Photon Counting

Module 5

SensL Technologies Ltd.
River View Business Park
Blackrock
Cork, Ireland

www.SensL.com

Contents

1 Objectives

This module explains the operation and structure of a photon counting detector. Photon counting detector characteristics which determine the performance of the device are also described. After reading this module the reader will:

- Understand the basic operation of a photon counting detector.
- Appreciate the characteristics which indicate the quality of the detector.

---

2 Photon Counting Detector Operation

Photon counting is the technique of choice when measuring low level optical signals. The intensity of the signal is measured by counting the number of detected photons during the measurement time (photon counting). In the past photomultiplier tubes (PMTs) were used to detect ultraweak optical signals, but they are bulky, delicate and require very high voltages (> 1 kV) [1]. Avalanche photodiodes (APDs) biased slightly below the breakdown voltage in the avalanche mode may be used for single photon detection. Avalanche photodiodes offer advantages such as smaller size and lower bias voltages. However, biased in avalanche mode APDs have limited gain which is affected by strong statistical fluctuations. Single photon detection is most efficiently obtained by biasing APDs above the breakdown voltage in Geiger-mode. APDs biased above breakdown are often referred to as Geiger-mode APDs (GM-APDs) or single photon avalanche photodiodes (SPADs). To date GM-APDs have been used successfully in applications such as flourescence microscopy for single molecule detection, quantum cryptography, range finding and 3-D imaging.

There are two basic silicon photon counting detector structures. The first, shown in Fig. ??(a), is as a shallow junction device while the second, shown in Fig. ??(b), is a reach through detector. The electric fields for both devices, when used as photon counters, are also shown in Fig. ??. The critical multiplication

field (shown by the line drawn through the electric field profiles) is the level the electric field must reach in order for photoelectron multiplication (impact ionisation) to occur. Below the critical field no multiplication occurs in the device.

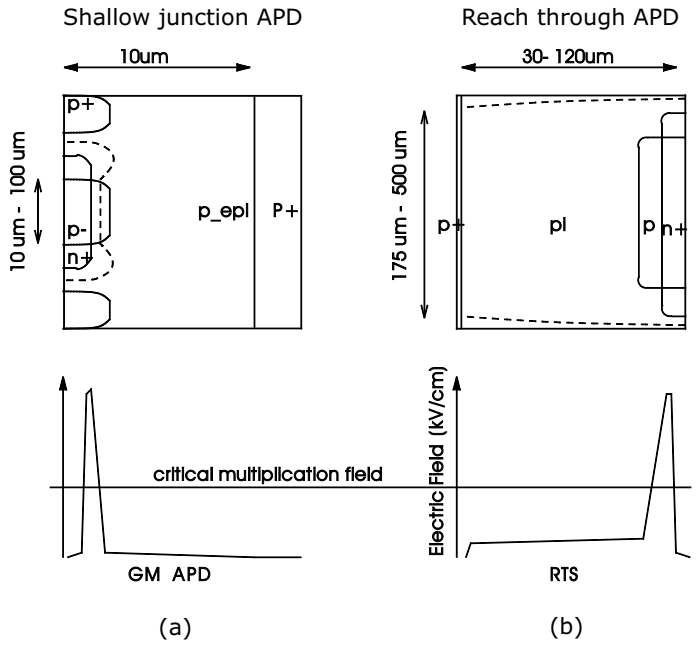

Figure 1: Photon counting detector structures, (a) shallow junction structure and (b) reach through structure.

The shallow junction structure consists of n^+ and p type diffusions into a p^- epilayer. The junction is located $1-2$ μm from the surface of the wafer, giving the detector increased sensitivity to short wavelengths $(400-650$ nm) which have short penetration depths in silicon (< 2.5 μm). The central p type implant, termed the high field region, defines the active area of the detector. This active area contains the high field depletion region where the multiplication process takes place. Edge breakdown is avoided by extending the n^+ region beyond the p type region, forming what is known as a guard ring around the active area. The guard ring has a higher breakdown voltage than the active area since the breakdown voltages are inversely proportional to the doping concentration of the p^- and p type regions respectively.

The reach-through structure (RTS) seeks to take advantage of the spectral

range of silicon. Because silicon has a bandgap energy of 1.12 eV, only photons of light with energies greater than 1.12 eV will be detected. Using the relation:

$$h\nu > E_g \quad or \quad \frac{hc}{\lambda} > E_g \tag{1}$$

where h is Planck's constant, E_g is the bandgap energy, c is the speed of light and λ is the wavelength of the light, it is found that silicon will not detect light with a wavelength of greater than 1100 nm. Furthermore, wavelengths shorter than 400 nm are absorbed very near the surface of the wafer where the carriers recombine before being detected. Silicon is therefore only responsive from 400 nm to 1100 nm. However, longer wavelengths of light require increased silicon thicknesses for the detection of photons. The reach through detector uses the $n^+ - p$ region as the multiplication region and a thick (30 μm to 120 μm) low doped p^- region on top to collect photons and encourage them to diffuse to the multiplication region with a high probability of initiating a breakdown.

3 Photon Counting Detector Characteristics

The quality of a single photon counting detector and its suitability for a particular application depends on several operating parameters of the device. Some of the more important parameters are described here.

3.1 Current-Voltage Characteristics

A GM-APD with a shallow junction device structure similar to that described above will typically have a primary breakdown voltage $\approx 30\ V$. Fig. 2 shows a simulation of the reverse current-voltage characteristic of one of these devices. From this image there are two observable break-down events, the first occurring around $30\ V$ reverse bias due to the active junction breakdown. The secondary breakdown occurs closer to $60\ V$ reverse bias and is due to the virtual guard ring that ensures uniform breakdown of the active junction.

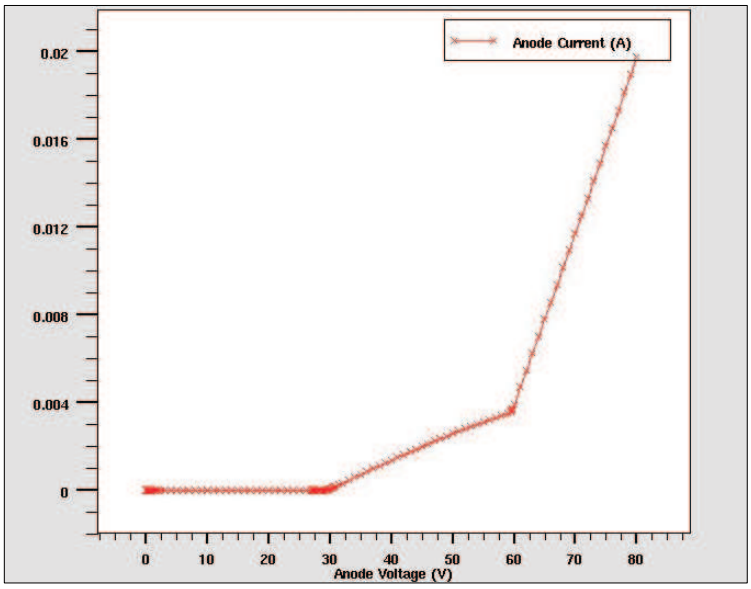

Figure 2: Simulation showing the reverse current-voltage characteristic of the device.

3.2 Dark Count Rate

Ideally a photon counting detector will have no dark count. The dark count is due to avalanche events triggered by thermally generated carriers. The dark count rate is a good measure of the quality of the device material, as it is also related to the quantity of generation centres present in the device. The dark count rate (DCR) increases as the active area of the device is increased and it can be reduced by cooling the device. Fig. 3 shows representative data for the dark count rate observed in 10 μm, 15 μm and 20 μm diameter devices. It is observed here that the dark count rate also increases with the excess bias applied above the breakdown voltage.

3.3 Photon Detection Probability

The photon detection probability (PDP) is the product of the device quantum efficiency and the avalanche initiation probability. It is therefore wavelength dependent. The PDP is the ultimate sensitivity of the device to photons of a specific

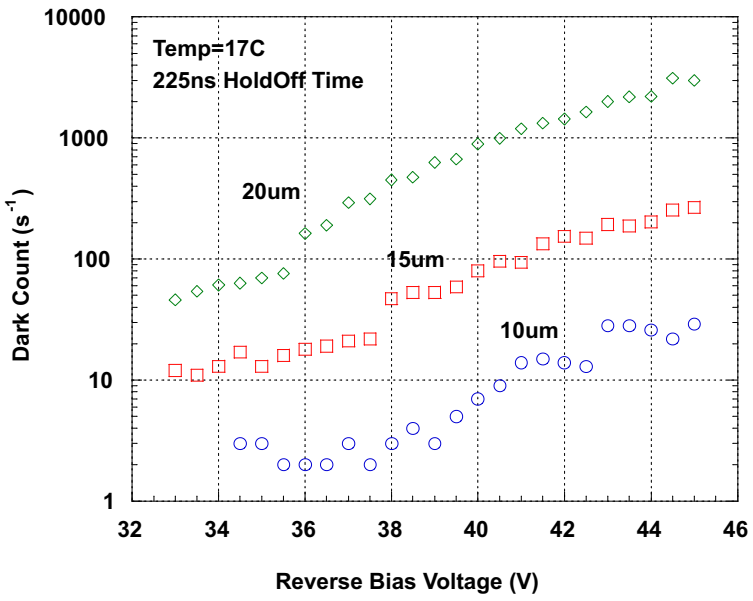

Figure 3: Dark count rate as a function of applied reverse bias for various sized detectors.

energy. Fig. 4 shows how the PDP for a silicon GM-APD typically varies with wavelength, while Fig. 5 shows that the PDP increases with increasing excess bias above the breakdown voltage.

To increase the PDP one needs to increase the quantum efficiency (by adding an anti-reflection coating, for example) or increase the avalanche initiation probability (by ensuring the peak electric field in the device is as close as possible to the critical breakdown field).

3.4 Photon Timing

Photon counting detectors are widely employed for the measurement of very faint optical signals. Biased in Geiger-mode, a photon counting detector can remain in a zero current state until a carrier succeeds in triggering an avalanche. When the carrier is photogenerated the leading edge of the avalanche current marks the arrival of a photon. The avalanche current flows until an external circuit lowers the bias below the breakdown voltage. The bias must then be restored allowing

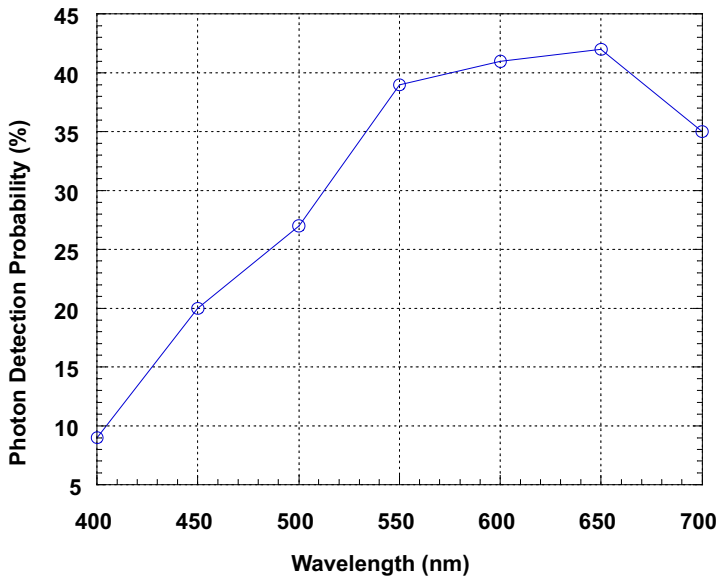

Figure 4: The photon detection probability for a silicon GM-APD as a function of wavelength.

the diode to detect another photon. The intensity of slowly varying signals is measured by counting the number of photons detected in a particular time, this is known as photon counting [2]. The intensity of fast optical signals is found by collecting a histogram of the arrival time of single photons, in a time correlated photon counting setup, this is termed photon timing [2].

In photon timing, the fastest signal detectable is limited by the timing performance of the photon counting detector itself. The timing response or resolution of a detector can be found by obtaining the statistical distribution of the delays from the true arrival of a photon at the device to the actual detection time. Fig. 6 shows a typical setup that would be used to find the timing resolution of a photodiode [3]. Picosecond laser pulses are sent directly to the detector and a time to pulse height converter (TPHC) measures the time between the emission of the photon, which is marked by a trigger signal synchronous with the pulse, and the photon detection which is defined by the avalanche current. A multichannel analyser (MCA) collects the results of repetitive measurements and the resulting histogram gives the timing resolution of the diode.

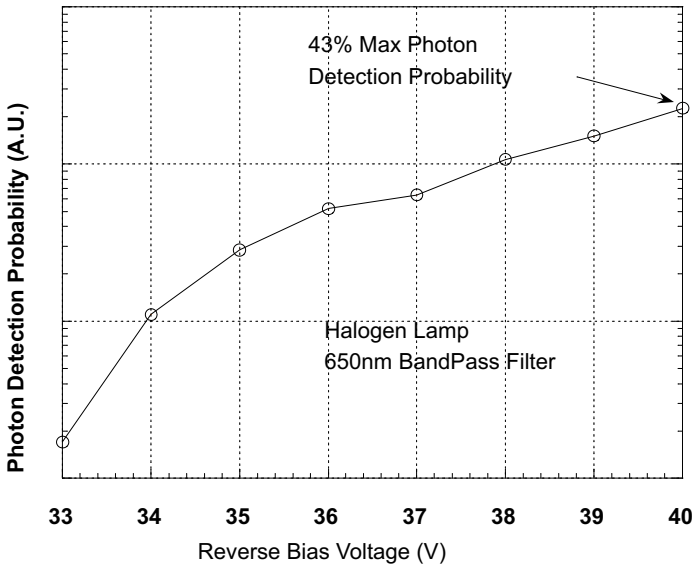

Figure 5: The photon detection probability for a silicon GM-APD as a function of bias voltage.

The timing resolution of shallow junction photon counting detectors is characterised by two components, namely a peak and a slow tail. The peak is due to carriers generated in the depletion region which experience high electric fields and are quickly detected. The timing resolution of the detector is given as the full width at half maximum of this peak [4]. In general, the thicker the depletion region the poorer the timing resolution [5]. The tail in the timing response is due to minority carriers, photogenerated in the bulk region beneath the junction that diffuse to the depletion layer. It has also been shown that the timing resolution achievable is dependent on the diameter or area of the detector [6]. This is due to the time taken for the multiplication process to spread from the point where the photon is absorbed to the whole of the active area. For optimum timing performance the sensitive area of the diodes must therefore be as small as possible. This, however, leads to low photodiode volume resulting in poor quantum efficiency and there is thus a trade-off between timing response and quantum efficiency.

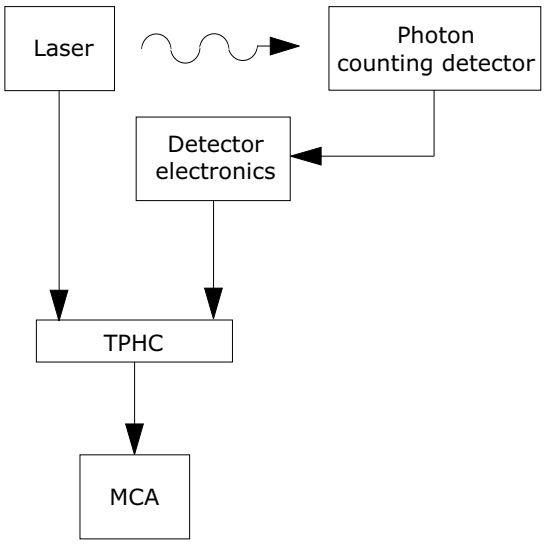

Figure 6: Experimental setup for the timing response measurement of a photon counting detector.

3.5 Dead Time

The process of detecting the onset of an avalanche event, registering a count, quenching the avalanche current, holding the device off for a preset time to allow trapped charge to be released and the subsequent resetting of the device above breakdown ensures that there is a relatively long period of time for which the device is inactive. This period of inactivity is known as the dead time of the device. The main effect of the dead time is the loss of signal while the device is off. The measured count rate is thus not a true reflection of the statistically true count rate. It is possible to correct the measured count rate by taking the dead time into consideration to arrive at the true signal rate.

3.6 Afterpulsing

A Geiger-mode avalanche photodiode relies on producing an avalanche current for every detected photon. However, avalanche currents in the Geiger-mode device will also be produced by thermally generated carriers in the active region of the device, leading to the dark-count. Fortunately dark count is easily quantified

and accounted for. However, there is another source of carriers that produce avalanching in the GM-APD and its effect is such that the detector will produce a larger number of counts than the number of real photon events in a specified time interval. This occurs because the avalanche process itself produces a flow of charge, some of which is trapped by trapping centres in the device. These traps will hold the charge for a time equal to the mean lifetime of the trap after which the charge will be released. If the device is biased above breakdown at this time the released charge can lead to the creation of an avalanche current, but this current is not related to the arrival of a new photon, instead it is correlated to a previous avalanche event and thus the count registered is a false count. These false counts are termed afterpulses and it is desirable that they are avoided in photon counting.

In the normal process of photon counting an active quench circuit is employed to sense the onset of the avalanche current, bring the device below breakdown to quench the avalanche current, keep the device below breakdown for a time (called the *hold-off time*) that exceeds the mean trap lifetime (thus avoiding afterpulses) and finally reset the device back above breakdown to wait for the next photon-arrival. The total time involved in this process is called the *dead time* of the device, as no further photons can be detected during this time. The quality of the GM-APD's output depends on the amount of afterpulsing present. This is relatively easy to analyse.

Assume that at time t_0 an avalanche is triggered. While the corresponding current pulse flows a certain number of electrons, say M, are trapped in *defect centres* or *trapping centres* in the device. Every trapped carrier is released after a statistical time delay. If τ_j is the delay after which the j^{th} electron is released, then the j^{th} electron is freed after a time $t_j = t_0 + \tau_j$. The last trapped electron is released a time τ_M after t_0. If τ_M is shorter than the dead time of the device then no afterpulsing will be observed. Hence, it is desirable that the hold-off time is set such that the total dead time is greater than the maximum time delay before the last trapped electron is released. Unfortunately, both the number of trapped electrons and their release-times are not known *a priori* as they are

statistical. Therefore, the statistical distribution of the delay times of the released electrons has to be considered. After the dead time has expired there is a non-zero probability that a trapped electron will be released. In other words, it is theoretically possible that a trapped electron can be released at any time, so it appears impossible to totally prevent afterpulsing from occuring.

The afterpulsing statistics are quite complex. The ultimate afterpulsing probability depends mainly on two statistical events: the number of trapped charges during an avalanche and the release time of these charges relative to the hold-off time set. These two events can be studied jointly as the probability of having a certain number of still unreleased electrons after the dead time has expired. It is important to realise that some electrons can be trapped and released at the edge of the gain region. These electrons may exit this region without first triggering an avalanche event, thus lowering the afterpulsing probability.

Let m_r be the mean release time of the traps and σ_r the standard deviation. Considering first a dead time less than the mean release time, $t_{dead} < m_r$. In this case there is a high probability that afterpulsing will occur. Before the dead time expires, a few of the trapped electrons will be released. Thus it is extremely likely that an afterpulse will be observed after the dead time interval, since most of the trapped electrons have yet to be freed. Since most of the trapped electrons are released after the dead time has expired the next electron released after the dead time triggers a new avalanche, causing an afterpulse. This new pulse exhibits the same behaviour as the pulse generated by the primary avalanche, generating additional afterpulses and the process repeats.

It can be noticed that when the dead time is less than the average release time of the trapped electrons, the detector may present a train of pulses as output even in the presence of only one photon event. Every pulse is separated from the previous by a time quantity equal to $t_{dead} + \delta$ where δ depends on the statistics of the electron which generates the afterpulse. It can be concluded that in this particular situation the dead time is too small as it does not prevent afterpulsing from occuring.

When the hold-off time is set to make the dead time much larger than the

average trap release time, the probability of afterpulsing occurring is lower. That is, it's very likely that all of the trapped electrons are released within the dead time interval. This does not guarantee that afterpulsing will not occur as there is still a non-zero probability (even if low) that some electron will be released after the dead time has expired. The longer the hold-off time used the lower is the probability of observing afterpulses, but the achievable photon counting rate is also reduced by the increased dead time. Therefore, a trade-off exists since the hold-off time has to be chosen to ensure a good photon detection rate while maintaining a low probability of afterpulsing.

3.6.1 Afterpulsing Statistics

As noted previously, electrons can get trapped in different energy levels in the structure of the semiconductor. The statistical description of afterpulsing starts from knowledge of the release time statistics. It has been observed that trapped carriers are subsequently released with exponentially decaying probability from their traps. Every trap centre contributes a different exponential probability density to global release statistics. That is, every trap centre is characterised by its own time constant. The time constant of the j^{th} trap level is given by:

$$\tau_j = \tau_0 e^{\frac{E_j}{k_B T}} \tag{2}$$

where E_j is the activation energy of level j, k_B is Boltzmann's constant, and T is the absolute temperature. The scaling value τ_0 is the time constant of a trap level having a zero activation energy. It is therefore a minimum value above which all time constants lie. At low temperatures all traps exhibit a larger time constant, thus the releasing process is slower and afterpulsing is more likely. The GM-APDs can operate at moderately high temperatures but this increases the thermal generation rate. Hence, a compromise in temperature must also be determined.

The afterpulsing probability density function (pdf) over a window which starts

from the end of the dead time after an avalanche is given by:

$$f_a(t) = B + \sum_{j=1}^{L} A_j e^{-\frac{t}{\tau_j}} \tag{3}$$

where L is the number of trap levels, τ_j is the time constant for the j^{th} trap centre, B is the primary count rate and A_j is the initial value (at the end of the dead time interval) of the amplitude of the j^{th} exponential component. The total afterpulsing probability is:

$$P_a^{TOT} = \sum_{j=1}^{L} A_j \tau_j \tag{4}$$

where each term of the sum is the afterpulsing probability for the j^{th} level since it represents the area of the corresponding exponential. Since every trap level operates independently from every other level the global probability density function is determined by the sum of the pdfs for every trap level. The probability density function of level j versus t means that, if the probability that afterpulsing occurs before time t has to be considered, the integral of the j^{th}-level pdf has to be calculated in the time interval starting from the end of the dead time up to t. Hence, probability that afterpulsing occurs due to level j at some time after the dead time has expired is the area of the corresponding exponential starting from the end of the dead time, as shown in Fig. 7.

By calculating the corresponding integral it can be found that this area is equal to $A_j \tau_j$. Therefore, by summing over all the L trap levels the afterpulsing probability is given as shown in Eq. 4. In order to obtain a value for P_a^{TOT} the amplitude values and the time constants have to be determined. This can be done experimentally employing several techniques, such as time-correlated carrier counting (TCCC). This experimental measurement uses thermally generated carriers as the primary carriers. The constant value B is measured separately and then subtracted. It is measured using a long dead time to ensure that afterpulsing is negligible. Once these values have been determined, the afterpulsing statistics are known. Once the global pdf has been determined, the mean value and standard deviation of the release time can be easily calculated. The dead time can be extended until a desired afterpulsing probability is obtained.

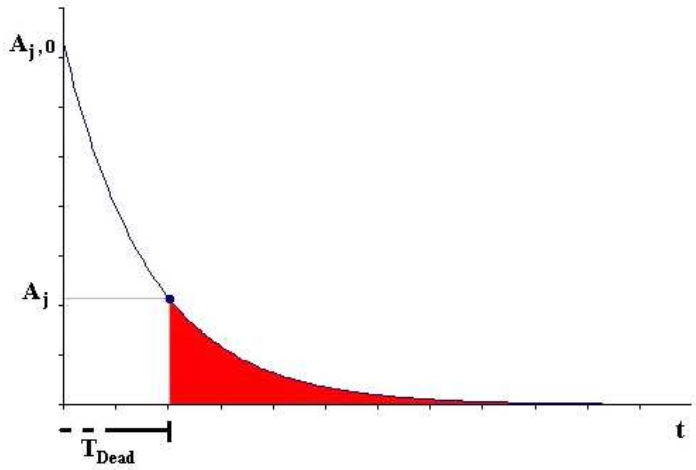

Figure 7: Afterpulsing probability due to some level j after some dead time, t_{dead}.

4 Summary

The operation of a photon counting detector has been explained. Items such as dark count, photon detection probability and afterpulsing, which demonstrate the detectors efficiency as a photon counter, have also been detailed. The reader is now advised to perform the dark count rate section of Experiment 2 together with Experiment 3.

5 Acknowledgements

SensL Technologies Ltd. would like to acknowledge Dr. Aoife M. Moloney of the School of Electronic and Communications Engineering in the Dublin Institute of

Technology and Dr. Alan P. Morrison of the Department of Electrical and Electronic Engineering in University College Cork, for writing and coordinating the PCEdu-1 Photon Counting Fundamentals Pack.

References

[1] A. Rochas, G. Ribordy, B. Furrer, P.A. Besse, and R.S. Popovic. First Passively-Quenched Single Photon Counting Avalanche Photodiode Element Integrated in a Conventional CMOS Process with 32 ns Dead Time. *Proceedings of the 5^{th} International Conference on Application of Photonic Technology, ICAPT 2002*, 2002.

[2] F. Zappa, M. Ghioni, S. Cova, C. Samori, and A.C. Giudice. An Integrated Active Quenching Circuit for Single Photon Avalanche Diodes. *IEEE Trans. on Instrumentation and Measurement*, 49(6):1167–1175, December 2000.

[3] A. Spinelli and A.L. Lacaita. Physics and Numerical Simulation of Single Photon Avalanche Diodes. *IEEE Trans. on Electron Devices*, 44(11):1931–1943, November 1997.

[4] A. Lacaita, M. Ghioni, and S. Cova. Double Epitaxy Improves Single Photon Avalanche Diode Performance. *Electron. Letters*, 25(25):841–843, June 1989.

[5] 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*, 14(7):360–362, July 1993.

[6] A. Lacaita and M. Mastrapasqua. Strong Dependence of Time Resolution on Detector Diameter in Single Photon Avalanche Diodes. *Electron. Letters*, 26(24):2053–2054, November 1990.