_{int} = 1s$ is taken. Since the PDP depends on the photon wavelength, we could select this wavelength according to the targeted application or an average PDP value could be used as representative over the range of interest or simply slecting the Peak PDP value. The dimensions of FOM_C is given by any unit length since SNR and $SMAX$ are dimensionless.

Photon-timing In photon timing applications, we want a device that can reconstruct very fast optical signal in the ps range and where one or few dozen independent detectors usually suffice.

We need to consider the Gaussian FWHM of the jitter histogram curve (Figure 8) which should be as narrow as possible for precise timing purposes. The SNR in this case is given by the peak value of the Gaussian distribution and the standard deviation of background level noise in Figure 8 [46] and considering that $SMAX$ is still to be maximized as well as $\sqrt{Area}$, it can be shown:

$$FOM_T = PDP\sqrt{\frac{Area}{DCR}}\frac{1 - P_{AP}}{T_{Dead}}\frac{1}{FWHM} = \frac{FOM_C}{FWHM} \tag{5}$$

Thus the requirements for the photon counting FOM are still taken into account but in addition, the timing jitter should be as low as possible which is quantified by its Gaussian FWHM. The latter is expressed in time units so FOM_T cane be expressed in m/s for example.

Photon imaging For photon-imaging applications, we want to acquire two or three dimensional images and thus the detector requires large arrays of pixels like SIPMs. The crosstalk and mainly the fill factor play then an important role. The sensor primary goal is still to count the number of incoming photons, so the FOM for imaging applications can be derived from (6) and further adding three to four fundamental parameters for SIPM based cameras: the fill factor FF, number of pixels N, maximum frame rate f_{max} and additionally crosstalk CT but this value is not always reported in literature [46]. As for cameras, the frame rate quantify the number of successive images the SPAD array can acquire per second and relies heavily on the shutter type and front-end electronics [53]. As such, the electronic readout circuitry design is of paramount importance for dSIPMs used for imaging purposes.

The figure of merit for photon imaging applications is therefore given by:

$$FOM_I = PDP\sqrt{\frac{Area}{DCR}}\frac{1 - P_{AP}}{T_{Dead}}.FF.N.f_{max}.CT = FOM_C..FF.N.f_{max}.CT \tag{6}$$

The timing jitter is not included for this FOM but several figure of merits can be evaluated for the same device; with PET for example where precise coincidence detection is important, FOM_T should also be estimated.

2.5 SPAD comparisons

Some state-of-the art SPAD devices are discussed using the FOMs stated previously to evaluate the current performances and to compare them later with the novel design developed at UCLouvain. For an easier and fair comparison, the SPADs presented here are designed to target the blue, UV, sub-UV wavelength range and it is the performance of the photoactive area that is estimated, without taking the fill factor into account as in (6) and (5). The selected devices comes from various background: some designed recently that are still in development but have been characterized, others a bit older but that exhibit good performances for the targeted applications and some commercial devices available in the market that are also implemented into arrays. The performances are summarized in Table 1. If a given SPAD does not disclose a certain parameter, the median value is taken based on the other available data.

Ref	**Technology**	**Peak PDP**	**DCR (RT)** [cps/μm^2]	**PAP**	**Tdead** [ns]	**Jitter** [ps]	V_{BD} [V]	V_{ov} [V]	**FOMC** [m]	**FOMT** [m/s]
[45] (2019)	0.14 μm CMOS BSI SOI	26.4 % (490 nm)	396.1	<0.1%	200	65 ps (405 nm)	11.3	3	0.066	1020
[54] (2020)	65 nm CMOS	50% (480 nm)	73	1%*	55*	158 (485 nm)	9.6	1	1.05	6666
[55] (2018)	0.18 μm CMOS	55%(400 nm)	0.23	1%*	100	23 (410 nm)	22.1	2	11.3	493635
[56] (2016)	NA	51.2%(400 nm)	0.75	5%	100	160 (420 nm)	26	3	5.62	35103
[57] (2014)	0.35 μm CMOS	55 % (420 nm)	0.079	4%	20	75 (780 nm)	25	6	94	1252359
[58] (2016)	NA	54% (450 nm)	0.55	1%*	55*	75*	53	3	13.1	174753
[59] (2017)	NA	50.5%(420 nm)	0.5	0.75%	45	75*	24.5	2.5	15.8	210686
[60] (2021)	NA	67% (420 nm)	0.71	<1 %	55	75*	32.5	6	14.5	192762

Table 1: Comparison of the SPAD performances for timing and counting applications targetting lower wavelength ranges. Because the PDP and DCR depends on the excess bias, the typical operating overvoltage is also stated.

We can notice first that the afterpulse probability is fairly low for all devices and plugging P_{AP} in (6), (5) has barely any influence on the FOMs. This is the case for most SPADs nowadays, over the years afterpulsing has beeen reduce to irrelevance thanks to the ability to control the dead time and to cleaner fabrication procedures.

The two first devices [45], [54] are the most recent and display the poorest performances, mainly due to their high DCR density characteristic. This imply that the fabrication process has room for improvement but above all these devices have the lowest breakdown voltage meaning a higher peak electric field and enhanced dark counts by tunnelling. For the BSI SOI SPAD, it also has the lowest PDP about a half of the other devices. For this design however, it is to be considered that it is in its earliest stages as well as the only one that is back-illuminated and in all fairness, it has decent properties compared with other BSI SPADs in standard CMOS technology and the the potentially high fill factor could give a high final PDE once a SIPM will be made.

The next devices of [55]–[57] show performances that are obviously better, mainly because the DCR density is much lower, the main reason being the high breakdown voltage that widen the depletion layer width and reduces the peak electric field strength. Impressively high characteristics are obtained for the design of [57] which performances also can be compared with custom SPAD designs. In this case the use of a 0.35 μm high-voltage technology with high level of cleanness and controlled substrate to reduce the number of defects allows to get a low DCR density. Moreover, the SPAD achieves a low dead time owing to a mixed passive acting quenching circuit.

Finally, in [58]–[60] we have high performances SPAD arrays that set the state-of-the-art devices available on the market that can be used for biophotonic applications. The PDP is again close to 50% and goes up to

67% for the SIPM of [60]. As for the previous devices with a high breakdown voltage, the DCR density is less than 1 cps/μm^2 and the afterpulsing negligible. Oddly enough, the jitter parameter is never given which is however important for high-performance timing applications as claimed by [59]. Therefore, the median value was taken.

3 Theory of the SPAD cell operation

The correct analysis of the characteristics and phenomena associated with the device under study requires a thorough understanding of its physics. In addition, it allows to make appropriate design choices consistent with the final application and make simulations saving both time and money.

Therefore, the subject matter of the SPAD physics that are relevant for later discussion of the device under study is presented in this chapter. Topics for this section include impact ionization and avalanche breakdown, absorption and photodetection probability, dark counts and finally quenching.

3.1 Impact ionization and breakdown

Impact ionization of free carriers and avalanche breakdown take place at high electric field strength in a diode in reverse bias mode. A free carrier injected in the depletion region can acquire enough kinetic energy from the field to ionize a silicon atom during a collision and create a secondary electron-hole pair. This new pair may in turns induce other e-h pairs by impact ionization and multiplication of free carriers occur in an exponential manner. This phenomenological description of avalanche breakdown is schematized on Figure 12.

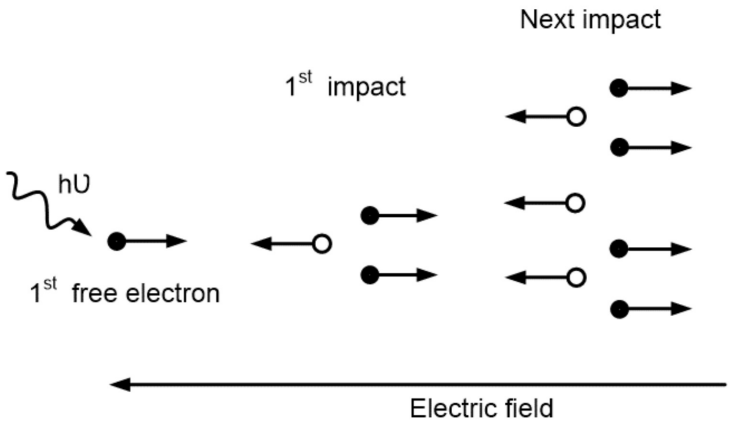


Figure 12: Impact ionization causing avalanche current in the presence of a high electric field where a number of free carriers are multiplied in each step. The electron injection in this case is due to photon absorption. From [61].

According to theory, important parameters to care about when coming up with a model are fundamentally given by the ionization coefficients which vary with the electric field and also temperature.

3.1.1 Avalanche multiplication

Figure 13 shows a p-n junction that is reverse biased where it is assumed that the electric field is high enough to create a secondary e-h pair by impact ionization. A primary e-h pair is injected in z (.e.g. photon

absorption) and since the electric field direction goes from N to P; electrons will drift towards z_n and holes towards z_p. At z' a secondary e-h pair is created by collision of the primary pair with a silicon atom.

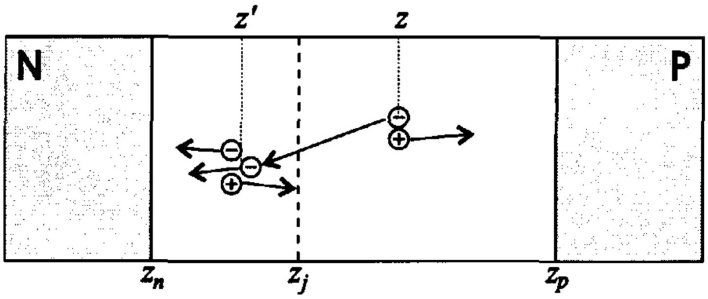


Figure 13: A reverse biased p-n junction in which an electron generated at z generates a secondary e-h pair at z'. z_n and z_p bound the avalanche region. From [62]

Avalanche multiplication happens because as more carriers are ionized, they will in turn collide with silicon atoms and multiply exponentially. Therefore, we introduce the multiplication gain $\mathbf{M(z)}$ representing the number of charge carriers that will flow out the depletion layer for one e-h pair generated at z [62] as well as the ionization coefficients $\alpha_{\mathbf{n}}, \alpha_{\mathbf{p}}$ [cm^{-1}] for electrons and holes respectively, defined as the number of electron-hole pairs generated by a carrier per unit distance traveled [63]. While traversing the depletion layer, both the electron and hole created at z may generate additional e-h pairs which themselves start new chains of ionizations, thus:

$$M(z) = 1 + \int_{z_n}^{z} \alpha_n M(z')dz' + \int_{z}^{z_p} \alpha_p M(z')dz' \tag{7}$$

From (7), we can get further differential equations (see [62]) and derive:

$$\begin{aligned} M_n &= \frac{1}{1 - \int_{z_n}^{z_p} \alpha_n e^{-\int_z^{z_p} (\alpha_n - \alpha_p)dz'} dz}, \\ M_p &= \frac{1}{1 - \int_{z_n}^{z_p} \alpha_p e^{-\int_{z_n}^{z} (\alpha_p - \alpha_n)dz'} dz}. \end{aligned} \tag{8}$$

The multiplication factor for pure electron injection $\mathbf{M_n}$ is the gain for electrons entering the depletion layer at the p-side of the junction and being multiplied across the whole avalanche region. The same goes for $\mathbf{M_p}$ but for holes injected at the n-side. Thus, we can view $M_{n,p} = \frac{I}{In_0, p_0}$ as the ratio of electron/hole current injected at the p/n-side with the total current I flowing across the junction. The multiplication factor $M_{n,p}$ can become infinite because there is positive feedback in the avalanche multiplication. Using (8) this gives the breakdown condition:

$$\begin{aligned} 1 &= \int_{z_n}^{z_p} \alpha_n e^{-\int_z^{z_p} (\alpha_n - \alpha_p)dz'} dz, \\ 1 &= \int_{z_n}^{z_p} \alpha_p e^{-\int_{z_n}^{z} (\alpha_p - \alpha_n)dz'} dz. \end{aligned} \tag{9}$$

Conditions in (9) are called ionization integrals and both are equivalents. In a computer algorithm the breakdown voltage can be determined using one of the ionization integral as will be shown shortly after. They depend on the electric field through the coefficients α_n, α_p. The simplified model above predicts

that the current will become infinite but in practice it is limited by the junction series and space charge resistances [62], [63] and also due to the fact that α_n, α_p depends on the temperature and decreases while the temperature rises with the total current I [64].

3.1.2 Ionization coefficients

So far, the breakdown conditions have been derived in (9) but it is not yet clear how it relates to the breakdown voltage. Looking closer at the ionization coefficients α_n, α_p and how they were defined, we realize that in order for a carrier to ionize a silicon atom it must acquire sufficient kinetic energy to produce an e-h pair by collision meaning that the electric field must be sufficiently high and that the distance between the collisions of the free carrier must be enough to allow acceleration to a sufficiently high velocity. Hence, the ionization coefficients are expected to be dependent on material parameters such as the bandgap, the electric field, the mean free path and temperature among other things. Several theories have been developed to model ionization rates [65]–[68], however in practice Chynoweth's experimental law [69] is often used:

$$\begin{aligned} \alpha_n(E) &= \alpha_{n,\infty} e^{-\frac{b_n}{|E|}}, \\ \alpha_p(E) &= \alpha_{p,\infty} e^{-\frac{b_p}{|E|}}. \end{aligned} \tag{10}$$

were the constants $\alpha_{i,\infty}$ and b_i for $i = n, p$ are determined by experiments. For silicon, data from Van Overstraeten and De Man [70] are frequently selected. Chynoweth's law can be used as a basis for more advanced models in simulations [71] and can be refined to take into account temperature dependence based on [72] for example.

Hence we can see the connection between breakdown voltage with ionization integrals of (9) giving the condition of avalanche breakdown and and the field dependence of the ionization coefficients in (10). The electric field and potential in the depletion layer are determined from the solutions of the Poisson equation, which for semiconductor devices can be generally expressed as [73]:

$$\nabla^2 \Phi = -\vec{\boldsymbol{\nabla}} \cdot \vec{E} = -\frac{q}{\varepsilon_r \varepsilon_0}[p - n + N_D - N_A + N_{tD} - N_{tA}] \tag{11}$$

where q is the elementary charge, p and n the hole and electron concentrations, respectively, and the other terms represent the concentration of additional, typically fixed, charges. These fixed charges can originate from charged impurities of donor (N_{D}) and acceptor (N_{A}) type and from ionized traps (N_{tD}) and (N_{tA}).

In order to find the breakdown voltage V_B, (11) can be solved numerically: when the electric field across the junction is high enough such that the breakdown conditions of (9) are met and the corresponding applied reversed bias can in turn be computed. Thus it can be deduced from the Poisson equation that the breakdown voltage will depend on the doping concentration and doping profile through N_A, N_D and also on the presence of ionized traps. The breakdown voltage can also be affected by the device geometry design as it can redirect the field lines; this can sometimes be a problem as SPADs are subject to premature edge breakdown.

3.1.3 Edge breakdown

For a p-n junction formed by diffusion of impurities into a bulk semiconductor, these impurities will diffuse downward and sideways (lateral diffusion) [74] and the junction will look as on Figure 14. Because of the lateral-diffusion effect, the junction consists of a central plane (or flat) region with approximately cylindrical edges (C region) with a radius of curvature r_j. In addition, if the diffusion mask contains sharp corners, the shape of the junction near the corner will be roughly spherical (S region) because of lateral diffusion.

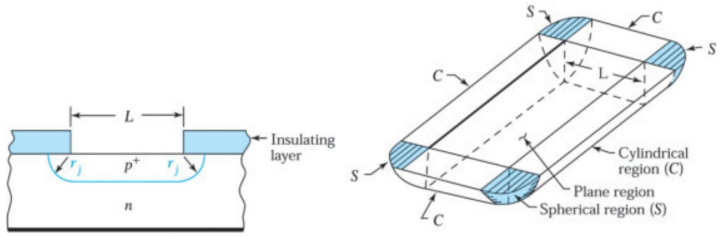


Figure 14: Depiction the doping profile with curved edges generated by diffusion. The electric field strength is more intense in the cylindrical and spherical region which are prone to premature edge breakdown. From [74]

In the C and S regions, the Poisson equation has to be solved in cylindrical and spherical coordinates respectively and it can be demonstrated that the field in these regions is higher than in the planar region [63], [75], this edge effect is enhanced if the radius r_j is small.

Because of the higher field at the edges, the breakdown voltage is smaller in these areas resulting in premature edge breakdown (PEB). With PEB, the SPAD cannot be biased above its breakdown voltage uniformly across the entire photosensitive area.

Hence, the edges of the diode must be protected in some ways by reducing the strength of the electric field. Firstly to overcome PEB and have the diode breakdown from its central photosensitive region, the preferred geometries for SPADs are circles, octagons, or rectangles with rounded corners [76]. Secondly in order to avoid PEB, a wide variety of SPAD guard ring structures exist to protect the SPAD edges, each with certain pros and cons [8]. The guard ring structure is often an isolation ring layer (lower doped region or shallow trench isolation for example) at the periphery of the active region that reduces the electric field. Such a structure is shown on Figure 15. The guard ring separates the central n-well and outer ring n-well forming a low n-density area in the gap for a diffused n-well structure [77]. Larger gaps are more efficient but takes more space and since the guard ring is inactive (a photon striking this region would not be detected) reducing the fill factor and overall photodetection efficiency.

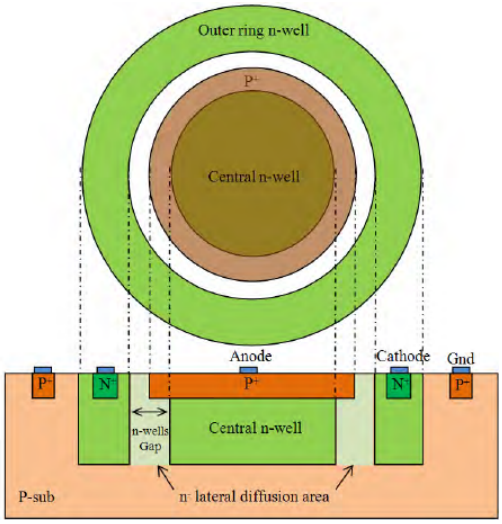


Figure 15: SPAD design structure exhibiting a diffused n-well guard ring. The top view shows the guard ring in white surrounding the active area. From [77]

3.2 Photodetection probability

3.2.1 Photon absorption

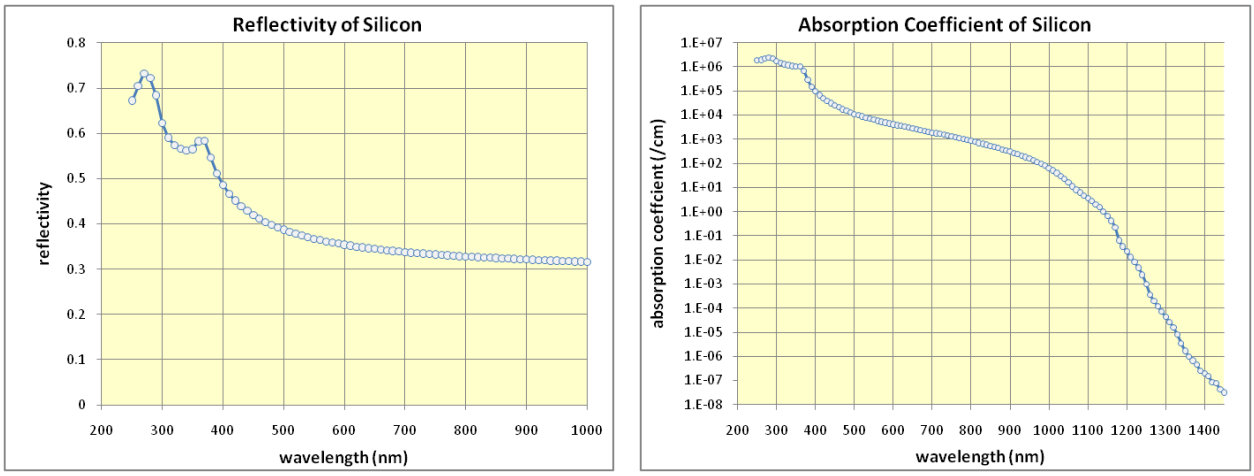
In order to engender an avalanche event, the photon must interact with the semiconductor materials in some way. Light-matter interaction can take place by several means: at the surface of the medium a fraction of the incident light is reflected and the rest transmitted. Inside the medium some of the radiation may be absorbed or scattered while the remainder passes through the sample. Some of the absorbed electromagnetic waves may be dissipated as heat or reemitted at a different frequency (photoluminescence) [78] and higher order processes involving multiple photons may occur [79]. In general, the strongest optical processes are reflection and absorption because they involve the lowest order of interaction between electromagnetic waves and elementary excitations inside the medium and we shall be concerned with these two processes only when discussing photon absorption in SPADs.

With some simplifications [80], light matter interaction in semiconductors can be described by an interaction Hamiltonian characterizing the coupling of the dipole moment $-q\mathbf{r}$ of an electron oscillating with the electric field of the light wave (electric dipole approximation). This interaction Hamiltonian will give the probability transition rate from an initial electronic state i to a final state f through the selection rules. Taking into account energy conservation and the matrix element $\langle f| H_{int} |i\rangle$ of the electric dipole perturbation, we can write $W(\omega)$ .i.e. the transition probability per unit time for photon absorption using Fermi's Golden rule as [78]:

$$\begin{aligned} W(\omega) &= \frac{2\pi}{\hbar} \sum_{\mathbf{k}_c \mathbf{k}_v} |\langle f| H_{int} |i\rangle|^2 \delta(E_f(\mathbf{k}_c) - E_i(\mathbf{k}_v) - \hbar\omega), \\ H_{int} &= -q\mathbf{r} \cdot \mathbf{E_0} cos(\omega t). \end{aligned} \tag{12}$$

where we consider absorption from valence band Bloch functions of wave vector $\mathbf{k_v}$ to conduction band states of wave vector $\mathbf{k_c}$. (12) shows first with the δ function that the energy gap between the initial states in valence band and final states in the conduction band is given by the photon energy $\hbar\omega$ (energy conservation) and therefore, the bandgap gives the minimum photon energy that can be absorbed. Secondly, absorption is more likely to happen for higher matrix elements $\langle f| H_{int} |i\rangle$ which characterize momentum conservation especially.

But momentum conservation with H_{int} in (12) implies that only absorption such that $\mathbf{k_v} = \mathbf{k_c}$ is allowed which is not enough to fully characterize every possibilities that are important to consider, in particular for indirect bandgap semiconductors such as silicon. This can be solved taking into account indirect interband absorption processes: these are phonons assisted-transitions where an electron transit from i to f via an intermediate state [81]. Such processes require to consider an electron-phonon interaction Hamiltonian in addition to the electric dipole in (12). Moreover, Pauli's principle dictates that the $i \rightarrow f$ transition is allowed only if the final states is unoccupied. This can be taken into account with the Fermi-Dirac distribution which gives the statistical occupation of electrons for a given energy. With these two additional considerations and Fermi's Golden rule, it is possible to express the absorption coefficient α including direct and indirect processes. While the theoretical study of the optical absorption coefficient could be beneficial for SPADs devices, it is as this discussion suggest not so straightforward, even more so if we consider the case of doped Si where complicated issues need to be taken into consideration [82]. Nevertheless, datas and tabulations for the optical properties of Silicon exist [83] and typical curves for Si are shown in Figure 16.



(a) (b)

Figure 16: **(a)** Reflectivity of a polished silicon wafer determined from the complex refractive index. **(b)** Absorption coefficient of silicon at T=300 K in cm^{-1} as a function of the wavelength. Sub-bandgap absorption could arise from doping and excitonic states. From [83]

.

From a design perspective, in order to mitigate light reflection we can add an anti-reflection coating layer (ARC) that would improve the optical transmission, but also helps eliminate the penetration of the standing wave into a shallow region close to the silicon surface, where the multiplication region of the SPAD is formed [84]. The absorption coefficient curve on Figure 16b demonstrates that as the wavelength gets shorter, the photon is more likely going to be absorbed shallower in Si. For comparison, if we take $\lambda = 400nm$ its absorption depth is about 0.1μm while for $\lambda = 900nm$ it is close to 30μm. This is important for design purposes: the targeted wavelength range is the main parameter to be considered when chosing the appropriate SPAD thickness and junction depth. Upon close inspection of the absorption coefficient data in [83], a sudden drop for a photon energy $\hbar\omega = 1.06eV$ is noticeable. This value corresponds to the bandgap in Si. Absorption below this threshold becomes negligible but still may occur notably due to dopants and impurities and excitonic states [78]. In order to quantify the probability that a photon with a given wavelength λ will be absorbed at a given depth x inside the material, we can use Beer-Lambert's law [12] and take the reflectivity into account:

$$P_{abs}(x) = [1 - R]\alpha e^{-\alpha x} \tag{13}$$

3.2.2 Avalanche triggering probability

In order to be detected a photon must be absorbed and subsequently must fire the avalanche process. Not every photogenerated carrier will be successful in generating a Geiger avalanche. It was assumed in the previous section that every injected e-h pair would induce an avalanche breakdown. In fact there is a nonzero probability that the primary carrier will lose energy by lattice scattering and recombination such that the sequence of ionizations stops before the entire junction goes into Geiger breakdown [85]. In addition, photons can be absorbed in the neutral regions above and below the depletion layer (see Figure 17). In this case a carrier that is generated in the neutral regions moves only by diffusion and, thus, has a limited probability to reach the depletion layer.

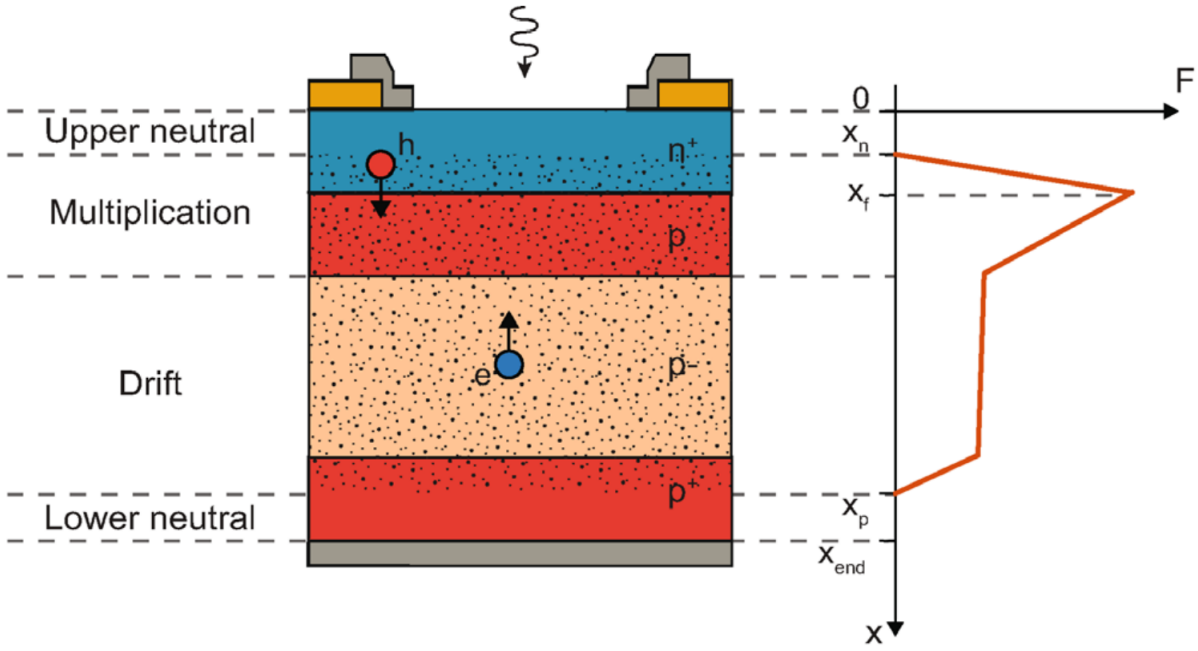


Figure 17: Basic structure of a SPAD, with corresponding electric field F on the right side. A photon absorbed in the depletion region $x_n < x < x_p$ has some probability that it succeeds in generating one or more daughter electron-hole pairs by impact ionization while a photon absorbed in the neutral regions must first diffuse to the drift and multiplication zone before it recombines. From [12]

.

The probability P_{diff} of the hole or electron diffusing to the depletion region before recombination will depend on the recombination process (mainly SRH and Auger) [23] and can be characterized by the diffusion length for holes, electrons $L_{e,h}$.. P_{diff} could be first approximated by [85]:

$$\begin{aligned} P_{diff} &= e^{-(x_n - x)/L_h} \; if \; 0 < x < x_n, \\ P_{diff} &= e^{-(x - x_p)/L_e} \; if \; x_p < x < x_{end}, \\ P_{diff} &= 1 \; if \; x_n < x < x_p. \end{aligned} \tag{14}$$

(14) is related to the probability of a hole in the upper neutral layer or an electron in the lower neutral layer created at x reaches the depletion region before it recombines and can be therefore considered equal to one when the carrier is located within the depletion region. The diffusion length characterize the typical distance travelled by a carrier before recombination and hence is linked to the recombination rate and diffusivity of the carrier. In reality, P_{diff} is more generally modeled by Fick's law and depending on the recombination process considered, several factors affect $L_{e,h}$ like temperature, local carrier concentration, concentrations of impurities and mobility.

As stated above, even if the carrier reach the depletion layer or is directly generated in the avalanche region, the avalanche breakdown process is not guaranteed. We define $P_e(x)$ as the probability that an electron starting at position x in the depletion layer will have an infinite number of descendents, .i.e., will trigger an avalanche. The function $P_h(x)$ for holes is analogously defined. Consequently, the probability $P_T(x)$ that either the electron or the hole initiates an avalanche is given by [86]:

$$P_T = P_e + P_h - P_e P_h \tag{15}$$

In order to obtain $P_e(x)$ and $P_h(x)$, the following system is to be solved numerically with the appropriate boundary conditions [85]–[87]:

$$\begin{aligned} \frac{dP_e}{dx} &= (1-P_e)\alpha_e(P_e + P_h - P_eP_h), \\ \frac{dP_h}{dx} &= -(1-P_h)\alpha_h(P_e + P_h - P_eP_h). \end{aligned} \tag{16}$$

where $\alpha_{e,h}$ are the ionization coefficients of (10). The presence of the ionizations coefficients in (16) indicates that $P_e(x)$ and $P_h(x)$ will be maximized for high $\alpha_{e,h}$ and thus the chance of detecting an avalanche event increases with the electric field strength and overvoltage.

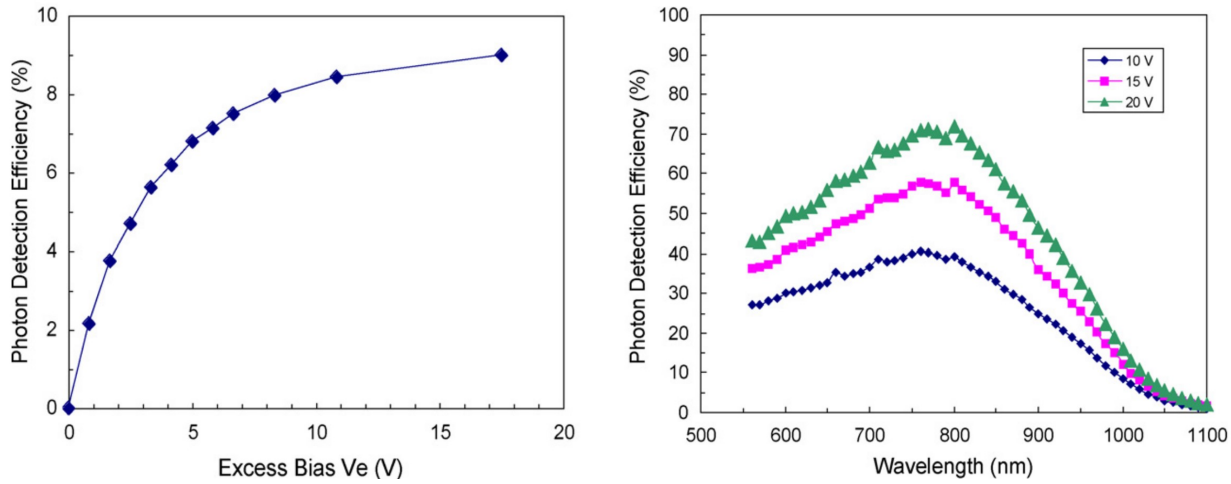
3.2.3 Photodetection probability

The photodetection probability is defined as the probability that a photon impinging on the active area of the detector succeeds in triggering an avalanche current that can be detected by the electronics [12]. It is therefore a combination of the probabilities relating to the absorption, diffusion and avalanche process and can be expressed as a product of these three terms, of which the diffusion term P_{diff} depends on where the photon is absorbed [12], [85]:

$$\begin{aligned} PDP &= \int_0^{x_{end}} P_{abs}(x)P_{diff}(x)P_T(x)dx \\ &= \int_0^{x_n} P_{abs}(x)P_{diff}(x)dx \times P_h(x_n) + \int_{x_n}^{x_p} P_{abs}(x)P_T(x)dx + \int_{x_p}^{x_{end}} P_{abs}(x)P_{diff}(x)dx \times P_e(x_p) \end{aligned} \tag{17}$$

where P_{abs}, P_{diff}, P_T come from (13),(14), (15), (16). The second equality subdivide (17) into three distinct absorption region. The first term when light is absorbed in the $0 < x < x_n$ domain characterizes the event where a photon is absorbed in the upper neutral region in Figure 17 creating a hole, the latter diffuse to the depletion zone and induce an avalanche breakdown starting at x_p. A similar interpretation can be given for the third term but for an electron instead of hole. For the middle term, the photon is absorbed in the depletion zone creating an e-h pair and the avalanche intitiation probability is given by P_T.

Typical PDE curves (obtained by multiplying the PDP by the fill factor) are displayed in Figure 18. Figure 18a shows that the PDE increases with the excess voltage bias above breakdown which comes from the increment of ionizations coefficients in P_e, P_h as discussed previously. As expected the PDE is zero below the breakdown voltage. Figure 18b demonstrates that SPADs are ususally calibrated to detect a specific wavelength range. As shown previously, longer wavelengths may be absorbed too deep in the device and the carrier has more chance to recombine before reaching the depletion region. This could also be the case for very short wavelengths absorbed near the surface depending on how deep is the P-N junction located inside the device. In any case, we can target the desired wavelength and tune the PDE by chosing the depletion region depth and width.



(a) (b)

Figure 18: **(a)** Photondetection efficiency at different excess bias for a thin SPAD at $\lambda = 850$ nm. **(b)** Photon-detection efficiency with respect to the wavelength for a thick SPAD for three different excess bias. From [52]

.

3.3 Dark count rate and afterpulse

A dark count is an avalanche event caused by non photogenerated carriers, which can be originated from four main causes: diffusion from neutral regions, thermal generation, band-to-band tunnelling (BTB) or by release from a charge trap (afterpulse) [6]. Thermal generation can be further subdivided into three processes: direct band-to-band thermal generation, trap assisted thermal generation and trap assisted tunneling generation. These phenomenons are represented on Figure 19

As shown on the same figure and for the diffusion process, the minority carriers are thermally excited in the bulk region of the p-n junction and could diffuse to depletion layer. However, because of high recombination rate for minority carriers in the p+ and n-well bulk regions, very few non-equilibrium minority carriers could be generated and drift to the depletion region. Therefore, the contribution of carriers direct thermal generation to DCR is negligible in these two bulk regions and DCR is mainly governed by thermally and BTB tunneling generated non-equilibrium carriers in the depletion layer [88].

Direct BTB thermal generation process is similar, except that the generated carrier do not need to diffuse. But due to the indirect bandgap of the Si material, indirect thermal generation in the depletion layer is overwhelmingly dominant over the direct BTB thermal generation, thus the latter process can be also ignored.

For trap assisted thermal generation, this corresponds to the recombination-generation mechanisms involving traps that can be modelled according to the well established SRH theory [23]

.

In the case of trap assisted tunnelling, not only the conventional SRH mechanism is involved but also tunneling via traps. For this process and as it is suggested on Figure 19, the emission of electrons and holes from a trap is assisted by a phonon instead of thermal emission over the entire trap depth.

Finally, the BTB tunneling mechanism describes transitions of electrons which tunnel directly from the valence band to the conduction band, this is a pure case of quantum tunnelling through a potential barrier [89].

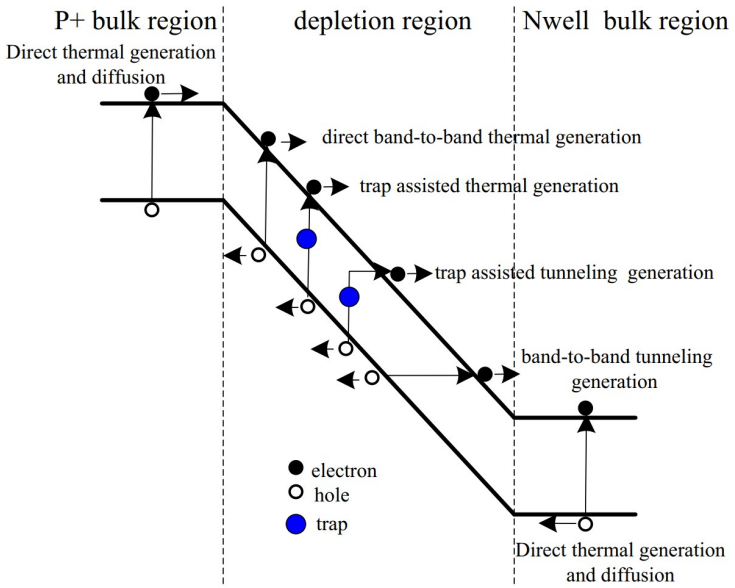


Figure 19: Schematic of several sources of dark counts in a typical p+/n-well SPAD device. The DCR phenomenons shown are diffusion with thermal generation in the p+ and n-well neutral regions, the tree thermal generation processes in the depletion layer and BTB tunnelling also in the depletion zone. From [88]

As discussed in the previous paragraph, direct thermal generation processes are negligible and we shall focus on the trap assisted, BTB tunnelling as well as afterpulse mechanisms later on.

3.3.1 Trap assisted carrier generation

We will start by first stating the main results of indirect generation-recombination phenomenons involving traps and SRH theory. A trap (or deep center) usually originates from a crystal flaw. Such defects can be caused by broken bonds, strain associated with displacement of atoms, and difference in electronegativity or core potentials between the impurity and host atoms [78]. These are not of the same types as dopant impurities which are conveniently called shallow donors/acceptors.

We consider for example an impurity (e.g. Au atom) that is introduced in the Si substrate which provide a deep level within the bandgap, meaning that the energy level E_t introduced by the impurity is far away from the conduction and valence band at the same time and lies close to the middle of the bandgap as can be seen on Figure 20. The generation-recombination processes shown in the figure consist of electron capture by the empty center in **1**, electron emission from the occupied center **2**, hole capture by an occupied center **3**, and hole emission by an empty center **4**. The **1** and **3** processes are recombination mechanisms since in both cases a free carrier is withdrawn from the conduction or valence band. On the other hand, the processes **2** and **4** are generation (or emission) mechanisms since in both cases a free carrier is added to the conduction or valence band.

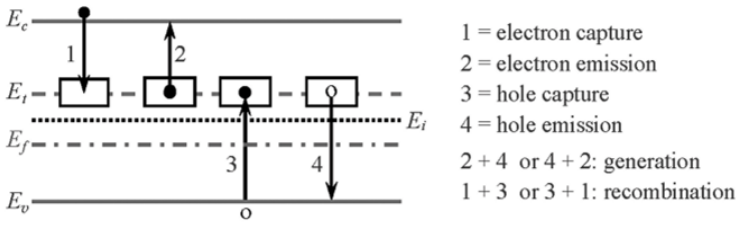


Figure 20: Generation and recombination in an indirect bandgap semiconductor; E_t is the trap level deep into the bandgap; **1**, **2**,**3** and **4** represent the generation and recombination processes. From [90]

The emission or recombination rate $R_{1,2,3,4}$ of the processes will depend on several factors. Firstly, the number of occupied and unoccupied states will be governed by Fermi-Dirac's statistics with the corresponding distribution (at equilibrium) and also depends on the trap concentration N_t [cm^{-3}]. Secondly and using the same logic, emission/recombination is more likely to happen with high free carrier concentration n, p [cm^{-3}]. Thirdly, the trapping potential is highly localized and interact with carriers only if they come in its vicinity and therefore the chance of a carrier interacting with the trap will also depends on the its capture cross sections for holes and electrons $\sigma_{p,n}$ [cm^2] and the carrier thermal velocity v_{th} [cm/s].

It can then be shown [23], that the net recombination rate $U_{SRH} = R_1 - R_2 = R_3 - R_4$ at steady state is:

$$U_{SRH} = \frac{np - n_i^2}{\tau_p[n + n_i e^{\frac{E_t - E_i}{k_b T}}] + \tau_n[p + n_i e^{\frac{E_i - E_t}{k_b T}}]} \tag{18}$$

where n_i is the intrinsic carrier concentration, E_i is the midgap (or intrinsic Fermi level) and $\tau_p = [\sigma_p v_{th} N_t]^{-1}$, $\tau_n = [\sigma_n v_{th} N_t]^{-1}$ are the hole and electron minority carrier lifetimes for n- and p-type Si. We can deduce from (18) that U is maximized when $E_t = E_i$ and thus that trap levels near the midgap are the most efficient recombination centers. When modelling dark counts, the SRH net recombination rate is the fundamental formula describing the DCR for the trap assisted thermal generation process.

The trap assisted tunnelling process also involves a deep level but in this case (18) will need to be modified to take tunnelling through a potential barrier into account. The process is represented in more detail on Figure 21. Instead of thermal emission over the entire trap depth $E_c - E_t$, which is the only escape mechanism possible in the absence of a field, carriers can also be emitted by thermal excitation over only a part of the trap depth (transition $P \rightarrow P'$ on Figure 21 similar to the SRH mechanism as seen before) followed by tunneling assisted by a phonon through the remaining potential barrier (transition $P' \rightarrow P$") [91]. On the same figure is also shown the Poole-Frenkel effect, which is the mechanism whereby in the case of Coulomb interaction between the free carrier and the trap, the effective trap depth is lowered (dashed line).

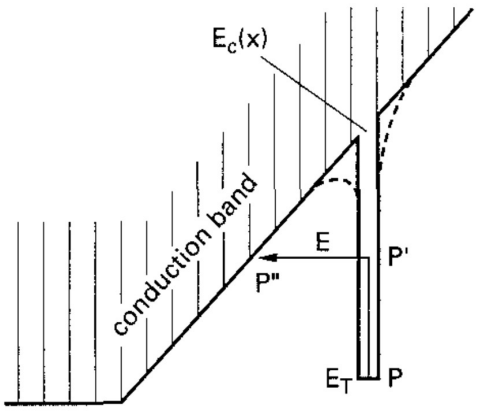


Figure 21: Generation and recombination in an indirect bandgap semiconductor; E_t is the trap level deep into the bandgap; **1**, **2**,**3** and **4** represent the generation and recombination processes. From [91]

In order to take the tunnelling effect into account, (18) is modified by introducing so-called "field-effect enhancement factors" [88] Γ_n and Γ_p for electrons and holes which accounts for the effects of tunnelling on both the density of captured carriers by a trap and the emission rate of carriers from a trap. It can then be demonstrated after some simplifications that [91]:

$$\begin{aligned} U_{SRH,TAT} &= \frac{np - n_i^2}{\frac{\tau_p}{1+\Gamma_p}[n + n_i e^{\frac{E_t - E_i}{k_b T}}] + \frac{\tau_n}{1+\Gamma_n}[p + n_i e^{\frac{E_i - E_t}{k_b T}}]}, \\ \Gamma_{n,p}(x) &= 2\sqrt{3\pi}\frac{|E(x)|}{F_\Gamma} e^{(\frac{E(x)}{F_\Gamma})^2}, \\ F_\Gamma &= \frac{\sqrt{24 m^*(k_b T)^3}}{q\hbar}. \end{aligned} \tag{19}$$

where $E(x)$ is the local electric field strength, m^* is the effective mass of the tunneling electrons for silicon. The approximation giving $\Gamma_{n,p}$ in (19) is valid if the applied electric field strength do not exceed 9×10^5V/cm and we also have $\Gamma_n = \Gamma_p = \Gamma$. The model of (19) neglect the Poole-Frenkel effect but this is hypothesis is not too harsh since this effect is much weaker than tunneling at a strong electric field [91].

The rate $U_{SRH,TAT}$ not only increase with temperature but also with the local electric field. This can be explained by the fact that as the applied field (or applied potential) gets larger, the band bending effect of the p-n junction is enhanced which in turn reduces the potential barrier width (distance of the $P' \rightarrow P$" transition in Figure 21). Moreover, we can notice that the SRH model of (18) is simply a limiting case of (19) when $E(x) = 0$ when the bands are completely flat and therefore $U_{SRH,TAT}$ includes both trap assisted thermal generation and trap assisted tunnelling processes at the same time.

Even if a carrier is generated inside the depletion region, it is still not guaranteed that avalanche breakdown will follow as discussed in subsection 3.2 and has a finite probability to occur. Hence the DCR must include this probability ((15) and (16)) and the DCR caused by traps related thermal and assisted tunneling generations can be obtained by integrating the relevant quantities over the whole avalanche multiplication region [88]:

$$DCR_{SRH} + DCR_{TAT} = \int_{x_n}^{x_p} P_{trig}(x) U_{SRH,TAT}(x) dx \tag{20}$$

3.3.2 Band to band tunnelling generation

Figure 22 shows the valence and conduction band of a reverse biased p-n junction. Under the applied electric field, a barrier between the valence and conduction band is created. Assuming a linear potential across the junction, the width of this barrier is E_g/qF [62] where E_g is the bandgap and F the local electric field. Thus, as the electric field strength increases, the width of the barrier gets thinner and BTB tunnelling is more likely to occur. If an electron starts this tunnelling process at x_1, then after the tunnelling process there will be an electron at $x_2 = x_1 + E_g/qF$ and a hole at x_1. This can be represented by the generation of an e-h pair in the middle of the gap $x = \frac{x_1+x_2}{2}$. The e-h pairs generation rate is dependent on the position of the carrier as an electron located outside the $x_n \leq x \leq x_p$ region cannot tunnel because there are no final (or initial) states into (or from) which the electron can do so. The same reasoning applies with the carrier energy and tunnelling is possible only if $E_{fn} < E < E_{fp}$.

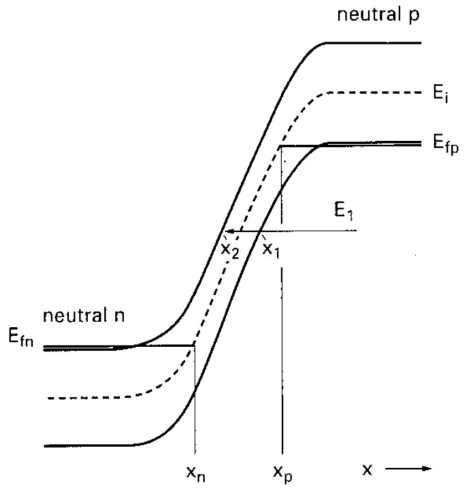


Figure 22: Schematic band diagram of a reverse biased p-n junction. The BTB tunneling mechanism is indicated by an electron of energy E_1 tunnelling from x_1 to x_2. The process is only possible in the region $x_n \leq x \leq x_p$. From [91]

According to [91], the generation rate due to BTB tunnelling is given by:

$$U_{BTB} = B|F(x)|^{\sigma} D e^{-\frac{F_0}{F(x)}} \tag{21}$$

where D is a statistical factor and B, σ and F_0 are definable parameters. In the standard model of [91], the values at room temperature of these parameters for silicon are $B = 4 \times 10^{14}\text{cm}^{-0.5}\text{V}^{-5/2}\text{s}^{-1}$, $F_0 = 1.9 \times 10^7$ Vcm^{-1} and $\sigma = 5/2$.

Similar to trap assisted generation, the DCR of SPADs induced by BTBT tunneling can also be given by:

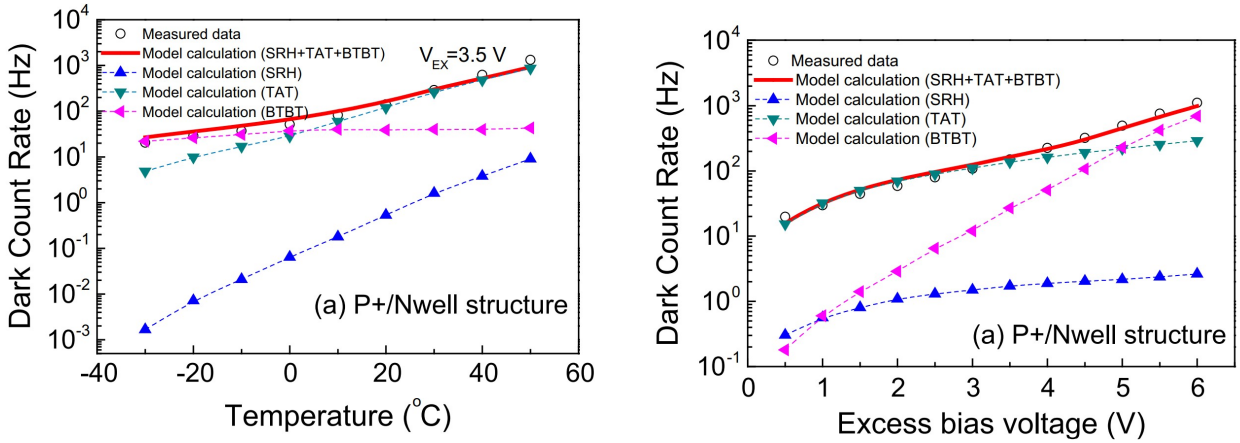
$$DCR_{BTB} = \int_{x_n}^{x_p} P_{trig}(x) U_{BTB}(x) dx \tag{22}$$

And thus we can sum up all the contributions to get the total DCR:

$$DCR_{TOT} = DCR_{TAT} + DCR_{SRH} + DCR_{BTB} \tag{23}$$

3.3.3 Temperature and Voltage bias dependence of DCR

In order to characterize the DCR behavior according to temperature and excess bias voltage. Figure 23 illustrates the DCR contributions of each phenomenons with respect to temperature and excess bias voltage for a p+/n-well SPAD structure from [88].



(a) (b)

Figure 23: **(a)** DCR dependence of temperature for a p+/n-well SPAD structure in standard 0.15 μm CMOS process. **(b)** DCR as a function of excess bias voltage at room temperature for a p+/n-well SPAD structure in standard 0.15 μm CMOS process. From [88]

The temperature dependence of Figure 23a shows in this case first a negligible SRH thermal contribution to DCR. More important it shows an exponential rise of the SRH and TAT contribution with temperature which is to be expected considering (19) accounting for the activation of traps with temperature. As for the BTB contribution, it remains constant at all temperature since this parameter does not change the tunnelling barrier significantly but is still slighlty dependent due to the variation of Si bandgap with temperature [92]. The effect of excess bias voltage illustrated on Figure 23b shows that the BTB process increases exponentially with the applied bias while TAT and SRH processes increase too but less significantly. Hence, SRH generation and TAT will dominate at high temperature and small excess bias and as the excess bias increases, BTB tunneling becomes the main DCR generation source, even more so at lower temperature. Overall, DCR is minimized when reducing temperature and excess bias voltage as suggested by the red plot on Figure 23.

In practice however and unlike in simulations it is not possible to directly separate the different DCR contributions. One way to quantify the influence of TAT and BTB tunnelling is to measure the DCR versus temperature at a fixed excess bias and use the Arrhenius equation [35]:

$$DCR = Ae^{-\frac{E_A}{k_b T}} \tag{24}$$

where A is a proportionality constant that depends on the trap density and E_A is the activation energy of the thermal process which represent the barrier height. The DCR logarithm is then plotted against the reciprocal of the temperature (Arrhenius plot) (see Figure 24).

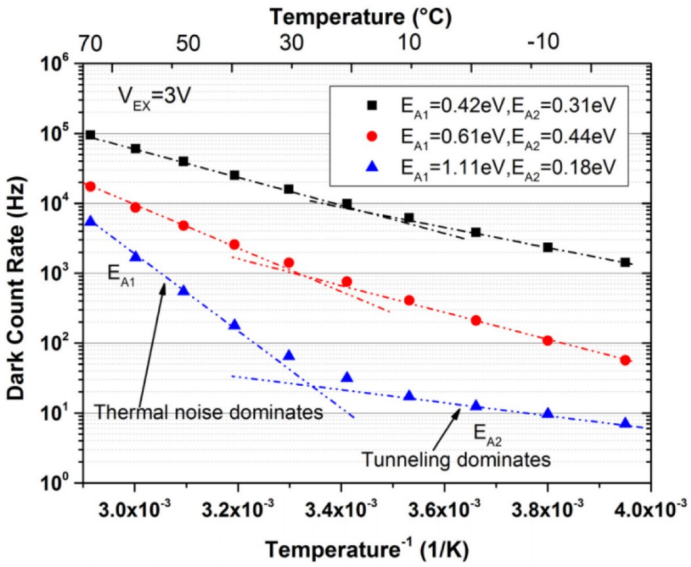


Figure 24: DCR versus temperature for 3 selected SPAD samples fabricated in standard CMOS 150nm technology at 3V excess bias; the dashed lines indicate the Arrhenius equation fit. From [35]

The Arrhenius plot of Figure 24 indicates that the DCR can be fitted by two straight lines of which the corresponding slopes represent the activation energies. Each of these activation energies characterize the dominating DCR process for a given range of temperature: the tunnelling contribution dominates at lower energy and is characterized by a smaller slope E_{A2} since it barley depends on the temperature while thermal processes are more active at higher temperature as discussed previously and are characterized by the steepest slope E_{A1}.

By looking at the E_A values one can get important information about the DCR dominating process at a given temperature [35], [92]. The direct thermal generation components (usually negligible) features the highest dependence on temperature, with an activation energy corresponding to the silicon bandgap $E_G = 1.21$ eV (E_{A1} for the blue plot in Figure 24), the pure SRH trap assisted thermal generation exhibits an intermediate activation energy $E_A = E_G/2 = 0.56$ eV and depending of the electric field we have TAT generation with an activation energy reduced to less than E_G (red and black plot). Finally, if the activation energy E_{A2} is much smaller than E_{A1} (blue plot) it is a strong indication of BTB tunnelling dominating the DCR. E_A also depends on the applied bias voltage because, as discussed, when the overvoltage gets higher, the amount of tunnelling to and from the traps is increased, which enhances deviation from pure SRH statistics [92].

Finally, it is to note that the doping level influence the DCR by means of the electric field profile across the junction. This is essentially because a high doping concentration results in a thin depletion region enhancing the electric field and in this case the BTB contribution dominates [76], [88].

3.3.4 Afterpulsing

Afterpulsing is a particular type of correlated dark count. During the avalanche process, a few carriers flowing over the depletion region are captured by trapping centers. After the avalanche pulse is fast quenched, the SPAD is recharged to excess bias voltage state. At this moment, if the captured carriers are detrapping, they will also initiate an avalanche event, which is called afterpulsing [30]. The larger the number of carriers trapped during a certain avalanche pulse, the shorter will be the time between the end of the avalanche pulse and the emission of the first trapped carrier, therefore the afterpulsing rate increases with the trap density

and total number of carriers crossing the junction (total charge of the avalanche pulse) [6].

Not very avalanche breakdown will display afterpulsing as it is also a probabilistic phenomenon. Thus the first step for modelling afterpulsing is to express a formula for the probability that a carrier is trapped during the avalanche and subsequently cause itself a junction breakdown when released. The model developed according to [62], [93] is discussed for electron traps but is entirely analogous for hole traps.

First, it is assumed that one type of electron trap i is present in the depletion layer. The probability that the electron is captured by a trap is given by the product of trap density with capture cross section $N_{t,i}\sigma_{n,i}$. After the capture process the electron is released and the probability of generating an avalanche breakdown is given by P_e in (15) (for holes, P_h is used). Then the probability that one electron generated at position x_g, is trapped and causes an afterpulse later equals:

$$P_{ale}(x_g) = \int_{x_n}^{x_g} N_{t,i}\sigma_{n,i}P_e(x)dx \tag{25}$$

where it is recalled that x_n is the position of the depletion layer boundary in the n-type side of the junction (see Figure 17). The electron can be generated randomly at any position x_g in the depletion layer. Consequently the probability that one random electron generated at a random position during breakdown causing subsequently an afterpulse can be calculated by averaging (25) over the spatial distribution of triggering probability P_{trig} in (15):

$$P_{ae} = \int_{x_n}^{x_p} P_{trig}(x_g)P_{ale}(x_g)dx_g \tag{26}$$

Using (26) the effective width of the high field effective region for electron trap afterpusling W_e (analoguous for holes W_h) can be defined and the afterpulsing probability of one electron can be rewritten:

$$\begin{aligned} W_e &= \int_{x_n}^{x_p} P_{trig}(x_g)(\int_{x_n}^{x_g} P_e(x)dx)dx_g, \\ P_{ae} &= N_{t,i}\sigma_{n,i}W_e \end{aligned} \tag{27}$$

For holes a similar expression can be derived. One can already notice that W_e depends solely on the avalanche probabilities for hole and electrons of (15) which themselves increases with the excess bias. Therefore, afterpulsing events will be more likely to take place with for higher overvoltages.

So far, only the probability that one electron generated during a breakdown event will cause an afterpulse has been calculated. The second step is to consider that many electrons are generated when Geiger breakdown takes palce and we want to calculate that probability that if n_{el} electrons are generated at least one of them will generate an afterpulse. The probability of k electrons being trapped and cause afterpulse when n_{el} electrons are generated in the depletion layer follows a binomial distribution $B(X = k) = \binom{n_{el}}{k}P_{ae}^k(1 - P_{ae})^{n_{el}-k}$. Hence if n_{el} electrons are generated during the avalanche, the total afterpulsing probability noted A_i for a given type of trap is equal to $1 - B(x = 0)$:

$$A_i = 1 - (1 - P_{ae})^n \approx n_{el}N_{t,i}\sigma_{n,i}W_e \tag{28}$$

The number of e-h pairs n generated during the avalanche can be estimated considering that in the SPAD equivalent circuit (subsection 3.4) the total capacitance C across the SPAD is discharged and the voltage across the diode drops from the breakdown voltage plus the excess bias V_e to the breakdown voltage:

$$n_{el} = \frac{CV_e}{q} \tag{29}$$

which shows for the second time that afterpulsing is enhanced with the overvoltage.

To fully quantify afterpulsing characteristics, we also need to analyze its time dependence. The emissions of first carrier events follow Poisson statistics [94] and the probability that a filled trap i will emit the captured charge carrier in a time interval dt is given by dt/τ_i, in which τ_i is the trap lifetime (time between capture of an electron and its release) [62]. Thus the traps display a release time following an exponential distribution with a time constant of τ_i and we get the afterpulsing probability by summing the contribution of the different traps:

$$\begin{aligned} P_{at}(t)dt &= \sum_i \frac{A_i e^{-\frac{t}{\tau_i}}}{\tau_i}, \\ \int_{t_1}^{t_2} P_{at}(t)dt &= \sum_i A_i [e^{-\frac{t_1}{\tau_i}} - e^{-\frac{t_2}{\tau_i}}]. \end{aligned} \tag{30}$$

where (t_1, t_2) is the time interval after which a first avalanche has been triggered at t_0 ($t_1, t_2 > t_0$) [95] and A_i is the total afterpulsing probability of (28).

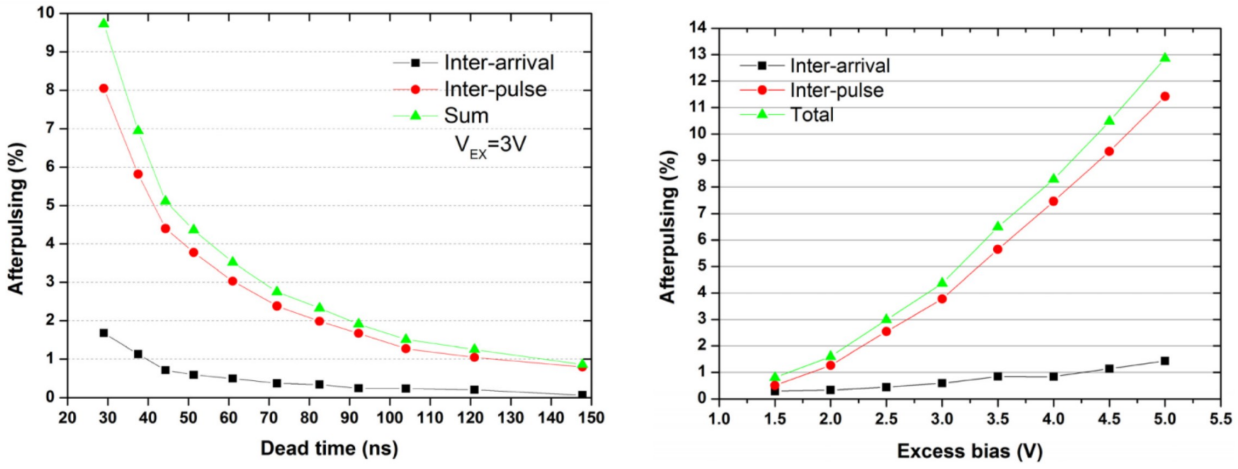
As for the DCR ,the afterpulsing statistics and τ_i is also a function of temperature and follows an Arrhenius type law [94]:

$$\tau_i = Ce^{-\frac{E_{Ai}}{k_b T}} \tag{31}$$

where E_{Ai} is defined as the trap activation energy. Therefore, reducing temperature will in turn maximize the trap lifetime in the device and afterpulsing will be delayed.

A common method that is used to characterize afterpulsing and minimize it as well is to apply to the SPAD cell a dead time after breakdown (time during which the SPAD is inactive. During the dead time a filled trap could be released but without having the opportunity to induce a Geiger breakdown, hence diminishing afterpulses counts and afterpulsing probability. According to (29), it was also stated that afterpulsing increases with overvoltage due to the larger current density, which increases the number of carriers crossing the device area and thus their probability to be captured.

To illustrate, the total afterpulsing probability with respect to the dead time is shown in Figure 25a (green plot). The probability decays exponentially since the beginning of time interval t_1 of (30) starts later. The time scale is in the nanosecond order: although the SPAD device may have different types of traps having different release time behavior from the subnanosecond to above microsecond range, for afterpulsing we are usually interested in the characterization of those dominating the nanosecond range, where the detector performance suffers the most and to ensure that they're not mixed with dark counts [95]. On Figure 25b is shown the afterpulsing probability as a function of the excess bias measured with a 50 ns dead time, which depicts a relatively linear dependence.



(a) (b)

Figure 25: **(a)** Afterpulsing as a function of dead time at 3 V overvoltage. **(b)** Afterpulsing as a function of excess bias with 50 ns dead time. From [35]

3.4 Quenching and equivalent electrical model

The avalanche breakdown inside the diode once started can be self-sustaining and remains active for a significant amount of time before coming to an end [3]. This is detrimental for the SPAD operation because it cannot withstand current this high for too long without important damage, furthermore it will be difficult to distinguish the counts from one another since their corresponding signal risk to overlap. Therefore, the SPAD must be adequately quenched after each avalanche which is done by reducing the voltage through the diode below or near the breakdown voltage. This quenching circuit and resulting pulse output for a SPAD cell is presented in the following section.

3.4.1 Equivalent electrical model for a single SPAD

The SPAD can be quenched by simply integrating a high quenching resistance (several of 100 kΩs) included with a parasitic capacitance. In comparison the internal diode resistance range from 500 Ω for SPAD types with a wide area and thick depletion layer and from a few hundred Ωs to various kΩs for devices with a small area and a thin junction [30]. The avalanche current will quench itself by developing a voltage drop on the high impedance resistor load (quenching resistor). These simple type of circuit is reffered to as passive quenching circuit [30] (PCQ) and is illustrated on Figure 26.

The intrinsic part of the diode (dashed rectangle on the figure) is described by the internal resistance of the space-charge region R_d and the inner depletion layer capacitance C_d while the quenching part is connected in series and characterized by its resistance R_L. A stray capacitance C_s and low value resistor R_s are also shown when the SPAD is connected to an external circuit for the purpose of pulse capture for example. Between the anode and cathode is applied a bias voltage V_A with the positive voltage on the cathode.

The switch models the avalanche event: when it occurs the switch closes and the total capacitance $C = C_d + C_s$ is discharged through the smaller resistance R_d. The voltage drop in the internal node is very close to the excess bias $V_E = V_A - V_{BD}$ and when the voltage across the junction reach the breakdown point V_{BD} where the avalanche can no longer self-sustain the switch opens up and the capacitors start to recharge.

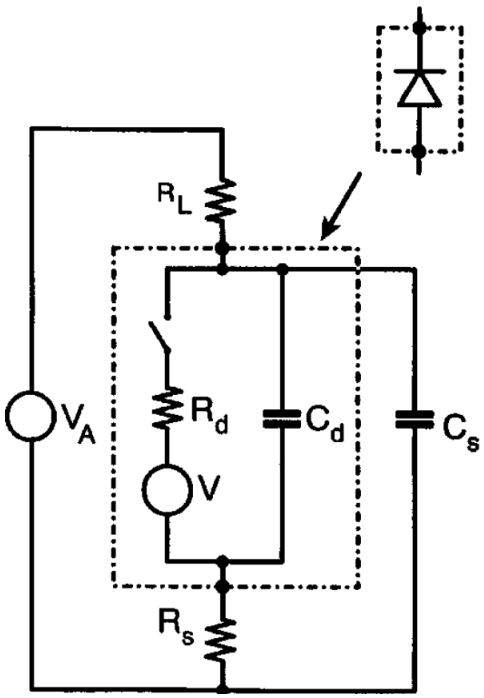


Figure 26: Traditional equivalent electrical model of a single SPAD micro-cell with R_d, C_d the internal diode resistance and capacitance, R_L the load (or quenching) resistance, R_s, C_s the low value resistance and stray capacitance and V_A the applied bias. The switch modelling the avalanche on/off state is located in series with R_s. From [30]

When the switch closes, the total capacitance is discharged through R_d and barely any current passes through R_L which can then be ignored in practice during the discharge phase. The current through the diode rises abruptly with an amplitude given by the overvoltage and diode resistance and then decreases exponentially with a time constant T_q corresponding to R_d and C [30]. We can then write:

$$\begin{aligned} V_{ex}(t) &= V_E e^{-\frac{t}{T_q}}, \\ I_d(t) &= \frac{V_{ex}(t)}{R_d}. \end{aligned} \tag{32}$$

where I_d is the diode current and V_{ex} the transient excess voltage. The quenching time constant is set by $T_q \approx CR_d$. The total resistance is in theory R_d in parallel with R_L but since the latter several order of magnitude bigger, it can be neglected. When the declining excess voltage approaches 0, the intensity of $I_d(t)$ becomes low and the number of carriers that traverses the avalanche region is then small. The avalanche process is statistical in nature and when the diode current I_d falls below a latching current level $I_q \approx 100\mu$A, the SPAD cell is fully quenched and the switch reopens meaning that I_d declines abruptly to 0 [30]. Because of the avalanche statistics, the I_q value (and therefore the switch opening moment) is not sharply defined as it is evidenced by the microplasma fluctuations of the diode turn-on/turn-off process [96]. The discharge process is represented on Figure 27 showing the device theoretical current and voltage pulse output. When avalanche breakdown takes place (the switch closes), the current rises almost instantaneously and then drops exponentially with the voltage. The figure also depicts the asymptotic values I_f, V_f of which the current and voltage tends if the switch was never to be open again. In that case I_d and the total voltage across the diode V_d tend to:

$$\begin{aligned} \lim_{t\to\infty} I_d(t) &= I_f = \frac{V_E}{R_d + R_q} \approx \frac{V_E}{R_L}, \\ \lim_{t\to\infty} V_d(t) &= V_f = V_{BD} + R_d I_f. \end{aligned} \tag{33}$$

However, when I_q is reached, the switch opens again, the current suddenly drops to 0 and voltage rises as the capacitors begins to charge back. The model also allows to estimate the total charge Q_{BD} contained in the avalanche breakdown pulse, a important parameter for afterpulsing. Assuming an asymptotic behavior $I_d(t) \to I_f$ and integrating over time (or simply considering the total charge stored in the capacitor C) we get $Q_{BD} \approx V_E C$ as in (29).

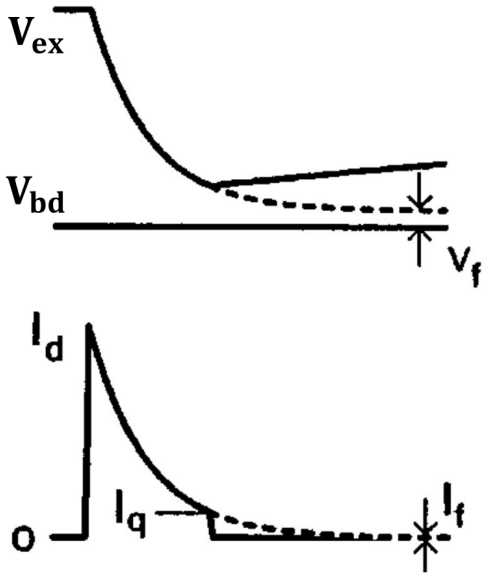


Figure 27: Voltage and current output waveform across the SPAD cell. The current rises and voltage drops the moment the switch in the equivalent circuit closes, then I_d decays exponentially towards I_f while V_d decreases towards V_f close to the breakdown voltage. Finally when I_q is reached the switch opens up. From [30]

The recharge phase corresponds to opening the switch in the diode equivalent circuit after discharge. During this phase, the capacitances are slowly recharged this time by the small current through the quenching resistor R_L. Almost no current flow through the diode and we can write for the transient excess voltage:

$$V_{ex}(t) = V_E(1 - e^{-\frac{t}{T_r}}) \tag{34}$$

where $T_r \approx CR_L$ corresponds to the SPAD characteristic recovery time constant, so that it takes $\sim 5T_r$ to recover the correct excess voltage within 1%. Due to the high quenching resistance, T_r is typically in the microsecond range while by comparison T_q on (32) is in the nanosecond order depending on R_d. The full output waveform for subsequent triggering is shown in Figure 28. Compared to the time scale of T_q, the discharge phase happens very briefly, which gives narrow pulses for the diode avalanche current (in (**a**)) and a sudden voltage drop in (**b**).

The recharge phase characterized by T_r is on the other hand much slower as attested by the diode voltage exponentially returning to its original value. A photon or dark count could retrigger the SPAD during this recovery phase however, meaning that the current pulse height will be shorter since it depends on the excess voltage at that moment (see (34)). It is also to note that if the diode fires at a voltage lower than V_A, it operates with lower photon detection efficiency and impaired photon-timing resolution [30]. This can generate issues with the device operation as it may never have the time to recharge enough leading to count rate saturation and paralyzable dead time behavior [97]. A solution to this problem is to consider using active quenching circuits which can accomodate this issue better.

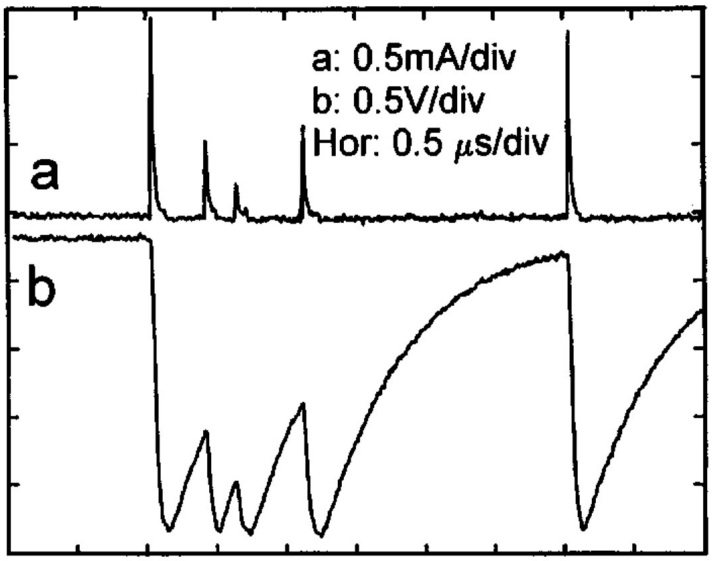


Figure 28: Pulse waveforms of a SPAD of that operates with a PQC. **a** represents the avalanche current I_d and **b** the diode voltage V_d. The pulse height of the current pulses will be shorter if the SPAD is retriggered during the recovery phase $t \lesssim 5T_r$. From [30]

To speed up the recharge phase, one could think of choosing a small quenching resistance. Unfortunately, it is not possible to go too low if we want a proper quenching behavior. This is because the diode current tends to the asymptotic value I_f during the discharge phase which as can be deduced with (35) increases when R_L gets smaller. But for the avalanche to stop being self-sustaining the current needs to fall below the latching current I_q and consequently if $I_f > I_q$ quenching takes place with a progressively longer delay making the dead time even longer [30]. To avoid this and as a rule of thumb the designer should ensure that the R_q value is at least 50 kΩ per applied excess bias V_E.

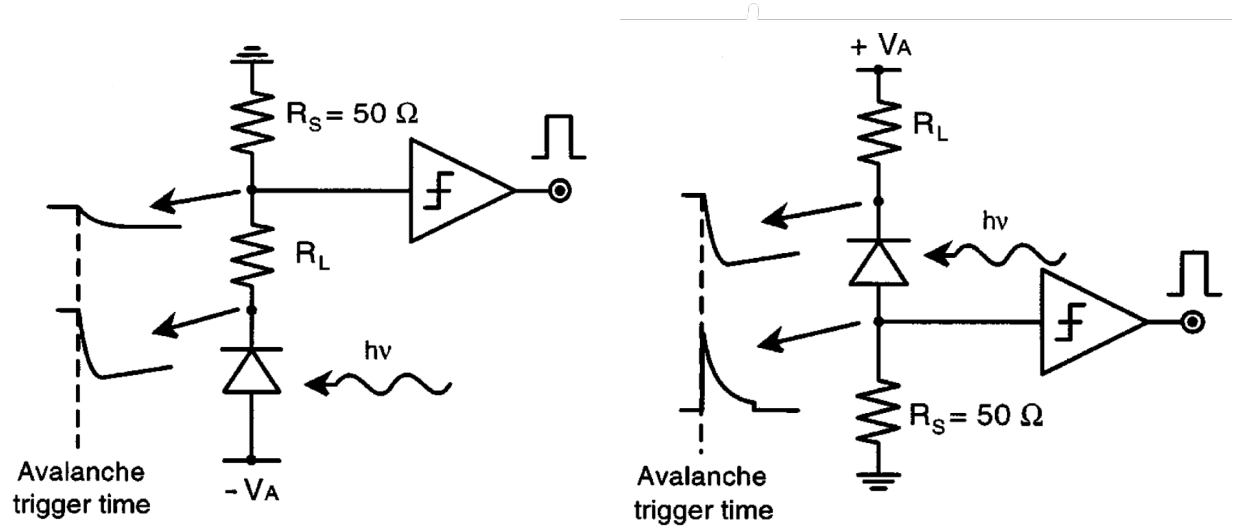
3.4.2 Output pulse

In order to capture the voltage output signal emitted by the SPAD, three main ways could be employed:

- **Voltage output mode**: This consists of picking up the voltage signal at the cathode, right between the SPAD and quenching resistor (shown in both Figure 29a and Figure 29b). By doing so, the voltage output waveform with falling edge of Figure 28**b** is obtained. However, for a SPAD with integrated quenching resistor it is generally not possible to probe the signal at this location.
- **Voltage quenched ouptut mode**: If the device have its own integrated quenching resistor, it is still possible to get the voltage output waveform. It can be done by inserting a low value resistor R_s in

series with the quenching (or load) resistor R_L on the ground lead of the circuit (see Figure 29a). The device is negatively biased at the anode and the signal is then picked between R_s and R_L and fed into an amplifier. The main drawback is that the two resistance act as a voltage divider and the output waveform is a attenuated replica of Figure 28**(b)**, so the appropriate value for the R_s resistance has to be chosen carefully.

- **Current output mode**: This mode is similar to the previous one but the R_s resistor is placed on the ground lead of the photodiode (Figure 29b) and the bias source is inverted in comparison with the voltage quenched output mode. If the voltage output is probed between R_s and the SPAD anode, the waveform of the pulse is directly that of the diode current and is therefore called current output mode. In order to capture a significant voltage pulse on R_s, the stray capacitance C_s of Figure 26 must be comparable to or greater than the intrinsic diode capacitance C_d, otherwise, only a small fraction of the avalanche current will flow through R_s while most part of the current flows within an internal loop within the diode.



(a) (b)

Figure 29: **(a)** Voltage quenched output mode. The SPAD is negatively biased at the anode and the signal is picked up between the high quenching (load) resistor R_L and low value resistor R_s. **(b)** Current output mode. The SPAD is positively biased at the cathode and the signal is picked up between the SPAD anode and low value resistor R_s so that the ouptut waveform ressembles to the current pulse. From [30]

The equation for the voltage output mode is the one given in (34), considering the discussion for the two other modes we get:

$$\begin{aligned} V_{volt} &= V_E \frac{R_s}{R_L}, \\ V_{current} &= V_E \frac{R_L}{R_d(1+\frac{C_d}{C_s})}. \end{aligned} \tag{35}$$

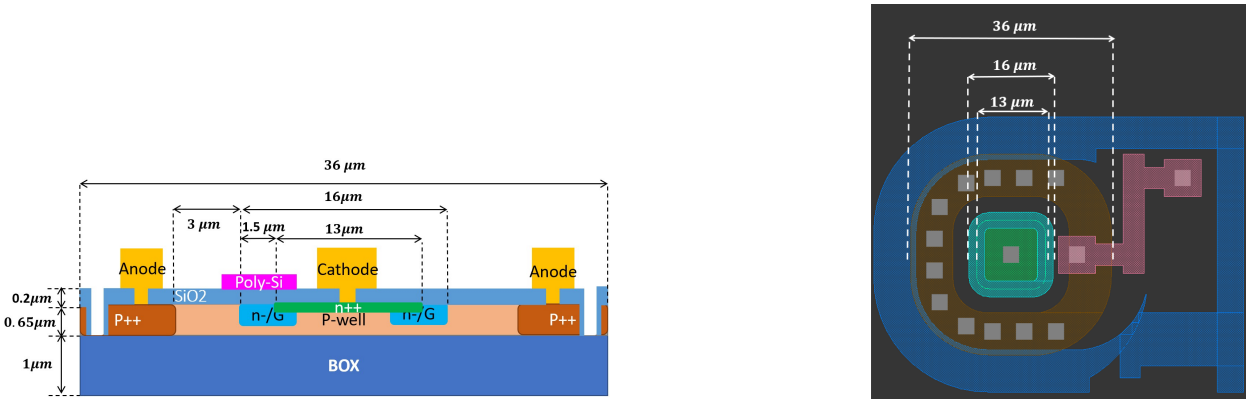
where V_{volt} corresponds to the signal amplitude for the voltage quenched output mode, and $V_{current}$ to the signal amplitude of the current output mode.

4 The BSI UV enhanced SOI SPAD

Recently the development of a new SPAD design had started at UCLouvain [98], [99]. This novel device is designed to target and optimize the photodection efficiency in the UV and sub-UV domain, this wavelength range is of great interests for several applications (fluorescence lifetime imaging microscopy, LIDAR, proton emission tomography) but requires a very shallow junction to absorb such short wavelengths. This can be quite challenging, it can be deduced from Figure 16b that the typical absorption depth in Silicon is about 100 nm for a wavelength λ=400 nm and go down to 5 nm if the wavelength is λ= 300 nm. Moreover, for PET applications it would be interesting to integrate these individuals SPADs into SIPMs arrays. In that case SOI technology comes in handy to make optimized SPAD micro-cells and circuitry separately (subsubsection 2.3.2) but then the device would need to be backside illuminated (BSI) and as such the bulk Si region below the junction has to be thin enough to minimize the diffusion length upon creation of an e-h pair. This means that for absorbing short wavelengths and eventhough the SiO_2 BOX layer is nearly transparent, it is required for the SPAD thickness to be reduced and extra care is mandatory when handling such a thin membrane to avoid compressive or tensile stress issues.

4.1 Design and structure

The essential design features are shown on Figure 30. A 1 μm BOX layer makes the SPAD backside under an ultrathin 0.65 Si μm Si film which is essential for good photodetectivity in the UV spectral range. The diode itself is made of a shallow n++/p-well junction and the doping profile is optimized to enhance the photodectivity in the targeted range. The guard ring consists of n-well implant that overlaps the junction, that way it is part of the photoactive area when the SPAD is backside-illuminated making an active area of 16 $\mu m \times 16 \mu m$. The p++ region (anode) is 7 μm wide and surrounds the central region. In the same area, a trench is realized to separate neighbouring micro-cells and avoiding crosstalk for the purposes of making SiPM arrays. It is worth noting that at this point, array structures have been fabricated as well, for a given SiPM the total area is $1000 \times 980 \mu m^2$ for 22 $\times$ 16 micro-cells so considering the SPAD photoactive area this gives a fill factor of 9.2% . Finally, a thin Polysilicon layer of 300 kΩ is integrated acting as the quench resistor.

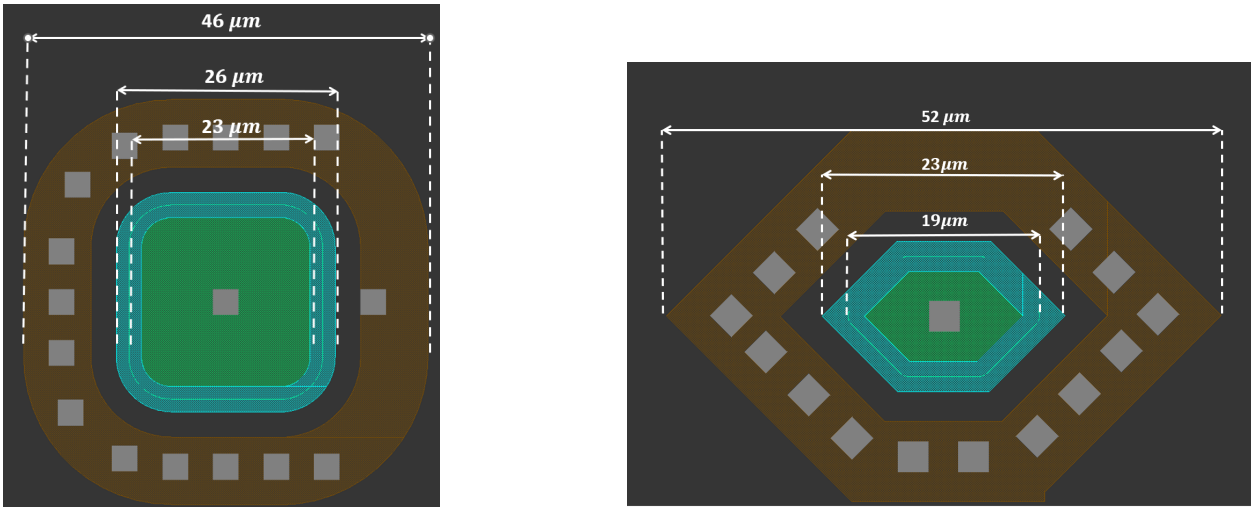


(a) (b)

Figure 30: **(a)** Vertical cross-section of the UV enhanced BSI SPAD cell. **(b)** Top view of the cell showing the square with rounded edges green n++ implant, blue guard ring and brown p++ region as well as the trench (blue), pink quench resistor and anode/cathode contacts in grey.

Two other design variations have been fabricated as well (Figure 31). For the first one, the shape remains identical but the active region is 10 μm wider while for the second design an hexagonal area is fabricated.

Both designs can be interesting for improving the fill factor. SiPM arrays were also fabricated for these modifications and the fill factor can therefore be calculated in the same manner as for the first model.



(a) (b)

Figure 31: **(a)** Top view of second SPAD design (shown without trench and quench resistor) similar to the first but with larger dimensions. **(b)** Top view of the third SPAD design with an hexagonal shape.

The dimensions for the three designs are shown in Table 2. An improvement of the fill factor is noticeable for the second and third design however it could be optimized, for example by reducing the distance between the guard-ring and anode.

Design	**Shape**	**Anode width** [μ**m**]	**Cathode width** [μ**m**]	**G-R width** [μ**m**]	**G-R overlap** [μ**m**]	**Gap** [μ**m**]	**Active area** [μ**m**^{2}]	**Fill factor**
36	Rounded square	36	13	16	1.5	3	256	9.2 %
46	Rounded square	46	23	26	1.5	3	676	16.5%
52	Hexagonal	52	19	23	2.5	4	374	12.5 %

Table 2: Dimension parameters showing the variation between each type of SPAD fabricated. The designs are labelled according to the anode width.

4.2 Device simulation

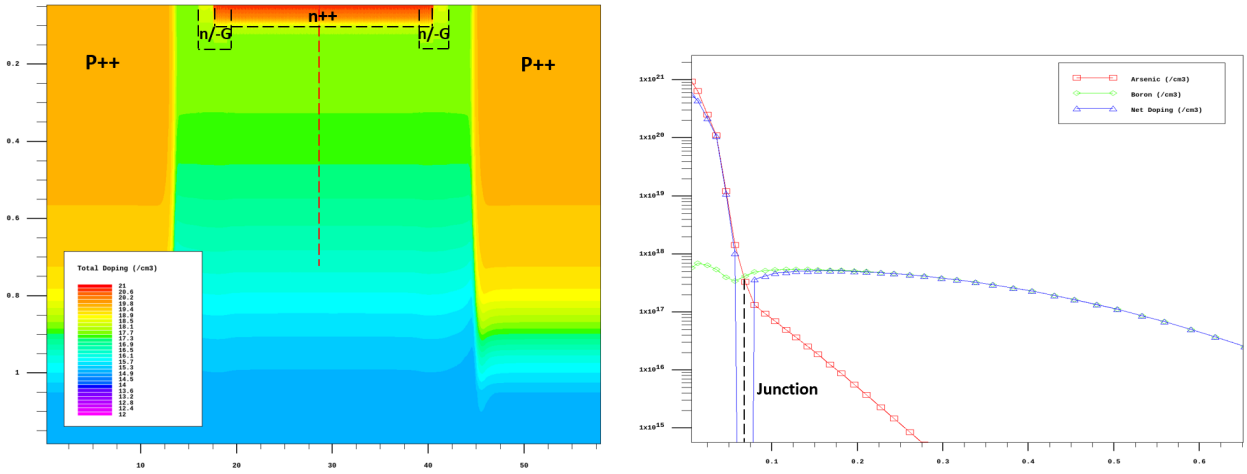
Simulations for the 36 SPAD cell have been carried out with Silvaco (Athena and Atlas [71], [100]) which is a TCAD assisted tool used for simulating electronic devices and their fabrication process. Firstly, it models the flow of the device fabrication steps and generates detailed geometric information and doping profile distribution of the device. Then, it uses the information of the first step to predict the device characteristics. The TCAD simulation involves the following steps:

- Device generation using SILVACO's process simulator (Athena) to virtually fabricate the device and simulate the doping implantation and diffusion steps
- Device simulation that solves the mathematical equations with specific physical models to describe the device characteristics and behavior
- Visualisation of all the simulation results by generating plots and diagrams

4.2.1 Optimized doping profile with Athena

Extra care is needed for the junction simulation as it needs to be optimized. The parameters for doping implantation and diffusion will influence the junction depth, its width and the electric field profile across the junction and consequently the breakdown voltage. For instance, a higher boron dose for the p-well implantation will increase the maximum electric field inside the depletion region which by itself results in an improved PDP but the reduction the depletion width will be narrower which implies based on (17) a reduction of the triggering probability contribution. Moreover dark counts are more likely to occur by tunnelling effect (see subsection 3.3).

Figure 32 shows the resulting doping profile across the middle of the junction, obtained after simulations of several steps of of thermal SiO_2 growth, ion implantation, and annealing in SILVACO Athena as done in [98]. The p-well gradient doping profile allows to get a strong electric field at the junction close to the surface and allows to sweep out the UV photocarriers more easily to the avalanche region. This results in a junction located at 115 nm from the surface and at 535 nm if the photon is absorbed at the backside (assuming a transparent BOX).



(a) (b)

Figure 32: **(a)** Cross sectional view of the total doping concentration simulated with Athena. The dashed red line across the middle of the junction shows the location of dopile profile cutline of Figure 32b **(b)** Doping profile showing the boron, arsenic and net doping concentration across the dashed cutline.

4.2.2 Device simulation with Atlas

Silvaco Atlas allows the user to chose from several models in order to simulate the device physics with great details. With all the possibilities, some compromise must be made between accuracy of the underlying physics and computational ressources with choices that are in adequacy with the device under study.

Simulation under dark conditions: Before performing the simulation of the SPAD under illumination (which is more time consuming), it is still possible to get important parameters of the device in dark conditions, such as the electric field profile, I-V curve with breakdown voltage and the avalanche probabilities P_T, P_e, P_h from (15). The main models used for the simulation (performed at room temperature) are as follows:

- Transport model: The classical drift-diffusion model is implemented by default for the transport equations. Thus the overall system to be solved is given by Poisson's equation, current continuity equations

and drift-diffusion equations established by Van Roosbroeck [101]. Of course, for avalanche multiplication a suitable impact ionization model is to be accounted for.

- Recombination: For the continuity equations, a good recombination model is mandatory in order to simulate the out of equilibrium phase. SRH theory is well established and widely used for recombination when deep levels but also shallow donors/acceptors that are located within the bandgap. The governing equation has been discussed in subsection 3.3. The model used here takes into account the impurity (doping) concentration and also Auger three-carrier recombination which is important at high current densities.

- Carrier statistics: The concentrations of carriers in the device also depends on the carrier statistics. Electrons in thermal equilibrium obey Fermi-Dirac statistics but in silicon the Boltzmann approximation can be used at room temperature which is the choice for the simulation. In the presence of heavy doping, the bandgap becomes doping dependent and decreases with increasing concentration. Thus considering the high concentration at the anode and cathode, a bandgap narrowing model is also used.

- Mobility: The lost of carrier momentum through various scattering processes is considerated. Atlas provides tables for the low field carrier mobility with respect to the doping concentration which is appropriate for the bulk part of the device but is valid at room temperature only. For the high-electric field region, a velocity saturation model is also accounted for.

- Impact ionization: Atlas allows to chose from models based on physical principles, but Chynoweth's law of (10) is preferable and the basis for the one employed during the simulation. In the Selberherr's impact ionization model, the datas for the ionizations coefficients are taken from measurements [70] and can be adapted with the temperature according to the semi-empirical method in [72], which relies on evaluating the carrier mean free path and the energy loss per phonon scattering, both temperature dependent.

For impact ionization, so called non-local models do exist [71], the main advantage is that contrary to the local model used the ionization coefficients are not overestimated and it allows a more accurate simulation of the breakdown voltage. However, it relies on an Energy Balance transport model, which is more delicate to implement and cost more resources than drift-diffusion transport. With this in mind, Table 3 sum up the models used during the simulation under dark conditions at room temperature.

Transport	**Recombination**	**Carrier statistics**	**Mobility**	**Impact ionization**
default (Drift-diffusion)	`consrh` (SRH)	`boltzman` (Boltzmann approx)	`conmob` (Low field mobility)	`selb` (Selberherr)
	`auger` (Auger)	`bgn` (Bandgap narrowing)	`fldmob` (Saturation velocity)	

Table 3: Summary of the main models used for simulating the device under dark conditions at room temperature

In order to find the breakdown voltage with the models given above, the voltage is gradually ramped at the cathode, a compliance limit is set at 1 μA to stop the simulation once the breakdown point is reached. The resulting I-V curve is shown below:

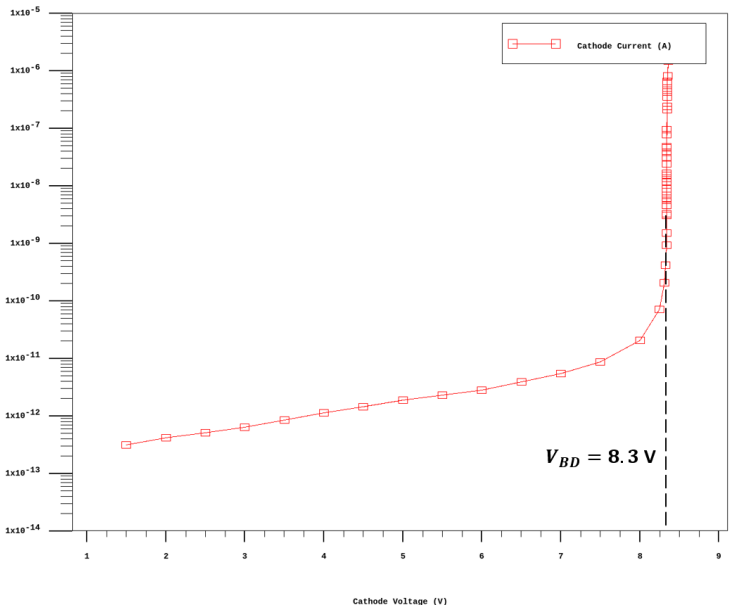
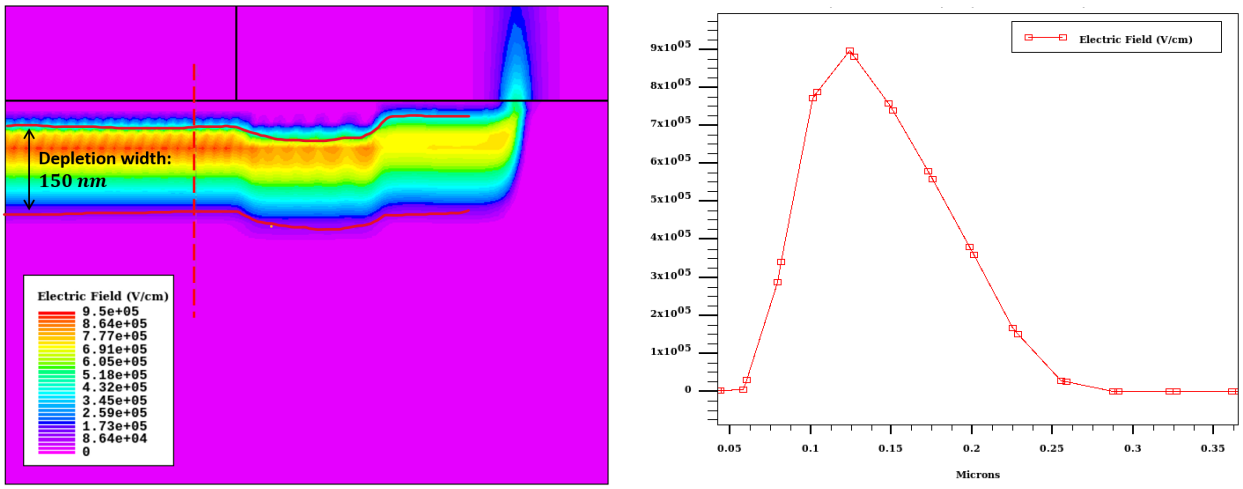


Figure 33: I-V curve of the SPAD under dark conditions

Below the breakdown point, the simulated dark current is rather low in the picoampere order and the compliance is reached at 8.3 V which we defined as the breakdown voltage.

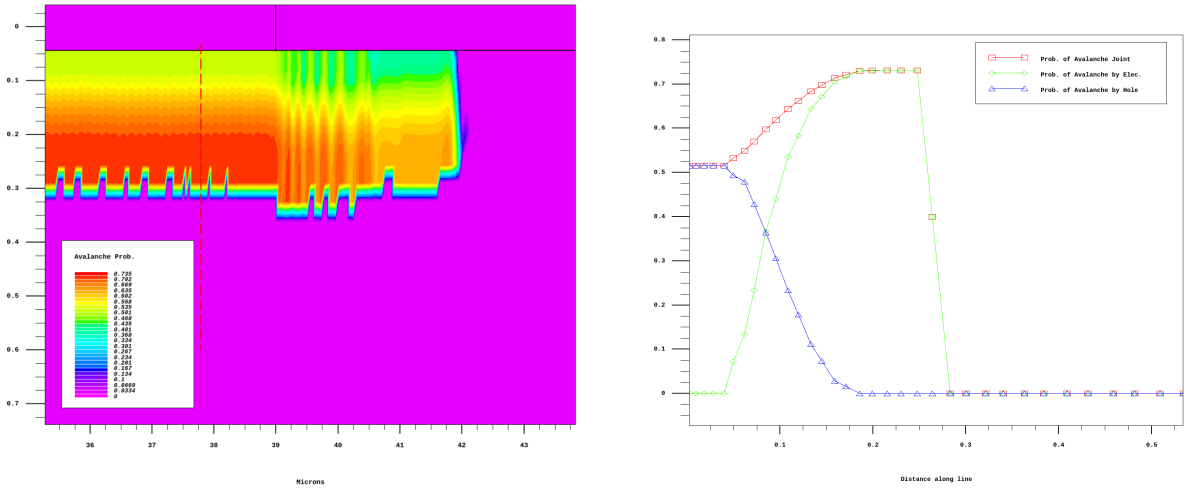
The resulting electric field profile at the breakdown point is shown on Figure 34, thanks to the guard-ring the field strength is attenuated at the edges avoiding premature edge breakdown. The depletion region is estimated based on the approximate locations where the electron and hole population decline abruptly. This results in a space charge region approximately 150 nm wide with a peak electric field of 9.5×10^5 V/cm at the junction, which ensures a good avalanche triggering probability thanks to the doping profile, although a wider depletion width could be made by reducing the implantation dose when forming the p-well. The peak electric field would be slightly reduced but still strong enough to induce a proper avalanche breakdown and additionally BTB tunnelling would decrease.



(a) (b)

Figure 34: **(a)** Zoom on the simulated electric field profile at the right side of the junction showing the width and depth of the depletion region at the breakdown point **(b)**. Electric field profile taken across the dashed cutline of Figure 34a

Finally, the avalanche triggering probabilities are computed at 3 V of excess bias by taking a cutline across the junction and shown in Figure 35b. As expected, the trigerring region matches the electric field strength profile of Figure 34 and peaks at 73% at the bottom of the depletion region where electrons start the breakdown process. The avalanche probability of holes , which is maximum at 55% at the top of the depletion region, is smaller than the one of electrons because of different ionization coefficients. Indeed, the $\alpha_{i,\infty}$ and b_i constants of (10) determined experimentally leads to a smaller ionization coefficient for holes given the same electric field strength.



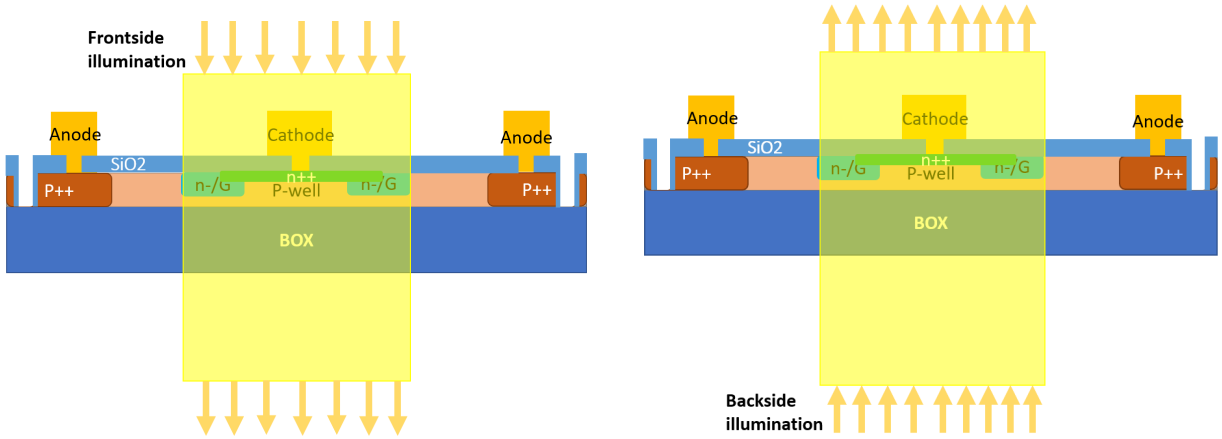
(a) (b)

Figure 35: **(a)** Cross sectional view of the device showing the distribution of avalanche triggering probability. **(b)** Avalanche triggering probabilities for holes, electrons and joint across the dashed cutline of Figure 35a.

Simulation under illumination Atlas also supports an optoelectronic simulator, essentially the parameters for the light rays are specified (wavelength, intensity, beam width, direction etc) and based on the material properties of the device (mainly the optical refractive indexes) the numbers of photogenerated carriers are computed.

The propagation of light can be described by several physical models. Ray tracing is used during the simulations and follows the laws of geometrical optics accounting for reflection and refraction but also absorption based on the refractive index provided by the software database.

To reproduce the device behavior under illumination, light source of intensity 1×10^{-4} W/cm^2 is placed above (frontside illumination) or below (backside illumination) the SPAD photoactive area (see Figure 36). In order to get the quantum efficiency and photodetection probability, the selected wavelengths range from 250 nm up to 600 nm and their respective refractive indexes are chosen from a data file given by Atlas (`Si110.nk`) that matches closely the one found in [83] (of which reflectivity and absorption coefficient are shown on Figure 16). For the indexes of other material, the default data file is taken and Table 3 remains the same for the main models.



(a) (b)

Figure 36: **(a)** Picture representing the device phtotoactive area illuminated from above during the optoelectronic simulation. **(b)** Picture representing the device phtotoactive area illuminated from below during the optoelectronic simulation.

The quantum efficiency is first computed. Atlas does not directly calculate quantum efficiency but it does calculate two useful quantities that can be used to get the quantum efficiency: the cathode current and source photocurrent. By definition the quantum efficiency is the ratio between the number of carrier detected divided by the number of incident photons on the detector which is equivalent to the ratio of cathode current difference $I_{C,\lambda} - I_{C,dark}$ by source photo current I_S:

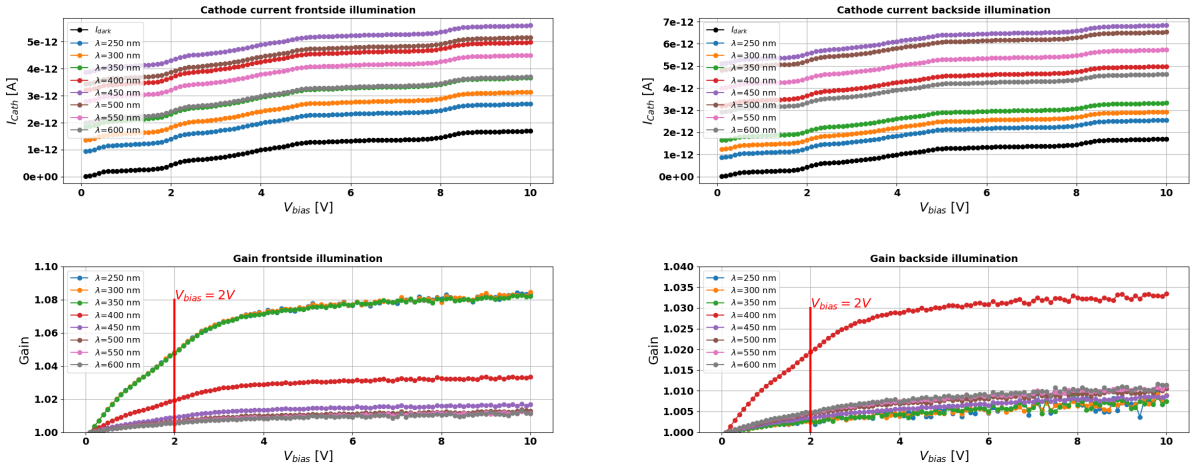
$$\begin{aligned} I_S &= q\frac{\lambda}{hc}P_\lambda, \\ QE &= \frac{I_{C,\lambda} - I_{C,dark}}{I_s}. \end{aligned} \tag{36}$$

where for the cathode current difference, $I_{C,\lambda}$ indicates the cathode current under illumination and $I_{C,dark}$ the cathode current under dark conditions, the difference between the two is taken so to get the current purely generated by the light source. For the source photocurrent, q is the elementary charge, h the Planck

constant, c the speed of light and P_λ defines the optical power. In order to compute this value as accurately as possible, any positive feedback mechanisms such as avalanche gain must be avoided. Consequently, it was chosen to turn off the impact ionization model during the simulation to not overestimate the cathode current so that the gain $G = 1$. The latter can be defined in this case for any bias voltage [102]:

$$G(V) = \frac{I_{C,\lambda}(V) - I_{C,dark}(V)}{I_{C,\lambda}(0.1) - I_{C,dark}(0.1)}. \tag{37}$$

Since the impact ionization phenomenon is deactivated, no intrinsic gain mechanism is present and QE can be computed at pretty much any voltages since $I_{C,\lambda} - I_{C,dark}$ is approximately constant. Figure 37 shows the I-V curves and gains of the different wavelengths under frontside and backside illumination; as expected the gain G stays relatively close to 1. Still, when computing the quantum effciency as in (36) the cathode current is taken at $V_{bias} = 2$ V to stay as close as possible from $G = 1$.



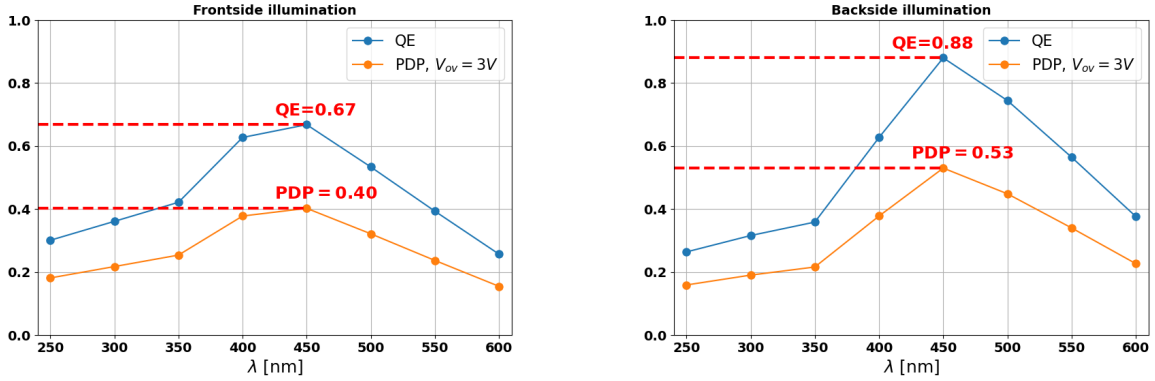
(a) (b)

Figure 37: **(a)** I-V curves and gains of the cathode current under frontside illumination for different wavelengths and impact ionization model turned off. **(b)** I-V curves and gains of the cathode current under backside illumination for different wavelengths and impact ionization model turned off.

According to (17), in order to get the photodetection probability we want ideally to compute the probability P_{diff} related to the diffusion part. This however depends on the doping profile and is not easy to simulate but it is implicitly computed when getting the QE since the software takes recombination outside the junction into account. Therefore, to get an estimation of the PDP the quantum efficiency is multiplied by the triggering probability P_T averaged over the depletion width which is essentially the formula given in (1), doing this we get an average of 0.6 for the joint triggering probability of holes and electrons. The QE and PDP computed are shown on Figure 38 for the BSI and frontside illumination case.

The peak Photodetection probability is 40% and 53% for the frontside-illuminated and backside-illuminated device respectively. For the sub UV range (250-350 nm) the quantum efficiency is slightly better when the SPAD is frontside illuminated which is to be expected since the junction is closer from the photogenerated carriers. The difference is still negligible suggesting that even when the photon is absorbed at the backside the losses by recombination when carriers diffuse to the depletion region is minor which is consistent with the thin Si body and long lifetime for electrons [103] given the doping concentration ($\sim 5 \times 10^{17}$ cm^{-3}) in the p-well region . Nonetheless, it is to bear in mind that the software consider the thick BOX to have an absorption coefficient of 0, which is questionable for a real device. In that case and based on absorption only,

it is expected to have better PDP than shown in Figure 38, however for this wavelength range the losses come mainly from the high reflectivity (see Figure 16a) and it is predicted that photogeneration could be drastically enhanced with the use of an antireflection coating layer. The drop in photodetection probability after 450 nm can be explained by the higher absorption depth > 1 μm and transmission coefficient meaning that a significant part of the light passes through the device without being absorbed in the Si body. From 450 nm however, the quantum efficiency of the BSI simulation is surprisingly high compared to frontside illumination. It is hypothesized that reflection at the metallic cathode is non negligible in that case. Indeed, as shown on Figure 36 under frontside illumination part of the incident ray would be reflected at the cathode preventing it from reaching the depletion region. Based on the same figure but for backside-illumination on the other hand, rays with longer wavelengths reflected at the metallic boundary extend their path into the thin Si body. If so, this effect is expected to be less important in a real device because in the 2D simulation all cross sections are identical and it is as if the cathode covers in the z-direction the entire photoactive area.



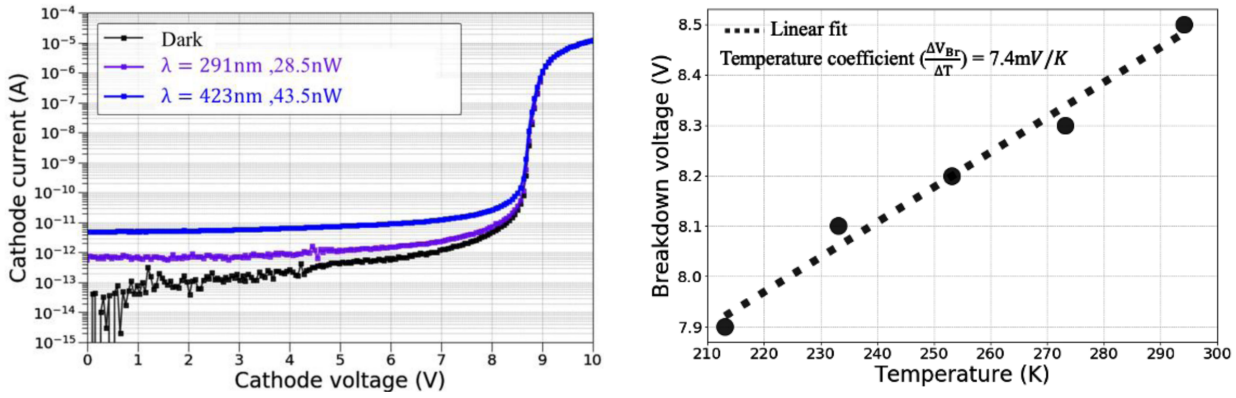
(a) (b)

Figure 38: **(a)** Quantum efficiency and photodetection probability at 3 V of overvoltage for the simulated device under frontside-illumination. **(b)** Quantum efficiency and photodetection probability at 3 V of overvoltage for the simulated device under backside-illumination

4.3 Characterization: First results

In [99], early attempts have been made to characterize the device breakdown voltage, PDP, DCR and afterpulse.

The first devices were made without the integration of a quench resistor, it is then easier to locate the breakdown point as the current rise is steeper. The I-V curves measured at the cathode is depicted on Figure 39a under dark and illuminated conditions at different wavelength and optical powers. The results match fairly well the simulation on Figure 33, confirming the model validity and the simulated electric field profile accuracy. Under illumination the dark current rises slightly which is expected. Because of the higher PDP and power, the light at λ =427 nm give rises to a slightly higher dark current than the 291 nm wavelength. The curves shows that the breakdown voltage is 8.5 V which is close from the one simulated at room temperature. As the temperature decrease, the breakdown voltage is expected to do the same because of the longer mean free path for carriers [72]. This trend for the device is shown on Figure 39b and the a variation of the breakdown voltage with the temperature of 7.4 mV/K was measured.



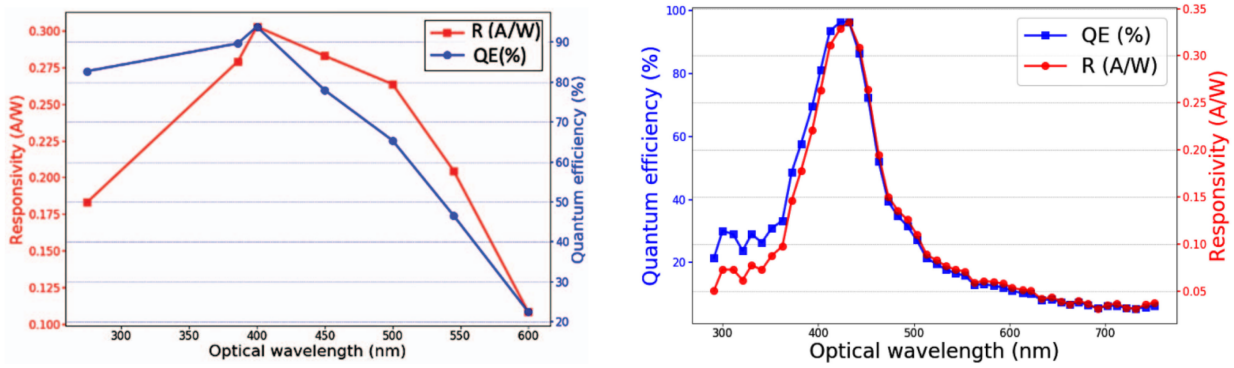
(a) (b)

Figure 39: **(a)** Current-voltage (I-V) measurement in the dark and under low light power at the wavelengths of 291nm and 423nm at room temperature **(b)** Breakdown voltage versus temperature. From [99]

Next, the quantum efficiency of the BSI SPAD that takes into account the light absorption and diffusion of carrier probability at the same time (see Equation 17) is measured. The method is similar to the one used for the simulations, however impact ionization cannot be "turned off" as for the quantum efficiency modelisation. Hence the SPAD cell is biased at 4 V to make sure that creation of carriers by impact ionization is minimized and the gain as defined in (37) does not deviate to far from one, going too low however would make the current measurement difficult. The light sources is given by LEDs of different wavelengths and powers while the cathode current is measured. One can also get the responsivity defined as the ratio of measured current by the optical power of the source.

The results are displayed on Figure 40 for frontside and backside illumination. The measurements shows good QE in the UV domain, better than the average state-of-theart device targetting the same range [98], [99] with a peak at $\lambda = 400$ nm and a QE of 82.75 % achieved at $\lambda = 275$ nm when frontside illuminated indicating that contrary to the simulated SPAD losses due to reflection are avoided somehow. In comparison, the BSI SPAD display a similar peak around the same wavelength but the performances drop more rapidly below. This is due to the UV photon being absorbed at the opposite side of the p-n junction, about 650 nm away and who has to diffuse that distance before getting the opportunity to induce an avalanche. This effect was negligible for the simulated device but defects and impurities arising from the real fabrication process would result in higher carrier recombination could explain this discrepancy. Moreover, eventhough the BOX layer beneath is theoretically nearly transparent, it is still thick ($\sim 1\mu$m) and absorbs some amount of light.

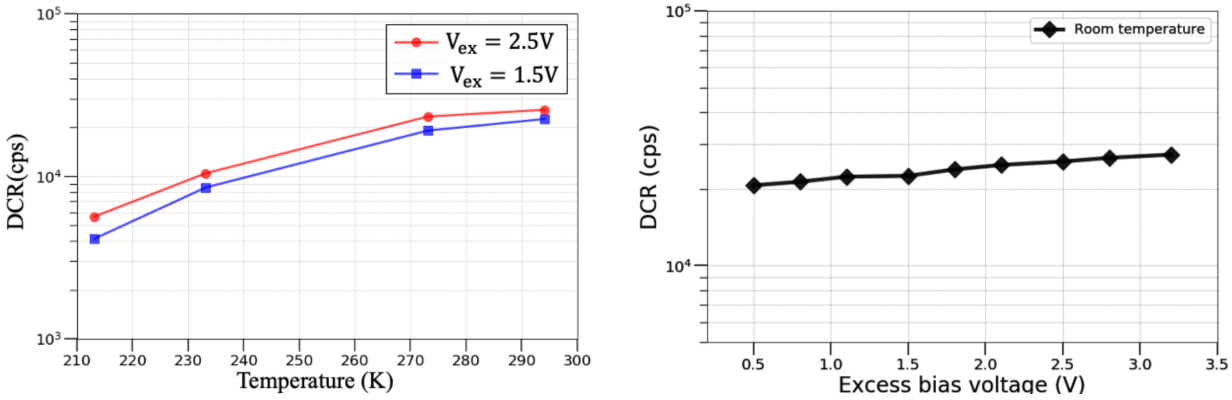
The PDP drop is more intense for longer wavelenghts as well in the BSI case. The reason here is that under fronstside illumination and for longer wavelengths that have their absorption depth below the thin silicon body between the SiO_2 and BOX layer (see Figure 30a), part of the light ray is reflected on the thick BOX and on the handle wafer beneath. This is an important difference for comparison with the computation done previously, the simulated device just had the BOX layer at the bottom with no handle wafer so part of the light could not be reflected back in the thin Si body that way. This recovery of the quantum efficiency by reflection is negligible under BSI: the ray travels through the BOX and suffers from the low light absorption in the thin silicon body of SPAD alone, reflection at the cathode contact is negligible since it actually covers only a small part of the photoactive area (see .e.g. Figure 31a).



(a) (b)

Figure 40: **(a)** Photoresponsivity and QE of the fabricated SPAD in the spectral range from 275 to 600 nm at 4V reverse bias under frontside illu mination at room temperature. **(b)** Quantum efficiency (QE) and responsivity (R) of the BSI in the spectral range from 291 to 751 nm at 4V reverse bias at room temperature.

Secondly, the DCR is analyzed with respect to the excess bias and temperature to investigate its physical origins. The DCR variation is displayed on Figure 41, the SPAD has been measured at room temperature and cooled down to -60°C, as for the excess voltage bias the latter ranges from 0.5 V to 3.25 V. The DCR with respect to temperature is shown for an overvoltage of 1.5 V and 2.5 V on Figure 41a; it is noticeable that the DCR is reduced at lower temperature and the slightly higher dependency on excess bias voltages of 1.5V and 2.5V at these temperatures is accentuated due to the BTB tunnelling contribution as discussed in subsection 3.3. Figure 41b shows that the DCR variation with the increase of excess bias voltage at room temperature exhibits a weak exponential dependency. This behavior can be related to a dominant contribution of trap-assisted process and lower band-to-band tunneling noise



(a) (b)

Figure 41: **(a)** DCR as a function of excess bias voltage at room temperature **(b)** DCR as a function of temperature at the excess bias voltages of 1.5V and 2.5V . From [99]

The BTB tunnelling main contribution at lower temperatures is further confirmed by looking at the Arrhenius plot in Figure 42. The linear fits exhibit an activation energy of 0.10 and 0.12 eV which is characteristic of the BTB tunnelling contribution as discussed in subsection 3.3. It is to be noted however that a corner temperature should be expected, which characterizes an increase of the activation energy at higher temperature indicating the predominance of TAT as suggested by Figure 41b.

Still, The DCR values at the excess bias voltages of 1.5V and 2.5V are 22.5 kcps and 25.62 kcps which divided by the SPAD 13μm $\times$ 13μm photoactive area (without guard-ring) gives 133.1 cps/μm^2 and 151.6 cps/ μm^2 at room temperature, respectively. This is relatively high compared to the median state-of-the art device of Table 1 (DCR/$\mu m^2 < 1$) but is in the same range of magnitude when compared to SPADs with the same breakdown voltage in the same table. In principle, in order to lower the DCR, the depletion layer should be widened by tuning the parameters during doping implantations and diffusion steps.

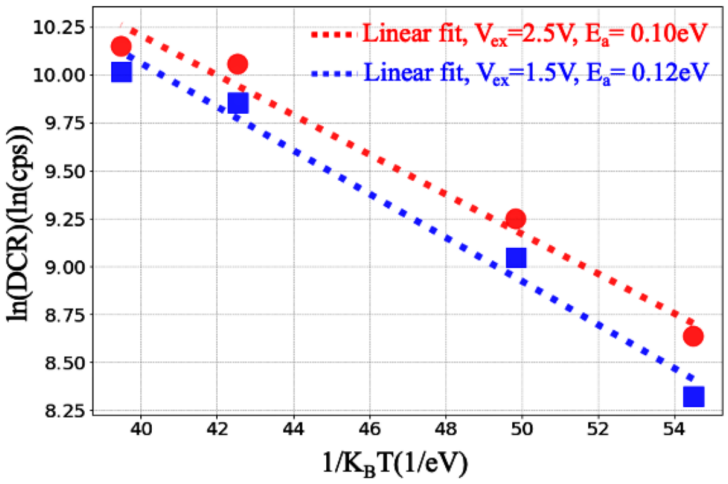


Figure 42: Arrhenius plots from of the DCR T=20°C to T=-60 °C at the excess bias voltages of 1.5V and 2.5V

5 Characterization of dark count on another device

In the previous section, the DCR was characterized. In the same paper [99] of which the same section is based, an attempt to characterize afterpulsing was also carried out. The main issue encountered was that the time resolution required to properly identify afterpulses ($\sim$ 100 ns) was not met due large pulse widths outputed by the shaping amplifier used during the measurement. However, it might be possible to get a better resolution as well as getting the pulse amplitudes by tuning the coupling capacitor used in the setup and signal processing methods, the setup and methods are developed further in subsection 5.2 and subsection 5.3. Additionally, it gives the opportunity to test other devices from different process batches and compare them with the results of subsection 4.3. Hence, the I-V curved are measured first for different types of devices.

5.1 I-V curves

For the three types of design presented in Table 2, the same fabrication process was performed on all wafers but SOI wafers coming from two different suppliers (Ultrasil [104] and SOITEC [105]) were used for the different batches. Additionally, half of the devices tested present a minor process variation: the trench shown in Figure 30a is not formed, instead the silicon surrounding the anode ring is etched leaving a so called 'mesa' structure. Accounting for the supplier and process variation, this gives twelve types of devices of which the I-V curves are displayed on Figure 43.

I-V curves and avalanche breakdown were measured for devices without quench resistor using an E4727B analyzer that can measure current in the pA range [106], It can be noted that the devices fabricated on the Ultrasil SOI wafer have a much higher dark current than expected, on the other hand the SOITEC devices have their dark current in the expected range of $10^{-13} \sim 10^{-12}$A, apart maybe for two devices. This could

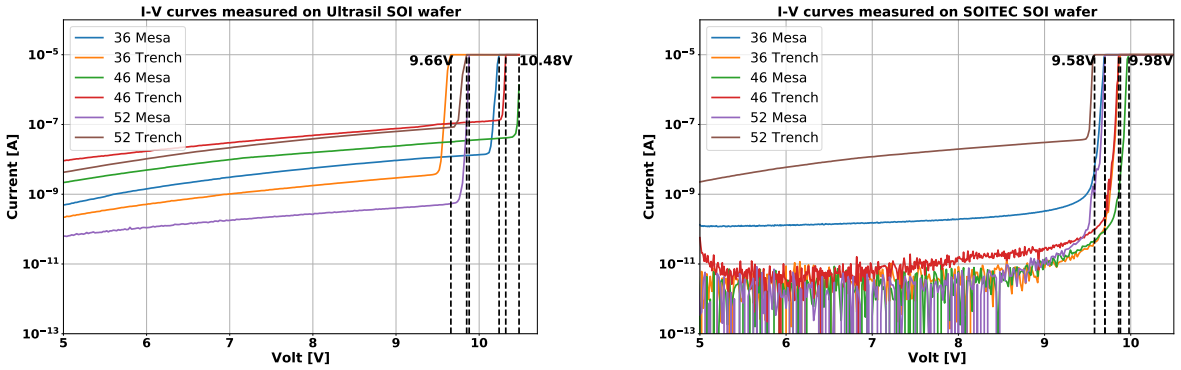
be due to the different BOX layer thickness of the SOI wafer, indeed there is some uncertainties of the BOX thickness across the Ultrasil wafer which can be noticed by the variation of color depending on the device location [107]. By contrast, the SOITEC substrate is perfectly uniform which is characterized by the same green color present all over the wafer where the BOX is exposed.

The breakdown voltage taken when the current reaches the compliance at 100 μA is a bit more uniform for the SOITEC wafer, close to 9.8 V, while for the other supplier it can exceed 10 V, typically for devices with the highest dark currents. This implies that the surface quality has some effects on the device properties and higher dark currents and more breakdown voltage variations are to be expected using Ultrasil wafers. Other than this parameter, it appears that the device dimensions and type ('mesa' or 'trench') has no significative influence.

Finally, compared to the device characterized in subsection 4.3, the breakdown voltage has been shifted from 8.6 V to about 9.8 V in average. This would typically imply a wider depletion region that could arise from a lower dose during one of the ion implantation process step and needs to be investigated. Still, the breakdown voltages are relatively not too far off from 8.6 V and more important, they do not vary too much from one another. Table of the breakdown voltages for the devices measured can be viewed in Table 4

	36 Trench	**46 Trench**	**52 Trench**	**36 Mesa**	**46 Mesa**	**52 Mesa**	**Average**
Ultrasil	9.66	10.32	9.85	10.24	10.48	9.88	**10.07 V**
SOITEC	9.88	9.86	9.58	9.70	9.98	9.70	**9.78 V**

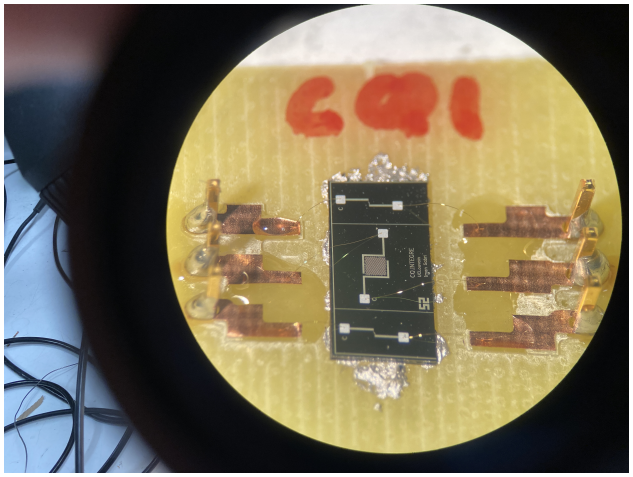
Table 4: Measured breakdown voltages for six types of devices for two different SOI wafer suppliers. The breakdown voltages are taken when the current compliance is reached at 100 μA



(a) (b)

Figure 43: **(a)** I-V curves of the devices measured on the Ultrasil SOI wafer. The dashed line shows the breakdown point for each SPAD, which range from 9.66 V to 10.48 V **(b)** I-V curves of the devices measured on the SOITEC SOI wafer. The breakdown voltages range from 9.58 V to 9.98 V

Next, some of the SPAD cells and SiPM arrays were removed from the wafer and wirebonded as can be visulaized in Figure **??** as an example.

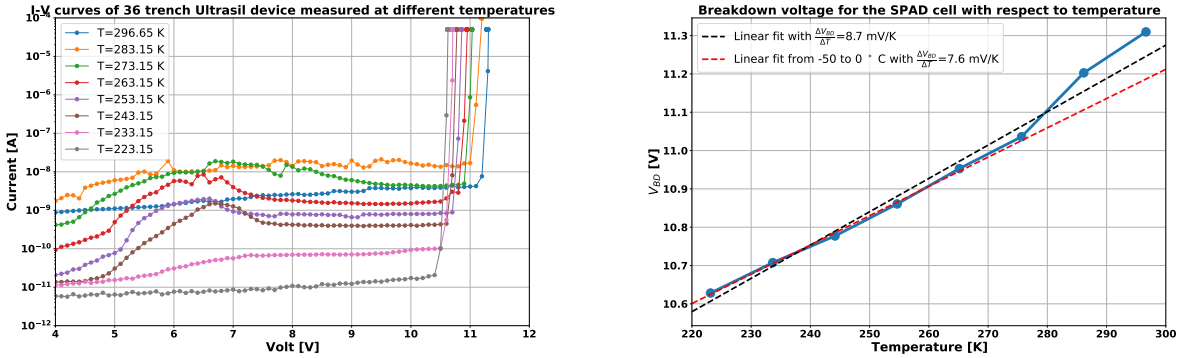


[Wirebonded SPAD device]Picture showing the wirebonded hexagonal (52) with trench SPAD type coming from an Ultrasil wafer. The top device is a SPAD cell without integrated quenching resistor, the middle one is a SiPM array where each micro-cell has its own quenching resistor and the bottom SPAD cell also has an integrated quenching resistor. The bonding pads are accesses to the cathode and anode.

Measurement of the breakdown voltages with respect to the temperature were made for a 36 Trench Ultrasil wirebonded SPAD cell first, of which the results are displayed on Figure 44. These measurements were made primarily so that when measuring dark counts in subsection 5.4 at different temperatures the same excess bias could always be applied for better comparison. The device without integrated quenching resistor was placed in a ESPEC SH-261 temperature chamber [108] and the temperature was cooled down to -50°C with a decrement of 10°C for each step. It was also ensured that SPAD was kept in the dark inside the chamber.

Figure 44 shows the resulting I-V curves and breakdown voltages for the different temperatures. The dark current decreases when lowering the temperature and the breakdown voltage does so linearly as expected. The breakdown voltage vs temperature graph displayed on Figure 44b exhibits a temperature coefficient of 8.7 mV/K which is close to the one obtained in [99] (equal to 7.4 mV/K shown in Figure 39b). It is interesting to notice when taking the measurement values from -50°C to 0°C only, we get a fit dislpaying even greater linearity with a temperature coefficient of 7.6 mV/K which nearly perfectly matches the result of [99].

It is to be noted that the breakdown voltage is further shifted to 11.3 V at room temperature. Other wirebonded devices measured shows the same behavior but the cause is still to be determined.



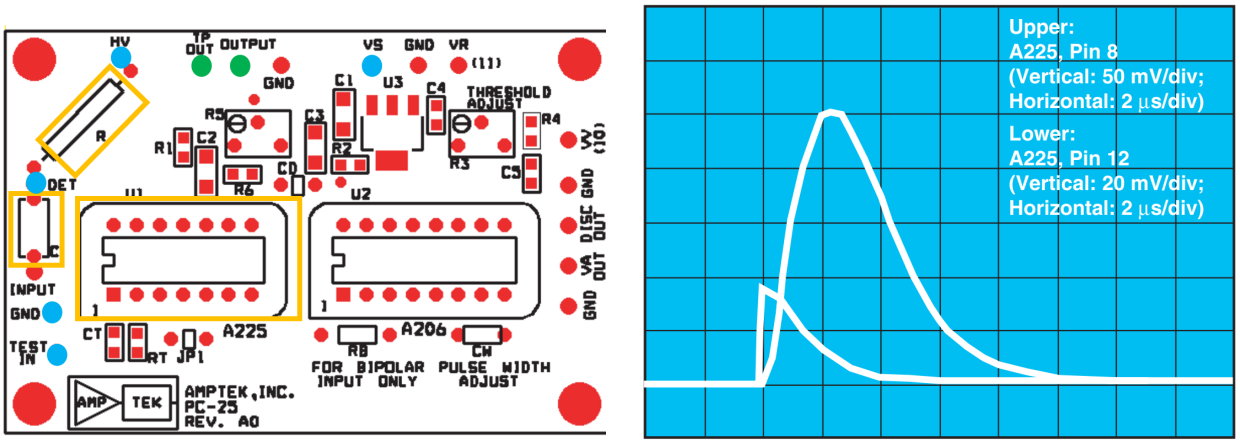
(a) (b)

Figure 44: **(a)** I-V curves of the 36 Trench Ultrasil SPAD measured under dark conditions at various temperatures **(b)** Breakdown voltage vs temperature linear fit of the same device. The breakdown points go from 11.31 V at room temperature to 10.63 V at -50°C

5.2 PCB setup and A225 shaping amplifier

For the purpose of capturing the SPAD output pulse, the PC-25 printed circuit board from Amptek [109] coupled with the A225 charge sensitive amplifier and shaping amplifier is used.

The printed circuit board (PCB) layout is shown on Figure 45a with the useful inputs, outputs and components related to the measurements. These are for inputs: the detector bias voltage source (`HV`), the detector input itself delivering the pulses (`DET`), the power source of the A225 amplifier (`VS`) and additionally an input to test circuit that can be connected to an arbitrary waveform generator for example. Besides the components already integrated to the PCB, a 100 kΩ resistor `R` is placed and acts in this setup as an external quench resistor for the device under study. A high voltage capacitor `C` also needs to be placed in order to feed its discharge to the A225 shaping amplifier, the capacitance can also be selected depending on our requirements and 1 nF was chosen in our case. Finally, to the shaping amplifier corresponds two outputs: `TPOUT` for the timing pulse output and `OUTPUT` for the shaping amplifier output, the latter will be most used since it is more suitable for pulse height analysis. The amplifier output pulse waveforms are depicted on Figure 45b.



(a) (b)

Figure 45: **(a)** PC-25 circuit board layout, the main inputs (`HV, DET, TEST IN,VS`) and ground connections employed are shown in blue, the important tunable components (`R,C`) are framed in yellow with the shaping amplifier and the green dots correspond to the two pulse outputs of the A225. **(b)** Pulse outputs of the A225 shaping amplifier, the Gaussian waveform (Pin 8) is the shaping amplifier output (`OUTPUT`) and the alternative (Pin 12) is the timing pulse output `TPOUT`.

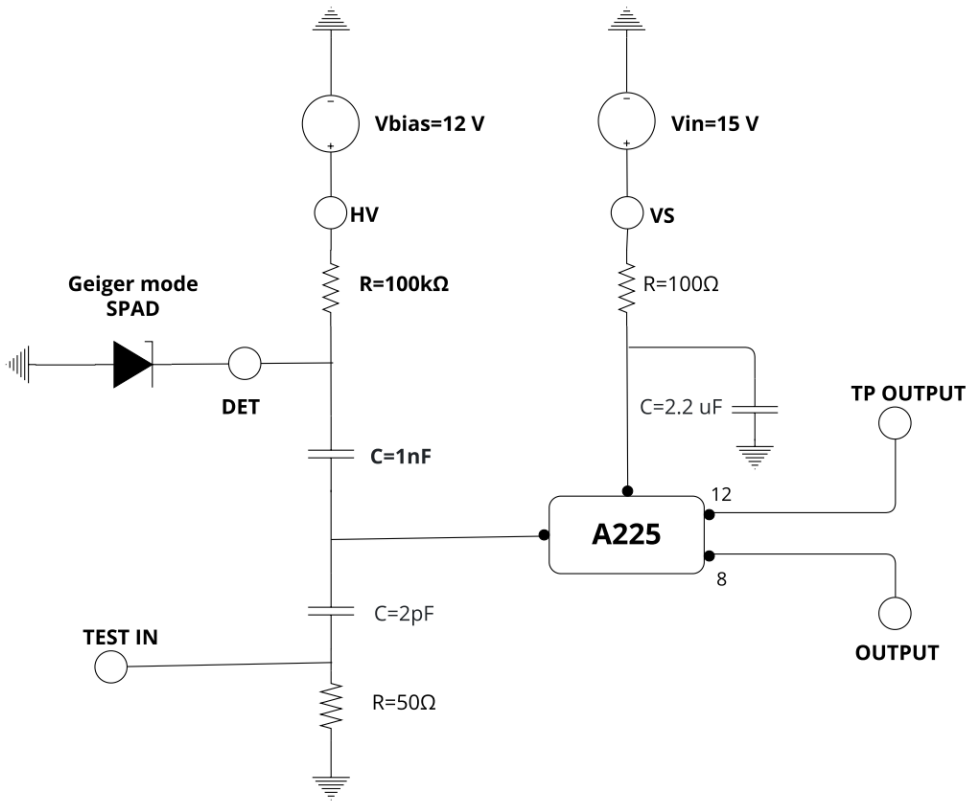


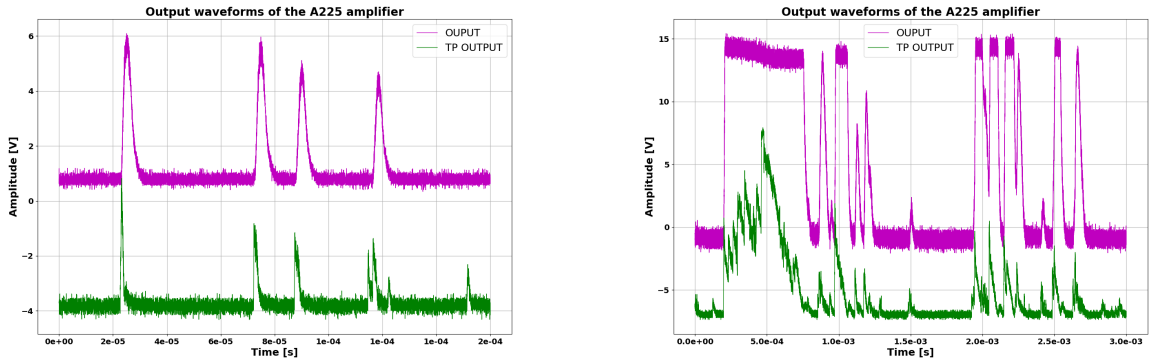
Figure 46: Circuit schematic of the setup used to capture the Geiger mode SPAD output pulse

The circuit schematic corresponding to the setup that has just been discussed is depicted on Figure 46. The A225 amplifier is powered by a source meter at 15 V through the `VS` input and the current is monitored as well to ensure the effective operation of the device.

A second source is required to bias the SPAD cell in Geiger mode, to do so the detector is biased positively at the cathode with another source meter through `HV` and the 100 kΩ `R` resistor acting as the external quenching resistor. This is the primary reason why a device without integrated quenching resistor was used: it allows to pick up directly the voltage waveform with its falling edge and characterized by (34) in the so called voltage output mode discussed in subsection 3.4. This is as a first step advantageous as the output pulse would not suffer of amplitude reduction caused by a voltage divider part (see Figure 29a). It is also conceivable to make the setup work in a current output mode as shown in Figure 29b with some modifications but the A225 shaping amplifier is designed to work with negative inputs polarities (.i.e. the input signal needs to have a falling leading edge) which is more difficult to adapt to the current mode output pulse with rising front edge.

The coupling capacitor (1 nF) is important to tune the charge feeding the capacitor and consequently the amplitude of the shaping amplifier output: the A225 has a fixed charge sensitivity of 5.2 V/pC [109] and the capacitor discharge is transferred to the amplifier input. Thus if the capacitance is too small for a given voltage step, the charge transfer will be insignificant and the shaping amplifier output pulse will not distinguishable from the amplifier noise. On the other hand if the capacitance is too high, the amplifier becomes too sensitive and the output signal is always saturated, limited by the 15 V source meter and it is not possible to accurately get the input pulse amplitude. The A225 can also be tested with a pulser on the `TEST IN` input where a small 2 pF test capacitor subsequently injects a test charge into the amplifier.

The output waveforms were acquired using a Tektronix MSO2024B oscilloscope with a sampling period of 5 ns. The waveforms are depicted on Figure 47 which shows outputs that are similar to Figure 45b. Figure 47a exhibits 'ideal' pulse outputs for the device under study at 0.8 V of excess bias. On Figure 47b on the other hand is illustrated the saturation case when the coupling capacitor is not adapted to the SPAD voltage step (here with an overvoltage of 5V using an Hamamatsu S13360-1325CS test device [58]).



(a) (b)

Figure 47: **(a)** Outputs of the A225 for the SPAD detector with 0.8 V of excess bias. The pulses correspond to dark counts and show that the coupling capacitor is well adapted to the device for this particular reverse bias. **(b)** Outputs of the A225 for the Hamamatsu test device reverse biased at 60 V (5 V of excess bias). The waveforms are characteristics of a saturated amplifier, implying that either the coupling capacitor should be smaller or the overvoltage should be diminished.

5.3 Pulse analysis and signal processing

Once the pulses shown in Figure 47a are acquired, we want to identify them further by obtaining their time stamps and heights. Signal analysis was carried out in a post-processing manner with Python, the peaks of the shaping amplifier output are identified and their true heights can be extracted using a pulse pile-up correction method.

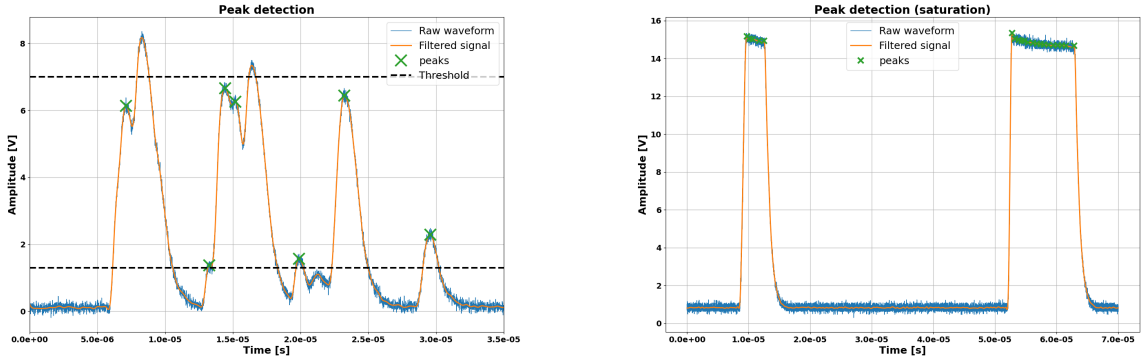
5.3.1 Peak detection

First a method is implemented to identify the pulses and their location in time domain.

The method essentially relies on locating the maximums (peaks) of the waveform, however in order for this method to be efficient, the signal must be initially smoothened. The SciPy library [110] was the main tool used for this purpose and signal processing in general. It is found that an IIR Butterworth low-pass filter method works best for this matter (`scipy.signal.iirfilter` method). The cutt-off frequency used is 2 MHz, which is a good compromise so the method filters out the noise without missing any peaks. It was noticed that initially the filtered signal was shifted compared to the raw waveform, to correct this the 'second-order sections' filter form output must be specified in the method.

Once the raw signal is filtered in an appropriate manner, it is straightforward to get the peaks using the `scipy.signal.find_peaks` method from the SciPy library. The SciPy function is quite robust and allows to specify useful parameters such the as upper and lower thresholds and the minimal time step between consecutive peaks. The applied filter and peak-finding method results can be visualized on Figure 48a, which allows to get the $(time, amplitude)$ locations for each peak.

One must be careful when applying this procedure to an acquired waveform because even though the selected coupling capacitor matches well our device requirements, saturated output pulses, although not frequent, can still be witnessed sometimes. When this happens, the method will overestimate the number of peaks detected since the flat pulse crest of the saturated signal is actually wavy when filtered and several peaks are mistakenly detected as can be seen on Figure 48b. Therefore, considering that saturation occurs for a minimal amplitude of V=15 V, the upper threshold value is chosen to be 14 V to leave a margin. Peaks above 14 V displaying saturation were still accounted for during the signal analysis, for the saturated pulses a method was implemented such that these are counted as one when the waveform rises above the 14 V threshold.



(a) (b)

Figure 48: **(a)** Application of the peak-finding method `scipy.signal.find_peaks` to the filtered signal obtained with `scipy.signal.iirfilter`. The peaks outside the lower and upper threshold bounds are ignored **(b)** Application of the peak-finding method to two saturated output pulses. The method return more peaks than expected because of the pulse waviness during the 'on' state.

5.3.2 Pulse pile-up correction

The peak heights can be obtained using the previous method, however as shown on Figure 48a, the signal can be subject to pile-up when two pulses overlap. The following procedure intends to correct this issue and get the true pulse height based on the shaping amplifier characteristic output function.

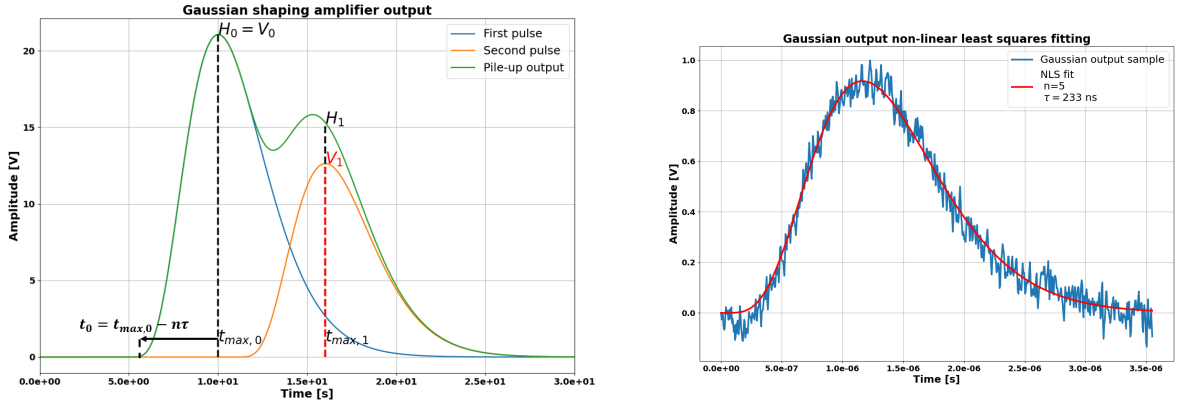
The A225 output `OUTPUT` is basically a semi-Gaussian pulse shaping amplifier, meaning that it is essentially a single CR high-pass filter which is followed by several stages of RC integrators. Amplifiers shaping in this way are have their output pulse and peak rising time given by [111]:

$$V_{OUT}(t) = V_0 \left(\frac{t}{\tau}\right)^n e^{-t/\tau}, \qquad t_{max} = n\tau. \tag{38}$$

where n corresponds to the number of RC integrators and $t_{max} = n\tau$ is the peaking time. The idea behind the pile-up correction method is represented on Figure 49a. Based on the peak locations ($(t_{max,i}, H_i)$ on the figure) and (38) it is possible to deduct the true height V_i of each pulse as:

$$V_i = H_i - V_{i-1} \left(\frac{t_{max,i} - t_{max,i-1} - n\tau}{\tau}\right)^n e^{-(t_{max,i} - t_{max,i-1} - n\tau)/\tau} \tag{39}$$

However, the parameters τ and n of the Gaussian shaping amplifier are to be determined first. This is done with the help of a non-linear least squares (NLS) algorithm from the SciPy library (`scipy.optimize.curve_fit`) that is applied to a normalized sample pulse so that it can find the parameter τ that fits best the data. The method requires a model function as an argument which is given by Equation 38. The integer parameter n is obtained simply through trial and errors. It is found that $n = 5$ and $\tau = 233$ ns minimize the least square error of the fit shown on Figure 49b. It is to be noted that the parameters have been found with one sample pulse, repeating the same process on multiple samples and taking the average for τ could give a value that is more representative.



(a) (b)

Figure 49: **(a)** Illustration of the pile-up correction method with two pulses. The output waveform is displayed in green and the corrected waveform of the second pulse with amplitude V_1 is shown in orange. **(b)** Non-linear least square fit of a normalized sample Gaussian pulse.

The pile-up correction method is tested on the waveform depicted on Figure 48a and the result is displayed on Figure 50. The method returns satisfying results, and the corrected Gaussian waveforms fit the original signal well enough. It was ensured that the correction procedure is robust enough so that it can be applied to longer samples with thousands of pulses. It is also adapted when a saturated pulse is encountered, in that case the amplitude is simply automatically set at 15 V. The time resolution (.i.e. minimum detectable time interval between pulses) is still limited however by the shaping time constant τ: if a second pulse occurs after a first one within a time interval smaller than the peaking time ($\Delta_t < n\tau = 1.165\mu$s), it will only be added-up to the first one and the two will be indistinguishable. This can be troublesome for afterpulsing detection given that the release time of a trap is in the hundreds of nanosecond order as disscussed in subsubsection 3.3.4.

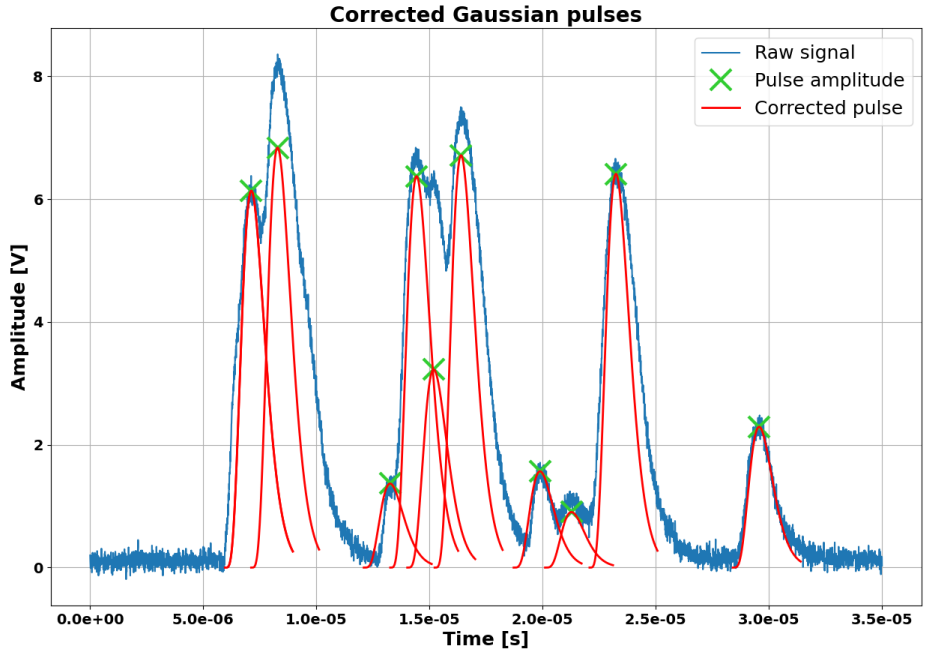


Figure 50: Waveform with application of the pile-up correction method. The corrected Gaussian pulses are displayed in red.

5.4 DCR characterization VS temperature

5.4.1 Inter-arrival times and pulse amplitude histograms

An attempt to characterize the DCR for temperature ranging from -50°C to 23.5 °C (same range as in subsection 5.1) was made in dark conditions inside the temperature chamber, or a test fixture for the room temperature measurement. For each temperature step, an overvoltage of 0.8 V was applied such that the A225 amplifier is not saturated. Applying the pulse processing method of the previous section, we can get the inter-arrival time between each pulse as well as their amplitude. The histograms for the device measured at room temperature are displayed on Figure 51, the same histogram shape is witnessed at any given temperature.

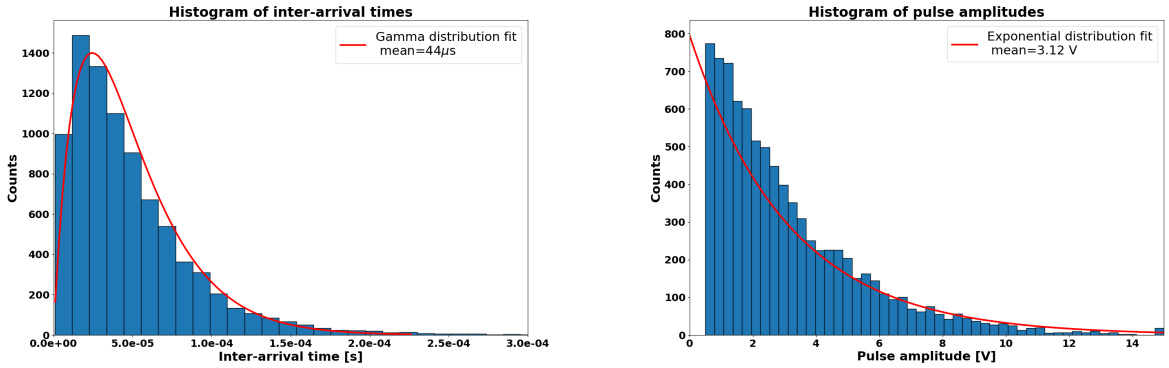
Based on the assumption that the DCR is Poissonian (each count is independent and characterized by a exponential probability density function, see subsection 3.3), it can be demonstrated [112] that inter-arrival times are characterized by a Gamma distribution with shape parameter $\alpha = 2$:

$$f(t_{inter}; 2, \beta) = t_{inter} e^{-\beta t_{inter}} \beta^2 \tag{40}$$

where the mean value is given by $\mu = \frac{2}{\beta}$. The histogram of Figure 51a matches well this distribution, validating the experimental setup and peak-finding method.

In contrast, the histogram distribution of pulse amplitudes is more peculiar and seems to follow an exponential distribution law. Indeed, the pulse amplitudes are expected to be more or less constant and the histogram should exhibit a plateau corresponding to a voltage step given by the excess bias .i.e. 1 photoelectron [113]. Higher pulses could be detected in the case of pile-up for an inter-arrival time $< 1.165\mu$s due

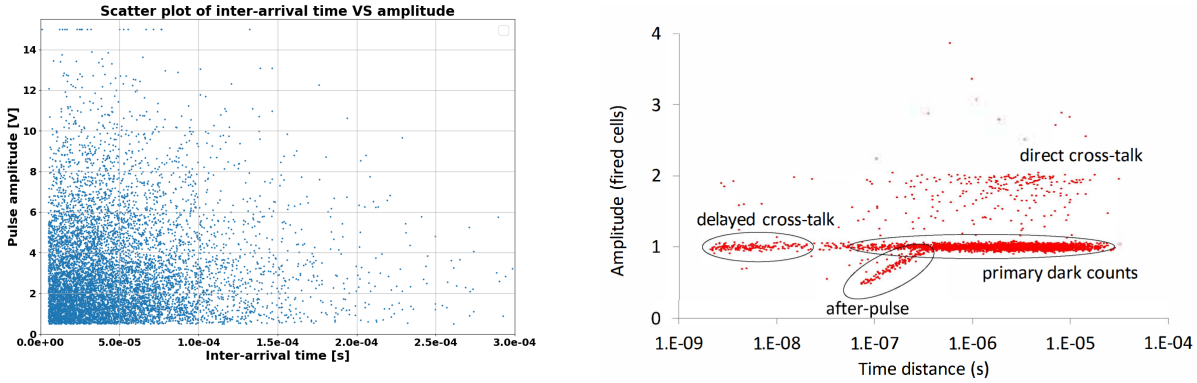
to the slow shaping time constant of the amplifier. However smaller pulses are more common which could in principle be explain by the slow recharge time of the SPAD T_r (see (34)). If this is the case, there should be a correlation between the inter-arrival time and the pulse height and thus for short inter-arrival times should correspond small amplitudes.



(a) (b)

Figure 51: **(a)** Histogram of inter-arrival times between consecutive pulses. The datas are fitted with a gamma probability density function. **(b)** Histogram of pulse amplitudes. The datas are fitted to an exponential distribution function.

The scatter plot of inter-arrival time vs pulse amplitude of Figure 52a does not display such correlation: it is true that to short inter-arrival times corresponds small pulses but this also the case for even longer time duration where small pulses are still more likely to occur. In comparison, Figure 52b shows the expected scatter plot (although in this case for a SiPM) characterized by a constant primary dark count plateau with 1 p.e amplitude and smaller pulse heights corresponding to afterpulses (for a amplifier with more accurate time resolution, afterpulses have a smaller height since they take place typically within the SPAD cell recharge time constant).



(a) (b)

Figure 52: **(a)** Scatter plot of inter-arrival time VS pulse amplitude for the device under study measured at room temperature. **(b)** Scatter plot of the pulse amplitude as a function its distance from the preceding event. For a SiPM, primary dark counts, afterpulses and crosstalks events can be characterized from the plot. From [113]

5.4.2 Characterization of the DCR with respect to the temperature

The DCR is computed based of the average inter-arrival time of the data, its evolution with respect to the temperature is displayed on Figure 53. The curve do not exhibit the expected behavior, the DCR should decrease in an exponential fashion because of the slower trap activity (see (24)) or be similar to the characterization done in Figure 41a. Instead it drops abruptly below room temperature and remains relatively constant about 16 kcps.

This would imply some mishap with the temperature chamber: the first measurement was made at -50 °C and the temperature was increased in between each measure. It could be possible that the device could not adapt its internal temperature fast enough as the chamber would have been heated up too rapidly. As for the DCR value at room temperature, it is clearly higher as this characterization was made inside the test fixture before even using the temperature chamber. It is interesting to note that the reverse path (cooling the device down) did not seem to be an issue during the breakdown voltage characterization in subsection 5.1 as suggested by the reasonable values obtained in Figure 44a.

Still, at room temperature and for an overvoltage of 0.8 V the DCR is 23 kcps for an photoactive area (without guard-ring) of 13 μm $\times 13\mu$m and thus the DCR density is 136 cps/μm^2 which is about the same value as in subsection 4.3 but for an excess bias of 2.5 V.

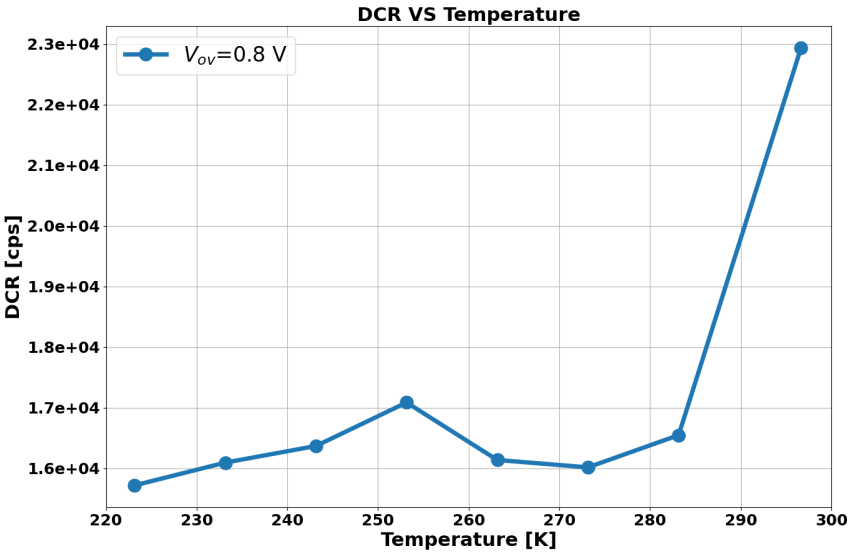


Figure 53: DCR as a function of temperature at the excess bias voltage of 0.8 V

6 Conclusion and future prospects

Ref	Technology	Peak PDP	DCR (RT) [cps/μm^2]	PAP	Tdead [ns]	Jitter [ps]	V_{BD} [V]	V_{ov} [V]	FOMC [m]	FOMT [m/s]
[45] (2019)	0.14 μm CMOS BSI SOI	26.4 % (490 nm)	396.1	<0.1%	200	65 ps (405 nm)	11.3	3	0.066	1020
[54] (2020)	65 nm CMOS	50% (480 nm)	73	1%*	55*	158 (485 nm)	9.6	1	1.05	6666
[55] (2018)	0.18 μm CMOS	55%(400 nm)	0.23	1%*	100	23 (410 nm)	22.1	2	11.3	493635
[56] (2016)	NA	51.2%(400 nm)	0.75	5%	100	160 (420 nm)	26	3	5.62	35103
[57] (2014)	0.35 μm CMOS	55 % (420 nm)	0.079	4%	20	75 (780 nm)	25	6	94	1252359
[58] (2016)	NA	54% (450 nm)	0.55	1%*	55*	75*	53	3	13.1	174753
[59] (2017)	NA	50.5%(420 nm)	0.5	0.75%	45	75*	24.5	2.5	15.8	210686
[60] (2021)	NA	67% (420 nm)	0.71	<1 %	55	75*	32.5	6	14.5	192762
This device (36 Trench BSI)	BSI SOI	60%* (423 nm)	93	?	?	?	11.2	0.8	?	?

Table 5: Comparison of the SPAD performances for timing and counting applications targetting smaller wavelength ranges. The PDP of the device is taken at an excess bias of 2 V in [99]

Above is shown the updated Table 1 for state-of-the-art SPAD cells targetting blue or UV, sub-UV wavelength ranges; measured parameters for our device has been added. Several properties are still to be measured in order to fully characterize the device.

For afterpulsing, the characteristic time constant of the shaping amplifier limits its resolution and for the dead time, the A225 with PCB output linearity must be thoroughly investigated to explain the results of Figure 52a as other variables might influence the pulse height other than just the amplitude of the input voltage step. Alternatively, a B1500A semiconductor analyzer with Waveform generator/Fast measurement unit module installed is at disposal [114]. The module has a sampling time of 50 ns and could theoretically measure the current output pulse directly. Moreover, its waveform generation capabilities makes it ideal for afterpulsing characterization with TCCC [94]. The main downside is that module can deliver a voltage of 10 V peak to peak maximum which is not sufficient given that the device exhibits now a breakdown point above 10 V. But it could be considered to apply a voltage offset to the device somehow.

The timing jitter would be typically measured by Time Correlated Single-Photon Counting (TCSPC) [115], which requires a pico or femtosecond LASER that emits in the wavelength range that could be detected by the SPAD.

More generally, the same types of measurement performed so far should be carried out on other devices as only a few have been fully characterized. This would allow to assess the reproducibility of the fabrication process and see if the slight variations of the designs has any effect on the performances. Especially, tests on SPADs with integrated quenching resistor and SiPM arrays must be done.

Looking even further, a focus on a digital SiPM implementation, front-end electronics and active quenching circuit design would be mandatory for imaging applications [13], [116], [117]. Particularly for PET, everything from the SPAD cell to the software programming needs to be optimized for efficient coincidence detection with scintillators and good imaging capabilities [14], [50], [118], so there are still plenty of milestones to reach before getting there.

References

[1] K. G. McKay, "Avalanche breakdown in silicon," *Phys. Rev.*, vol. 94, pp. 877–884, 4 May 1954. DOI: 10.1103/PhysRev.94.877.

[2] A. G. Chynoweth and K. G. McKay, "Internal field emission in silicon $p - n$ junctions," *Phys. Rev.*, vol. 106, pp. 418–426, 3 May 1957. DOI: 10.1103/PhysRev.106.418.

[3] K. S. Champlin, "Microplasma fluctuations in silicon," *Journal of Applied Physics*, vol. 30, no. 7, pp. 1039–1050, 1959. DOI: 10.1063/1.1776976.

[4] R. J. McIntyre, "Theory of microplasma instability in silicon," *Journal of Applied Physics*, vol. 32, no. 6, pp. 983–995, 1961. DOI: 10.1063/1.1736199.

[5] R. H. Haitz, "Model for the electrical behavior of a microplasma," *Journal of Applied Physics*, vol. 35, no. 5, pp. 1370–1376, 1964. DOI: 10.1063/1.1713636.

[6] R. H. Haitz, "Mechanisms contributing to the noise pulse rate of avalanche diodes," *Journal of Applied Physics*, vol. 36, no. 10, pp. 3123–3131, 1965. DOI: 10.1063/1.1702936.

[7] H. Ruegg, "An optimized avalanche photodiode," *IEEE Transactions on Electron Devices*, vol. 14, no. 5, pp. 239–251, 1967. DOI: 10.1109/T-ED.1967.15937.

[8] G.-F. Dalla Betta, L. Pancheri, D. Stoppa, H. Robert, and R. Justin, "Avalanche photodiodes in submicron cmos technologies for high-sensitivity imaging," *Advances in Photodiodes*, G.-F. Dalla Betta, Ed., London: IntechOpen, 2011, ch. 11, pp. 225–248. DOI: 10.5772/15178.

[9] D. Renker, "Geiger-mode avalanche photodiodes, history, properties and problems," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 567, no. 1, pp. 48–56, 2006. DOI: https://doi.org/10.1016/j.nima.2006.05.060.

[10] A. Rochas, M. Gosch, A. Serov, *et al.*, "First fully integrated 2-d array of single-photon detectors in standard cmos technology," *IEEE Photonics Technology Letters*, vol. 15, no. 7, pp. 963–965, 2003. DOI: 10.1109/LPT.2003.813387.

[11] Y. Li and J. Ibanez-Guzman, "Lidar for autonomous driving: The principles, challenges, and trends for automotive lidar and perception systems," *IEEE Signal Processing Magazine*, vol. 37, no. 4, pp. 50–61, 2020. DOI: 10.1109/MSP.2020.2973615.

[12] 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. 2 000 102, 2020. DOI: 10.1002/qute.202000102.

[13] C. Bruschini, H. Homulle, I. M. Antolovic, S. Burri, and E. Charbon, "Single-photon avalanche diode imagers in biophotonics: Review and outlook," *Light: Science & Applications*, vol. 8, 87, Dec. 2019. DOI: 10.1038/s41377-019-0191-5.

[14] C. Bruschini, E. Charbon, C. Veerappan, *et al.*, "Spadnet: A fully digital, scalable, and networked photonic component for time-of-flight pet applications," vol. 9129, May 2014, p. 912 913. DOI: 10.1117/12.2051952.

[15] P. Lecoq and S. Gundacker, "Sipm applications in positron emission tomography: Toward ultimate pet time-of-flight resolution," *The European Physical Journal Plus*, vol. 136, no. 3, p. 292, Mar. 2021. DOI: 10.1140/epjp/s13360-021-01183-8.

[16] D. Bronzi, "Front-side and back-side illuminated spad arrays for 2d imaging and 3d ranging," Ph.D. dissertation, Politecnico di Milano, Nov. 2014. DOI: 10.13140/RG.2.1.3329.9047/1.

[17] SensL Technologies Ltd., *An introduction to the silicon photomultiplier (sipm)*, 2017. [Online]. Available: https://www.onsemi.com/pub/collateral/and9770-d.pdf (visited on 06/18/2022).

[18] F. Zappa, A. Tosi, A. Dalla Mora, and S. Tisa, "Spice modeling of single photon avalanche diodes," *Sensors and Actuators A: Physical*, vol. 153, no. 2, pp. 197–204, 2009. DOI: https://doi.org/10.1016/j.sna.2009.05.007.

[19] E. Nocerino, "The semiconductor multiplication system for photoelectrons in a vacuum silicon photomultiplier tube and related front end electronics," Ph.D. dissertation, University of Naples Federico II, Nov. 2016.

[20] T. Frach, G. Prescher, C. Degenhardt, R. de Gruyter, A. Schmitz, and R. Ballizany, "The digital silicon photomultiplier — principle of operation and intrinsic detector performance," *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*, 2009, pp. 1959–1965. DOI: 10.1109/NSSMIC.2009.5402143.

[21] F. Acerbi and S. Gundacker, "Understanding and simulating sipms," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 926, pp. 16–35, 2019, Silicon Photomultipliers: Technology, Characterisation and Applications. DOI: https://doi.org/10.1016/j.nima.2018.11.118.

[22] I. Malass, W. Uhring, J.-P. Le Normand, N. Dumas, and F. Dadouche, "Evaluation of size influence on performance figures of a single photon avalanche diode fabricated in a 180 nm standard cmos technology," *Analog Integrated Circuits and Signal Processing*, vol. 89, pp. 69–76, Oct. 2016. DOI: 10.1007/s10470-016-0763-8.

[23] M. Tyagi, *Introduction to Semiconductor Materials and Devices*. Wiley India Pvt. Limited, 2008.

[24] S. Jahromi and J. Kostamovaara, "Timing and probability of crosstalk in a dense cmos spad array in pulsed tof applications," *Opt. Express*, vol. 26, no. 16, pp. 20 622–20 632, Aug. 2018. DOI: 10.1364/OE.26.020622.

[25] W. Kindt, H. van Zeijl, and S. Middelhoek, "Optical cross talk in geiger mode avalanche photodiode arrays: Modeling, prevention and measurement," *28th European Solid-State Device Research Conference*, 1998, pp. 192–195.

[26] A. Zhukov, E. Popova, and N. Gerasimenko, "Methods for suppressing optical crosstalk between the cells of a silicon photomultiplier array," *Semiconductors*, vol. 50, no. 13, pp. 1729–1733, 2016. DOI: https://doi.org/10.1134/S1063782616130133.

[27] W. Yu, Y. Zhang, M. Xu, and X. Lu, "Dark count in single-photon avalanche diodes: A novel statistical behavioral model∗," *Chinese Physics B*, vol. 29, no. 4, p. 048 503, Apr. 2020. DOI: 10.1088/1674-1056/ab7808.

[28] A. Lacaita, M. Mastrapasqua, M. Ghioni, and S. Vanoli, "Observation of avalanche propagation by multiplication assisted diffusion in p-n junctions," *Applied Physics Letters*, vol. 57, pp. 489–491, 1990. DOI: 10.1063/1.103629.

[29] 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.

[30] S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa, "Avalanche photodiodes and quenching circuits for single-photon detection," *Applied Optics*, vol. 35, no. 12, pp. 1956–1976, Apr. 1996. DOI: 10.1364/AO.35.001956.

[31] M. Ghioni, A. Gulinatti, I. Rech, F. Zappa, and S. Cova, "Progress in silicon single-photon avalanche diodes," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 13, no. 4, pp. 852–862, Jul. 2007. DOI: 10.1109/JSTQE.2007.902088.

[32] A. Lacaita, S. Cova, M. Ghioni, and F. Zappa, "Single-photon avalanche diode with ultrafast pulse response free from slow tails," *IEEE Electron Device Letters*, vol. 14, no. 7, pp. 360–362, 1993. DOI: 10.1109/55.225573.

[33] A. Spinelli, M. Ghioni, S. Cova, and L. Davis, "Avalanche detector with ultraclean response for time-resolved photon counting," *IEEE Journal of Quantum Electronics*, vol. 34, no. 5, pp. 817–821, 1998. DOI: 10.1109/3.668769.

[34] "Sap500: Silicon apds for photon counting applications," *Photonics News 24*, 2019. [Online]. Available: https://www.lasercomponents.com/us/news/sap500-silicon-apds-for-photon-counting-applications-1/.

[35] H. Xu, L. Pancheri, G.-F. Dalla Betta, and D. Stoppa, "Design and characterization of a p+/n-well spad array in 150nm cmos process," *Opt. Express*, vol. 25, no. 11, pp. 12 765–12 778, May 2017. DOI: 10.1364/OE.25.012765.

[36] M. Gersbach, J. Richardson, E. Mazaleyrat, *et al.*, "A low-noise single-photon detector implemented in a 130nm cmos imaging process," *Solid-State Electronics*, vol. 53, no. 7, pp. 803–808, 2009, Papers Selected from the 38th European Solid-State Device Research Conference – ESSDERC'08. DOI: https://doi.org/10.1016/j.sse.2009.02.014.

[37] R. M. Field, J. Lary, J. Cohn, L. Paninski, and K. L. Shepard, "A low-noise, single-photon avalanche diode in standard 0.13m complementary metal-oxide-semiconductor process," *Applied Physics Letters*, vol. 97, no. 21, p. 211 111, 2010. DOI: 10.1063/1.3518473.

[38] J. Richardson, L. Grant, and R. Henderson, "Low dark count single-photon avalanche diode structure compatible with standard nanometer scale cmos technology," *Photonics Technology Letters, IEEE*, vol. 21, pp. 1020–1022, Aug. 2009. DOI: 10.1109/LPT.2009.2022059.

[39] L. Pancheri and D. Stoppa, "Low-noise single photon avalanche diodes in 0.15 m cmos technology," *2011 Proceedings of the European Solid-State Device Research Conference (ESSDERC)*, 2011, pp. 179–182. DOI: 10.1109/ESSDERC.2011.6044205.

[40] 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, 2012. DOI: 10.1109/LED.2012.2187420.

[41] 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, p. 459, Mar. 2016. DOI: 10.3390/s16040459.

[42] 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, 2012. DOI: 10.1109/LED.2012.2214760.

[43] Y. Zou, D. Bronzi, F. Villa, and S. Weyers, "Backside illuminated wafer-to-wafer bonding single photon avalanche diode array," *2014 10th Conference on Ph.D. Research in Microelectronics and Electronics (PRIME)*, 2014, pp. 1–4. DOI: 10.1109/PRIME.2014.6872751.

[44] 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*, vol. 23, no. 10, pp. 13 200–13 209, May 2015. DOI: 10.1364/OE.23.013200.

[45] 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, 2019. DOI: 10.1109/JSTQE.2019.2918930.

[46] 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.

[47] P. Martelli, M. Brunero, A. Fasiello, F. Rossi, A. Tosi, and M. Martinelli, "Single-spad implementation of quantum key distribution," *2019 21st International Conference on Transparent Optical Networks (ICTON)*, Jul. 2019, pp. 1–3. DOI: 10.1109/ICTON.2019.8840199.

[48] A. Chiuri and F. Angelini, "Fast gating for raman spectroscopy," *Sensors*, vol. 21, no. 8, 2021. DOI: 10.3390/s21082579.

[49] M. Buttafava, G. Boso, A. Ruggeri, A. Dalla Mora, and A. Tosi, "Time-gated single-photon detection module with 110 ps transition time and up to 80 mhz repetition rate," *Review of Scientific Instruments*, vol. 85, no. 8, p. 083 114, 2014. DOI: 10.1063/1.4893385.

[50] L. H. C. Braga, L. Gasparini, L. Grant, *et al.*, "A fully digital 8×16 sipm array for pet applications with per-pixel tdcs and real-time energy output," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 1, pp. 301–314, 2014. DOI: 10.1109/JSSC.2013.2284351.

[51] D. Bronzi, F. Villa, S. Bellisai, S. Tisa, G. Ripamonti, and A. Tosi, "Figures of merit for CMOS SPADs and arrays," *Photon Counting Applications IV; and Quantum Optics and Quantum Information Transfer and Processing*, J. Fiurásek, I. Prochazka, and R. Sobolewski, Eds., International Society for Optics and Photonics, vol. 8773, SPIE, 2013, pp. 21–27. DOI: 10.1117/12.2017357.

[52] 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.

[53] F. Guerrieri, S. Tisa, A. Tosi, and F. Zappa, "Two-dimensional spad imaging camera for photon counting," *IEEE Photonics Journal*, vol. 2, no. 5, pp. 759–774, 2010. DOI: 10.1109/JPHOT.2010.2066554.

[54] X. Lu, M.-K. Law, Y. Jiang, X. Zhao, P.-I. Mak, and R. P. Martins, "A 4-µm diameter spad using less-doped n-well guard ring in baseline 65-nm cmos," *IEEE Transactions on Electron Devices*, vol. 67, no. 5, pp. 2223–2225, 2020. DOI: 10.1109/TED.2020.2982701.

[55] S. Parent, M. Côté, F. Vachon, *et al.*, "Single photon avalanche diodes and vertical integration process for a 3d digital sipm using industrial semiconductor technologies," *2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)*, 2018, pp. 1–4. DOI: 10.1109/NSSMIC.2018.8824571.

[56] C. Piemonte, F. Acerbi, A. Ferri, *et al.*, "Performance of nuv-hd silicon photomultiplier technology," *IEEE Transactions on Electron Devices*, vol. 63, no. 3, pp. 1111–1116, 2016. DOI: 10.1109/TED.2016.2516641.

[57] F. Villa, D. Bronzi, Y. Zou, *et al.*, "Cmos spads with up to 500 m diameter and 55% detection efficiency at 420 nm," *Journal of Modern Optics*, vol. 61, no. 2, pp. 102–115, 2014. DOI: 10.1080/09500340.2013.864425.

[58] Hamamatsu Photonics, *Hamamatsu mppc s13360 series*, 2016. [Online]. Available: https://www.hamamatsu.com/eu/en/product/optical-sensors/mppc/mppc_mppc-array.html.

[59] Semiconductor Components Industries, LLC. "Onsemi j-series sipm." (2017), [Online]. Available: https://www.onsemi.com/products/sensors/photodetectors-sipm-spad/silicon-photomultipliers-sipm/j-series-sipm.

[60] "Broadcom afbr-s4n66p024m sipm." (2021), [Online]. Available: https://www.broadcom.com/products/optical-sensors/silicon-photomultiplier-sipm/afbr-s4n66p024m.

[61] S. Hasan, S. A. Shawkat, S. Amer, S. Islam, N. McFarlane, and G. Rose, "Modeling emerging semiconductor devices for circuit simulation," *Computational Models in engineering*, K. Volkov, Ed., IntechOpen, London, May 2019. DOI: 10.5772/intechopen.85873.

[62] W. J. Kindt, "Geiger mode avalanche photodiode arrays: For spatially resolved single photon counting," Ph.D. dissertation, Delft University of Technology, 1999. [Online]. Available: https://repository.tudelft.nl/islandora/object/uuid%3A300807f8-97db-4e6b-a77e-a19f45ce6e40.

[63] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*. John Wiley & Sons, Dec. 2006.

[64] J. Singh, *Semiconductor Optoelectronics: Physics and Technology*. McGraw-Hill, 1995.

[65] P. A. Wolff, "Theory of electron multiplication in silicon and germanium," *Phys. Rev.*, vol. 95, pp. 1415–1420, 6 Sep. 1954. DOI: 10.1103/PhysRev.95.1415.

[66] W. Shockley, "Problems related to p-n junctions in silicon," *Solid-State Electronics*, vol. 2, no. 1, pp. 35–67, 1961. DOI: https://doi.org/10.1016/0038-1101(61)90054-5.

[67] G. A. Baraff, "Maximum anisotropy approximation for calculating electron distributions; application to high field transport in semiconductors," *Phys. Rev.*, vol. 133, A26–A33, 1A Jan. 1964. DOI: 10.1103/PhysRev.133.A26.

[68] B. K. Ridley, "Lucky-drift mechanism for impact ionisation in semiconductors," *Journal of Physics C: Solid State Physics*, vol. 16, no. 17, pp. 3373–3388, Jun. 1983. DOI: 10.1088/0022-3719/16/17/020.

[69] A. G. Chynoweth, "Ionization rates for electrons and holes in silicon," *Phys. Rev.*, vol. 109, pp. 1537–1540, 5 Mar. 1958. DOI: 10.1103/PhysRev.109.1537.

[70] R. V. Overstraeten and H. D. Man, "Measurement of the ionization rates in diffused silicon p-n junctions," *Solid-State Electronics*, vol. 13, no. 5, pp. 583–608, 1970. DOI: https://doi.org/10.1016/0038-1101(70)90139-5.

[71] *Atlas user's manual: Device simulation software*, Silvaco, Inc, 2016.

[72] P. Mars, "Temperature dependence of avalanche breakdown voltage temperature dependence of avalanche breakdown voltage in p—n junctions†," *International Journal of Electronics*, vol. 32, no. 1, pp. 23–37, 1972. DOI: 10.1080/00207217208938266.

[73] R. Entner, "Modeling and simulation of negative bias temperature instability," Ph.D. dissertation, Vienna University of Technology, 2007.

[74] S. M. Sze and M.-K. Lee, *Semiconductor Devices: Physics and Technology*. John Wiley & Sons Singapore Pte. Limited, Aug. 2012.

[75] S. Sze and G. Gibbons, "Effect of junction curvature on breakdown voltage in semiconductors," *Solid-State Electronics*, vol. 9, no. 9, pp. 831–845, 1966. DOI: https://doi.org/10.1016/0038-1101(66)90033-5.

[76] D. P. Palubiak and M. J. Deen, "CMOS SPADs: Design Issues and Research Challenges for Detectors, Circuits, and Arrays," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 20, no. 6, pp. 409–426, Nov. 2014. DOI: 10.1109/JSTQE.2014.2344034.

[77] B. Zhang, Z. Li, and M. Zaghloul, "A single-photon avalanche diode in cmos 0.5 µm n-well process," *Proceedings of IEEE Sensors*, Oct. 2012, pp. 1–4. DOI: 10.1109/ICSENS.2012.6411333.

[78] P. Yu and M. Cardona, *Fundamentals of Semiconductors*. 2010. [Online]. Available: https://link.springer.com/book/10.1007/978-3-642-00710-1.

[79] A. Yariv, "The application of time evolution operators and feynman diagrams to nonlinear optics," *IEEE Journal of Quantum Electronics*, vol. 13, no. 12, pp. 943–950, 1977. DOI: 10.1109/JQE.1977.1069267.

[80] A. Yariv, *Quantum electronics*, 3rd ed. Wiley, 1991.

[81] B. K. Ridley, *Quantum processes in semiconductors*, 3nd. Oxford: Clarendon Press, 1993.

[82] C.-Y. Tsai, "Absorption coefficients of silicon: A theoretical treatment," *Journal of Applied Physics*, vol. 123, no. 18, p. 183 103, 2018. DOI: 10.1063/1.5028053.

[83] M. A. Green, "Self-consistent optical parameters of intrinsic silicon at 300 k including temperature coefficients," *Solar Energy Materials and Solar Cells*, vol. 92, 1305–1310, 2008. DOI: 10.1016/j.solmat.2008.06.009.

[84] 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.*, vol. 60, no. 25, pp. 7815–7820, Sep. 2021. DOI: 10.1364/AO.432219.

[85] M. Mazzillo, A. Piazza, G. Condorelli, *et al.*, "Quantum detection efficiency in geiger mode avalanche photodiodes," *IEEE Transactions on Nuclear Science*, vol. 55, no. 6, pp. 3620–3625, Dec. 2008. DOI: 10.1109/TNS.2008.2007483.

[86] W. Oldham, 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.

[87] R. 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, 1973. DOI: 10.1109/T-ED.1973.17715.

[88] Y. Xu, P. Xiang, and X. Xie, "Comprehensive understanding of dark count mechanisms of single-photon avalanche diodes fabricated in deep sub-micron cmos technologies," *Solid-State Electronics*, vol. 129, pp. 168–174, 2017. DOI: https://doi.org/10.1016/j.sse.2016.11.009.

[89] E. O. Kane, "Theory of tunneling," *Journal of Applied Physics*, vol. 32, no. 1, pp. 83–91, 1961. DOI: 10.1063/1.1735965. eprint: https://doi.org/10.1063/1.1735965. [Online]. Available: https://doi.org/10.1063/1.1735965.

[90] S. K. Saha, *FinFET Devices for VLSI Circuits and Systems*. CRC Press, 2020.

[91] G. Hurkx, D. Klaassen, and M. Knuvers, "A new recombination model for device simulation including tunneling," *IEEE Transactions on Electron Devices*, vol. 39, no. 2, pp. 331–338, 1992. DOI: 10.1109/16.121690.

[92] C. Piemonte and A. Gola, "Overview on the main parameters and technology of modern silicon photomultipliers," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 926, pp. 2–15, 2019. DOI: https://doi.org/10.1016/j.nima.2018.11.119.

[93] W. Kindt and H. W. Van Zeijl, "Modelling and fabrication of Geiger mode avalanche photodiodes," *IEEE Transactions on Nuclear Science*, vol. 45, no. 3, pp. 715–719, Jun. 1998. DOI: 10.1109/23.682621.

[94] 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, 1991. DOI: 10.1109/55.116955.

[95] H. Mahmoudi, M. Hofbauer, B. Goll, and H. Zimmermann, "Noise and breakdown characterization of spad detectors with time-gated photon-counting operation," *Sensors*, vol. 21, no. 16, 2021. DOI: 10.3390/s21165287.

[96] O. Marinov, M. J. Deen, and J. A. Jimenez Tejada, "Theory of microplasma fluctuations and noise in silicon diode in avalanche breakdown," *Journal of Applied Physics*, vol. 101, no. 6, p. 064 515, 2007. DOI: 10.1063/1.2654973.

[97] L. Neri, S. Tudisco, F. Musumeci, *et al.*, "Note: Dead time causes and correction method for single photon avalanche diode devices," *Review of Scientific Instruments*, vol. 81, no. 8, p. 086 102, 2010. DOI: 10.1063/1.3476317.

[98] I. S. Alirezaei, N. Andre, and D. Flandre, "Enhanced ultraviolet avalanche photodiode with 640-nm-thin silicon body based on soi technology," *IEEE Transactions on Electron Devices*, vol. 67, no. 11, pp. 4641–4644, 2020. DOI: 10.1109/TED.2020.3017699.

[99] I. Sabri Alirezaei, N. Andre, A. Sedki, P. Gerard, and D. Flandre, "An ultra-thin ultraviolet enhanced backside-illuminated single-photon avalanche diode with 650 nm-thin silicon body based on soi technology," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 28, no. 2, pp. 1–10, 2022. DOI: 10.1109/JSTQE.2021.3129274.

[100] *Athena user's manual: 2d process simulation software*, Silvaco, Inc, 2005.

[101] W. Van Roosbroeck, "Theory of the flow of electrons and holes in germanium and other semiconductors," *Bell System Technical Journal*, vol. 29, no. 4, pp. 560–607, 1950. DOI: https://doi.org/10.1002/j.1538-7305.1950.tb03653.x.

[102] B. Tabbert and A. Goushcha, "Optical detectors," *Springer Handbook of Lasers and Optics*. Springer Berlin, Heidelberg, Jan. 2012, pp. 543–619. DOI: 10.1007/978-3-642-19409-2_9.

[103] M. Tyagi and R. Van Overstraeten, "Minority carrier recombination in heavily-doped silicon," *Solid-State Electronics*, vol. 26, no. 6, pp. 577–597, 1983. DOI: https://doi.org/10.1016/0038-1101(83)90174-0.

[104] "Ultrasil llc." (2022), [Online]. Available: https://www.ultrasil.com/get_quote_all_products_new.aspx.

[105] "Produits-soitec." (2017), [Online]. Available: https://www.soitec.com/fr/produits.

[106] Keysight, *E4727b advanced low-frequency noise analyzer: Datasheet*, 2020. [Online]. Available: E4727B%20Advanced%20Low-Frequency%20Noise%20Analyzer (visited on 06/18/2022).

[107] B. Pires, A. Silva, A. Moskaltsova, *et al.*, "Multilevel process on large area wafers for nanoscale devices," *Journal of Manufacturing Processes*, vol. 32, Feb. 2018. DOI: 10.1016/j.jmapro.2018.01.024.

[108] ESPEC Corp., *Bench-top type temperature (humidity) chambers*, 2021. [Online]. Available: https://www.espec.co.jp/english/products/book/sh/#target/page_no=1 (visited on 08/10/2022).

[109] A. Inc, *A225 charge sensitive preamplifier shaping amplifier*. [Online]. Available: https://www.amptek.com/internal-products/a225-charge-sensitive-preamplifier-shaping-amplifier (visited on 08/10/2022).

[110] P. Virtanen, R. Gommers, T. E. Oliphant, *et al.*, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020. DOI: 10.1038/s41592-019-0686-2.

[111] S. Byun, *Med phys 4ra3,4rb3,6r03 course notes: Chapter 6 pulse processing*, Sep. 2021.

[112] K. Siegrist, *Virtual laboratories in probability and statistics, chapter 13: The poisson process*, 2009. [Online]. Available: http://math.bme.hu/~nandori/Virtual_lab/stat/poisson/Gamma.xhtml.

[113] C. Piemonte, A. Ferri, A. Gola, *et al.*, "Development of an automatic procedure for the characterization of silicon photomultipliers," *2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC)*, 2012, pp. 428–432. DOI: 10.1109/NSSMIC.2012.6551141.

[114] K. Technologies, *B1530a waveform generator/fast measurement unit: User's guide*. [Online]. Available: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjS7K2389H5AhVPr6QKHflxCo0QFnoECAoQAQ&url=https%5C%3A%5C%2F%5C%2Fwww.keysight.com%5C%2Fus%5C%2Fen%5C%2Fassets%5C%2F9018-01706%5C%2Fuser-manuals%5C%2F9018-01706.pdf&usg=AOvVaw0rktJnTx-Y1BY1tzsDqMwR (visited on 08/10/2022).

[115] D. Phillips, R. Drake, D. OConnor, and R. Christensen, "Time correlated single-photon counting (tcspc) using laser excitation," *Instrumentation Science Technology*, vol. 14, pp. 267–292, Sep. 1985. DOI: 10.1080/10739148508543581.

[116] "Fully digital arrays of silicon photomultipliers (dsipm) – a scalable alternative to vacuum photomultiplier tubes (pmt)," *Physics Procedia*, vol. 37, pp. 1546–1560, 2012, ISSN: 1875-3892. DOI: https://doi.org/10.1016/j.phpro.2012.03.749.

[117] S. Hatefi Hesari, M. A. Haque, and N. McFarlane, "A comprehensive survey of readout strategies for sipms used in nuclear imaging systems," *Photonics*, vol. 8, no. 7, 2021. DOI: 10.3390/photonics8070266.

[118] T. Szcze˛śniak, M. Grodzicka, M. Moszyński, M. Szawłowski, D. Wolski, and J. Baszak, "Characteristics of scintillation detectors based on inorganic scintillators and sipm light readout," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 702, pp. 91–93, 2013, PET/MR and SPECT/MR: New Paradigms for Combined Modalities in Molecular Imaging, ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2012.09.011.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
École polytechnique de Louvain
Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/epl**