Design and characterization of a p+/n-well SPAD array in 150nm CMOS process

HESONG XU, 1,2,* LUCIO PANCHERI, 2 GIAN-FRANCO DALLA BETTA, 2 AND DAVID STOPPA 1,2

1Division of Integrated Radiation and Image Sensors, Fondazione Bruno Kessler, Via Sommarive 18, Trento, 38123, Italy
2Department of Industrial Engineering, University of Trento, Via Sommarive 9, Trento, 38123, Italy
*xuhesong@fbk.eu

Abstract: This paper reports on characterization results of a single-photon avalanche diode (SPAD) array in standard CMOS 150nm technology. The array is composed by 25 (5 × 5) SPADs, based on p’/n-well active junction along with a retrograde deep n-well guard ring. The square-shaped SPAD has a 10µm active diameter and 15.6µm pitch size, achieving a 39.9% array fill factor. Characterization results show a good breakdown voltage uniformity (40mV max-min) within each chip and 17mV/°C temperature coefficient. The median DCR is 0.4Hz/µm², and the afterpulsing probability is 0.85% for a dead time of 150ns at 3V excess bias voltage. The peak PDP is 31% at 450nm wavelength and a good uniformity (1.1% standard deviation) is observed for the array at 5V excess bias. The single SPADs exhibit a timing jitter of 52ps (FWHM) and 42ps (FWHM) under a 468-nm and a 831-nm laser, respectively. The crosstalk probability as a function of pixel-to-pixel distance and excess bias voltage is presented, and random telegraph signal (RTS) noise is also discussed in detail.

© 2017 Optical Society of America

OCIS codes: (040.1345) Avalanche photodiodes (APDs); (040.5160) Photodetectors; (040.3780) Low light level; (280.4788) Optical sensing and sensors; (130.5990) Semiconductors; (130.6010) Sensors.

References and links
1. L. H. C. Braga, L. Gasparini, L. Grant, R. K. Henderson, N. Massari, M. Perenzoni, D. Stoppa, and R. Walker, “A fully digital 8×16 SiPM array for PET applications with per-pixel TDCs and real-time energy output,” IEEE J. Solid-State Circuits 49(1), 301–314 (2014).
2. N. Krstajic, J. Levitt, S. Poland, S. Armeen-Beg, and R. Henderson, “256 × 2 SPAD line sensor for time resolved fluorescence spectroscopy,” Opt. Express 23(5), 5653–5669 (2015).
3. C. Niclass, M. Soga, H. Matsubara, S. Kato, and M. Kagami, “A 100-m Range 10-Frame/340 × 96-Pixel Time-of-Flight Depth Sensor in 0.18-µm CMOS,” IEEE J. Solid-State Circuits 48(2), 559–572 (2013).
4. A. Rochas, A. R. Pauchard, P.-A. Besse, D. Pantic, Z. Prijic, and R. S. Popovic, “Low-noise silicon avalanche photodiodes fabricated in conventional CMOS technologies,” IEEE Trans. Electron Dev. 49(3), 387–394 (2002).
5. M. A. Karami, H. J. Yoon, and E. Charbon, “Single-photon Avalanche Diodes in sub-100nm Standard CMOS Technologies,” in Proc. Intl. Image Sens. Workshop IISW (2011).
6. J. A. Richardson, E. A. G. Webster, L. A. Grant, and R. K. Henderson, “Scaleable Single-Photon Avalanche Diode Structures in Nanometer CMOS Technology,” IEEE Trans. Electron Dev. 58(7), 2028–2035 (2011).
7. L. Pancheri, G.-F. Dalla Betta, L. H. Campos Braga, H. Xu, and D. Stoppa, “A single-photon avalanche diode test chip in 150nm CMOS technology,” in 2014 International Conference on Microelectronic Test Structures (ICMTS) (2014), pp. 161–164.
8. S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa, “Avalanche photodiodes and quenching circuits for single-photon detection,” Appl. Opt. 35(12), 1956–1976 (1996).
9. H. Finkelstein, M. J. Hsu, and S. C. Esener, “STI-Bounded Single-Photon Avalanche Diode in a Deep-Submicrometer CMOS Technology,” IEEE Electron Device Lett. 27(11), 887–889 (2006).
10. R. K. Henderson, L. A. Grant, and J. A. Richardson, “Reduction of Band-to-band Tunneling in Deep-submicron CMOS Single Photon Avalanche Photodiodes,” in Proc. IEEE Int. Image Sens. Workshop (2009), pp. 1–4.
11. E. A. G. Webster and R. K. Henderson, “A TCAD and spectroscopy study of dark count mechanisms in single-photon avalanche diodes,” IEEE Trans. Electron Dev. 60(12), 4014–4019 (2013).
12. N. V. Loukianova, H. O. Folkerts, J. P. V. Maa, D. W. E. Verbagt, A. J. Mierop, W. Hoekstra, E. Roks, and A. J. P. Theuwissen, “Leakage current modeling of test structures for characterization of dark current in CMOS image sensors,” IEEE Trans. Electron Dev. 50(1), 77–83 (2003).
1. Introduction

Single-photon avalanche diodes (SPADs) integrated with readout electronics in CMOS processes have been proposed for a wide range of applications because of their cost effectiveness and high functionality. Examples of such applications include positron emission tomography (PET), time-resolved fluorescence spectroscopy and time-of-flight (TOF) range finding [1–3]. Very different SPAD arrays have been fabricated with dedicated circuitry, tailored for the specific system requirements. In most of the cases, high fill-factor and low noise are essential requirements for the array design.

After the first demonstration of SPADs in a CMOS process [4], a series of works has been conducted to push the fabrication process node into deep submicron (DSM) technology, down to 65nm [5]. However, this approach potentially increases the noise of SPADs because of high electrical field, and thus limiting the dynamic range for photon detection applications. To alleviate this issue, small-area SPADs were fabricated for low noise, since there is less probability of having defects in the active area [6]. On the other hand, small SPADs exhibit lower pixel fill-factor and detection efficiency can be dramatically reduced. We have demonstrated that the normalized photon detection probability of devices with 5µm nominal active area is reduced to 20% with respect to 20-µm devices [7]. Therefore, a thorough design and characterization of SPAD arrays featuring high fill factor and low noise are essential to develop application-specific single photon imagers.

In several application, where a large pixel size can be accepted, it is sometimes convenient to employ a small array of SPADs instead of a single SPAD in each pixel. In TOF ranging and scintillator readout, for example, it is possible to threshold on the signal from two or more cells to greatly reduce the effect of Dark Count Rate and ambient on the measurement [1,3]. In these cases, optical cross talk can become a non-negligible source of correlated noise.

In this paper, we present a SPAD array fabricated in standard CMOS 150nm process with a complete characterization of the device in terms of noise, uniformity, efficiency and timing resolution. In Section 2, the design of the SPAD array and the readout circuits are presented. In section 3, we focus on the full characterization of performance in terms of breakdown
voltage ($V_{\text{BD}}$), dark count rate (DCR), afterpulsing, photon detection probability (PDP),
timing jitter, crosstalk and random telegraph signal (RTS) noise. The conclusions are
drawn in Section 4.

2. SPAD array design

This section presents the SPAD structure, readout circuits and array design.

2.1 SPAD structure and array design

The essential elements of a SPAD are a p-n junction, working as active region, and a guard
ring, preventing early periphery breakdown [8]. Figure 1 shows the cross section of a SPAD,
implemented in standard 150nm CMOS process, on which we have based our design. The
15.6-µm SPAD was realized in a square shape formed by a p⁺/n-well junction with an active
diameter of 10µm [7]. Shallow trench isolation (STI), which is a feature of deep submicron
processes, was first introduced as guard ring to improve fill factor and reduce pixel sizes [9].
However, it generally introduces high dark signal due to defects and charge trapping. To
avoid this issue, a deep n-well layer is adopted to realize a lower doped region as guard-ring,
preventing edge breakdown. Another benefit of this structure is an improvement of fill-factor
as SPADs can share their cathodes by placing the devices in a single deep n-well. To prevent
sharp edge introducing hot spots of high electrical field, the device corners are cut at 45
degrees. A metal shield is finally formed to avoid light transmission to non-active area. Each
SPAD array is composed of 25 (5 × 5) devices, which can be separately selected through
readout circuits. The fill-factor, defined as the ratio of active area to total area of the SPAD
array, is 39.9%.

![Cross section of the p+/n-well SPAD with a deep n-well guard ring. The shadowed areas indicate approximately the avalanche region.](image)

2.2 SPAD pixel and array readout

Each SPAD is integrated together with a pixel circuit, as shown in Fig. 2. The anode of the
SPAD is connected to a transistor M2, which acts as a passive quenching resistor and its
value can be adjusted by the gate voltage $V_{\text{Q}}$. A transistor M1, in series with M2, acts as a
switch to enable or disable the recharge of the SPAD. When an avalanche occurs in the
multiplication region of the SPAD, the diode delivers a current pulse and the anode voltage
quickly rises and then decreases, driven by transistor M2. The analog voltage signal is
converted to a 1.8V digital pulse through an inverting Schmitt-trigger comparator. A NOR
logic is employed to enable the output selection. An arbitrary subset of the SPADs can be
selected, while only two devices can be connected at a time through a 25-to-2 multiplexer
(MUX).
3. Characterization results

Ten test chips, each containing one 5 × 5 array (25 SPADs), have been fabricated in 150nm CMOS process. A printed circuit board (PCB) provides power supply, bias and reference voltage for the SPADs, while the acquisition was conducted with an external ripple counter or a digital oscilloscope connected to the chip output. A series of measurements have been performed to characterize the SPAD arrays.

3.1 Breakdown voltage

There are several methods to measure the breakdown voltage of a SPAD. The simplest way is to sweep the bias voltage while recording the output count rate, and subtract the Schmitt comparator’s threshold from the voltage where the count rate has a sharp transition. This method has a limited accuracy, since the comparator threshold voltage is affected by process variation. Thus, to precisely measure the breakdown voltage of single SPAD within one array, we monitored the device count rate as a function of applied voltage above the threshold and then extracted the breakdown voltage by extrapolation. The SPADs were illuminated with a blue LED and a diffuser to obtain a uniform count rate for all the devices. The light density was adjusted so that the count rate of all the devices was around 100kHz, high enough compared to DCR and sufficiently low in order not to affect the measurement linearity. The count rate as a function of bias voltage was obtained for 25 SPADs within one single chip, as shown in Fig. 3. The number of counts is initially zero and starts increasing as the anode voltage is approaching the Schmitt comparator’s threshold. Then, the counts rate can be treated as a linear function of the bias voltage in a certain region. The breakdown voltage was estimated as the crossing point of the linear fit and the x-axis. With this method, the inter-chip breakdown voltage variation is roughly 0.04V (max-min), which is negligible compared to the excess bias voltage (V_{EX}) used in all the applications. Characterization shows an 18.01V average breakdown and 0.5V standard deviation for chip-to-chip variation.
The breakdown voltage variation versus temperature was measured and it is shown in Fig. 4, indicating a 17mV/°C temperature coefficient.

### 3.2 Dark count rate

Even in the absence of illumination, an avalanche can be triggered by a dark count, which is due to non-photon generated carriers [8]. The dark count rate (DCR) measurement involves the characterization of primary pulses (primary DCR) and secondary pulses (afterpulsing). Primary DCR could be due to thermal generation effects, trap-assisted or band-to-band tunneling. The latter processes are typically the main dark signal generation mechanism in DSM process [10], while afterpulsing is correlated to primary avalanche, and is caused by trapping and de-trapping of carriers. To precisely characterize primary DCR, a relatively long dead time of 1µs was used, in order to make afterpulsing effect negligible. The DCR of 250 SPAD samples from 10 test chips were measured at room temperature, as shown in Fig. 5.
The observed median DCR is 0.4 Hz/µm² at 3V excess bias, and the mean DCR is 11.4 Hz/µm² in the same condition due to the 20% noisy SPADs.

To better investigate primary dark counts contributions, DCR versus temperature for three SPADs, exhibiting different behavior (indicated in Fig. 5) was measured at 3V excess bias and presented in Fig. 6. The breakdown voltage variation was also taken into consideration by tuning the bias voltage at each temperature. The activation energy (E_A) of SPADs was calculated using Arrhenius equation:

$$DCR = A \exp\left(-\frac{E_A}{k_B T}\right)$$

(1)

where $k_B$ is Boltzmann’s constant, $A$ is a constant and $T$ is the absolute temperature in Kelvin [11]. If $E_A$ is similar to the band gap energy, thermal carriers injected from the neutral regions are the main DCR factor; while if $E_A$ is much smaller than $E_G/2$, tunneling dominates the DCR [12]. One can observe that for low-DCR SPADs, the transition point between the two regions is around 27°C. Since activation energy $E_{A2}$ (0.18eV) is much smaller than $E_{A1}$ (1.11eV), tunneling is the dominant generation process in the −20°C-27°C range, and relatively independent of temperature. For SPADs featuring high DCR, however, no transition region is observed and trap-assisted tunneling dominates in all temperature ranges.
Fig. 6. DCR versus temperature for selected SPADs at 3V excess bias; the dashed lines indicate Arrhenius equation fit.

In Fig. 7, the DCR versus temperature for the low-DCR SPAD at different excess bias is shown. One can observe that activation energy $E_{A2}$ was reduced at higher excess bias, resulting in higher dark counts. It is also of interest to note that the cross point of Arrhenius equation fit (dashed lines) shifted to higher temperature when increasing excess bias, indicating that, as expected, the contribution of band-to-band tunneling increases at high voltages.

![Graph](image)

Fig. 7. The DCR versus temperature for a low-DCR SPAD at different excess bias, and the dashed lines indicate Arrhenius equation fit.

### 3.3 Afterpulsing

During an avalanche some carriers may be captured by traps in the depletion region, and released with a statistical delay, retriggering another spurious avalanche that increases the
SPAD noise. Since traps are due to impurities and crystal defects, a cleaner manufacturing process could reduce the afterpulsing probability.

In Fig. 8, the statistical avalanche behavior is sketched. When an avalanche occurs, the anode voltage of SPAD (V_A) exceeds the threshold voltage (V_th) of the Schmitt-trigger comparator, resulting in a digital pulse (V_D). The primary pulse indicates an avalanche caused by a photon or a primary dark pulse, as shown in Fig. 8(a). Because of the statistical nature of this phenomenon, afterpulsing may happen during the recovery phase or even after a longer time. In Fig. 8(b), the afterpulsing occurs after a certain time of the primary pulse, resulting in two consecutive pulses. However, a longer digital pulse is generated if the afterpulsing happens during the recharge phase as shown in Fig. 8(c). In the even worse case, two or more afterpulsing may occur, introducing a much longer digital pulse.

![Fig. 8. The primary pulse (PP) and Afterpulsing (AP) behaviour in the passive quenching circuit.](image)

Firstly, we measured the afterpulsing in Fig. 8(b) following the procedure described in [13]. We acquired digital output signals and recorded the inter-arrival time (i.e. the time between the pulse and its previous one). The histogram of inter-arrival times is shown in Fig. 9. We could extrapolate that pulses occurring after 10μs are all primary pulses by fitting the late part of the spectrum with a Poisson distribution. It is evident that all pulses of the early part (<0.8μs) are afterpulses, while the middle part of the spectrum contains a mixture of primary pulses and afterpulses. The afterpulsing probability is calculated from the histogram by subtracting the integration of the Poisson distribution from the total counts.

![Fig. 9. Inter-arrival time histogram measured at 3V excess bias with 50ns dead time.](image)

During the measurement, we also noticed a variation of the output pulses width (ex. from 50ns to 150ns) when we preset the dead time as 50ns. It is evident that afterpulsing occurs during the recharge phase, and its effect is to increase the digital pulse width, as indicated in Figs. 8(c) and 8(d). Thus, a complete characterization should also include this “inter-pulse” contribution as afterpulsing. Figure 10 shows the histogram of acquired pulse widths. It is
clear to separate primary pulses (50ns dead time on average) and afterpulsing. The width variation of primary DCR is mainly due to the timing jitter of the electronics, following a Gaussian distribution. Then the afterpulsing probability contribution due to the increase of the digital pulse width is calculated by subtracting the integration of the Gaussian distribution from the total counts.

![Fig. 10. Histogram of pulse widths, for SPAD of 50ns dead time, at V_{EX} = 3V, room temperature.](image)

The total afterpulsing versus excess bias was measured on a SPAD with 50ns dead time, as shown in Fig. 11. The observed increase is due to larger current density, which increases the number of carriers crossing the device area and thus their probability to be captured [8].

![Fig. 11. Afterpulsing as a function of excess bias with 50ns dead time.](image)

Afterpulsing effect can be significantly reduced by increasing the dead time, at the expense of limiting the maximum photon count rate. Figure 12 shows afterpulsing probability as a function of dead time at 3V excess bias voltage. With 150ns dead time, afterpulsing was reduced to 0.85% at 3V excess bias.
3.4 Photon Detection Probability

The photon detection probability (PDP) was measured at different bias conditions and for incident light wavelengths from 400nm to 1100nm, as shown in Fig. 13. The SPAD achieves a maximum PDP of about 31% at 450nm wavelength when biased at 5V excess bias. The PDP decreases for longer wavelengths due to the increasing absorption length of photons than cannot be detected by the SPAD shallow junction. PDP non-uniformity, which may reduce the dynamic range of a SPAD imager [14], was 1.1% (standard deviation) between the pixels in one array for the peak PDP at 5V excess bias.

3.5 Timing jitter

The measurement of the SPAD timing jitter was conducted with the time-correlated single-photon counting (TCSPC) technique, where a laser source operated at 1 MHz emitting light pulses and attenuated by neutral density filters to single photon level was used. The histograms of the time difference between the laser output trigger and the SPAD rising edge were collected, and the jitter was estimated in terms of full width half maximum (FWHM). Characterization results with two different lasers (PicoQuant GmbH with peak emission wavelength at 468nm and 831nm respectively) are presented in Fig. 14, where the timing response of a commercial SPAD (Custom process MPD PDM series [15]) is shown for comparison. By deconvolution of the jitters contribution from the lasers (63ps for blue and
44ps for infrared), the presented SPAD exhibits 52ps FWHM jitter for blue laser and 42ps for infrared laser at 4V excess bias. The secondary and tertiary peaks observed in the histograms are present in the shape of the laser-emitted optical pulses at high intensity.

![Graph showing timing jitter of SPAD and the reference MPD device at 4V excess bias voltage in response of 468-nm and 831-nm laser pulses.]

**3.6 Crosstalk**

When an avalanche event is triggered, a large number of hot carriers is generated in the SPAD high-field region. These carriers may recombine and emit photons, which can be absorbed thus triggering an avalanche in an adjacent SPAD [16]. Moreover, crosstalk can also be caused by the lateral diffusion of carriers, which also becomes a correlated noise source. In the literature, different techniques had been exploited to reduce crosstalk, such as increasing the distance between the devices, tailoring process steps (optical trenches), and gated-on mode circuitry, at the expense of smaller fill-factor [16–18].

![Graph showing crosstalk probability mapping for 24 SPADs with reference to SPAD (1,1) at 4.5V excess bias.]

Fig. 14. Timing jitter of SPAD and the reference MPD device at 4V excess bias voltage in response of 468-nm and 831-nm laser pulses.

Fig. 15. Crosstalk probability mapping for 24 SPADs with reference to SPAD (1,1) at 4.5V excess bias.
Coincidence measurements were performed to evaluate the correlation between the SPADs output signals, whose time stamps were recorded by a digital oscilloscope with 10GSa/s sampling rate. If digital pulses occur in a defined observation time window (2 ns in our measurement), then it was considered as a crosstalk event. This procedure was repeated for all 24 SPADs to obtain the crosstalk mapping with reference to SPAD (1,1), as shown in Fig. 15. The neighboring SPADs (SPAD (1,2) and SPAD (2,1)) exhibit a crosstalk probability lower than 2.5%, while the corner SPAD (SPAD (2,2)) has a crosstalk around 0.5%. It is of interest to note that crosstalk probability of SPAD (3,3) is 0.3%, possibly due to indirect optical crosstalk [15]. The crosstalk probability increases with excess bias for nearest SPADs as presented in Fig. 16. This behavior is due to both an increasing of charge flowing in the junction and a higher triggering probability.

![Crosstalk probability as a function of excess bias voltage.](image)

**Fig. 16.** Crosstalk probability as a function of excess bias voltage.

### 3.7 Random Telegraph Signal (RTS) noise

During the course of the above-mentioned investigation of DCR, temporal fluctuations of DCR between different states have been observed in two SPADs. Similar fluctuations, which are known as random telegraph signal (RTS) noise, have been previously observed in irradiated SPADs [19]. This instability is mainly due to deep level traps and defects causing superposition of several levels of fluctuations. Figure 17 shows the DCR measurement of three SPADs recorded for 10 hours at room temperature. The first two devices exhibit an RTS-like behavior, while the third one is free of RTS noise. The transitions between DCR levels are random and unpredictable for both cases. To clearly estimate these levels, the DCR histogram is presented in Fig. 18. One may notice that SPAD#1 has a main DCR level at around 254 Hz and multilevels ranging from 1.5 kHz to 3 kHz, while SPAD#2 has two distinct DCR levels at 1.48 kHz and 1.77 kHz, and multilevel ranging from 2 kHz to 2.9 kHz. In the end, the standard SPAD (SPAD#3) has a Poisson distribution of DCR with an average of 156 Hz.
Fig. 17. Different forms of RTS in two SPADs and normal behavior of SPAD#3 at 5V excess bias.

Fig. 18. DCR histogram of the three SPADs at 5V excess bias.
4. Comparison with the state-of-the-art CMOS SPADs

The performance of the SPAD is summarized in Table 1, compared with other SPADs fabricated in deep submicron CMOS technologies.

| Technology | This work | [20] | [21] | [22] | [23] | [24] |
|------------|-----------|------|------|------|------|------|
| Junction, GR structure | p+n-well, DNW virtual GR | p+n-well, p-well GR | p+n-well, p-type implants GR | DNW/p-sub, p-epi GR | p-well/DNW, virtual GR | p-well/DNW, p-epi GR |
| Active area (diameter) | 97.12 µm² (10µm) | 113.1 µm² (12µm) | 58 µm² (9µm) | 50 µm² (8µm) | 32 µm² (6.4µm) | 27.6 µm² (5.25µm) |
| Shape | Square | Circular | Octagonal | Circular | Square |
| Mean $V_{bd}$ (@RT) | 18.01 V | 11.3 V | 12.8 V | 20 V | 14.9 V | 12 V |
| Mean $V_{ex}$ | 3V | 3 V | 1 V | 2 V | 0.4 V | 3 V |
| Median DCR (@RT) | 39 Hz | 0.4Hz/µm² | 0.24kHz/µm² | 10 Hz | 0.17Hz/µm² | 0.36Hz/µm² | 18 Hz | 0.31Hz/µm² | 11kHz | 0.4kHz/µm² |
| Afterpulsing (@RT) | 0.85% (@150-ns dead time) | <0.1% (@200-ns dead time) | - | <4% (@500-ns dead time) | 0.375% (@15-ns dead time) | - |
| PDP peak | 31% (@450nm) | 25.4% (@490nm) | 28% (@490nm) | 72% (@560nm) | 44% (@690nm) | 27.5% (@425nm) |
| Timing jitter | 52ps (@468nm) | 65ps (@490nm) | 128ps (@V_{ex} = 2V) | <60ps (@V_{ex} = 6V) | 51ps (@690nm) | 136ps (@V_{ex} = 2.4V) |
| Crosstalk adjacent SPADs (@4.5V $V_{ex}$) | <2.5% | - | - | - | - | - |

DNW: deep N-well, GR: guard ring, $V_{bd}$: breakdown voltage, $V_{ex}$: excess bias voltage, RT: room temperature.

5. Conclusion

In this work, we have presented a SPAD array fabricated in standard 150nm CMOS process and its in-depth electro-optical characterisation. The array is composed by 25 (5 × 5) SPADs, which have 10µm active-area and 15.6µm pitch sizes, achieving a 39.9% array fill factor. The realized SPAD has a median DCR of 0.4Hz/µm² at 3V excess bias, which is comparable to or better than other devices with similar structures, fabricated in DSM technologies. The device achieved a peak PDP larger than 30% at 450nm wavelength, and shows very good uniformity within the array. With a dead time of 150nm, the SPAD exhibits low afterpulsing probability of less than 1%, and a timing jitter of 52ps FWHM and 42ps FWHM under at 468-nm and 831-nm wavelengths, respectively. We believe this array can be adopted in a number of applications requiring high fill factor, low noise and excellent timing jitter.