A CMOS Compatible Single-Photon Avalanche Diode

A. C. Giudice, M. Ghioni, S. Cova, F. Zappa
Politecnico di Milano - DEI
P. Leonardo da Vinci,32 Milano (Italy)
giudice@elet.polimi.it

E. Sciacca*, S. Lombardo, E. Rimini*
CNR-IMM Catania (Italy)
esciacca@imetem.ct.cnr.it

D. Sanfilippo, C. Difranco, G. Fallica
ST Microelectronics
Stradale Primosole 50
Catania (Italy)
delfo.sanfilippo@st.com

Abstract

Single-Photon Avalanche Diodes (SPAD) are a solid-state alternative to photomultipliers (PMT, vacuum tube detectors) for measuring very faint and/or fast optical signals. A SPAD device compatible with a standard CMOS technology is presented. The detector noise is very low: the dark counting-rate at room temperature is 30-40 c/s for 10μm-diameter devices and less than 5kc/s for 40μm-50μm devices. High time resolution is achieved, up to 60ps.

1. Introduction

Photon-counting and photon-timing techniques are widely used for accurate measurement of weak optical signals and fast luminescence [1] or fluorescence decays [2] (in the nanosecond or picoseconds range), in various applications such as optical component characterization [3], non-invasive testing of VLSI circuits [4] and DNA-separation [5].

Among available photodetectors, only two devices are able to detect single photons: photomultiplier tubes (PMTs) and avalanche photodiodes (APDs). PMTs operating in the visible spectrum range have good quantum efficiency and low noise at room temperature. PMTs for red and near-infrared wavelength range have lower quantum efficiency and require cooling of the photocathode. The intrinsic time resolution ranges from 1ns (ordinary PMTs) to less than 30ps (ultrafast micro-channel plate multipliers, MCPs) [6]. However, the PMT are bulky fragile, sensible to disturbances, they require high voltages power supply (2-3 kV) and their cost is fairly high. Semiconductor APDs have instead the typical advantages of solid-state devices (small size, low bias voltage and power consumption, ruggedness and reliability, suitability to build integrated systems, etc.). They also have higher photon detection efficiency than PMTs, specially in the infrared.

Single Photon Avalanche Diodes (SPADs) are p-n junctions biased above the breakdown voltage V_b with junction electric field high enough to initiate an avalanche multiplication process. When the diode is biased at operative level the diode current is zero for a considerably long time (hundreds of nanoseconds or even milliseconds), until a minority carrier is generated and succeeds in triggering the avalanche process. When the carrier is photogenerated, the leading edge of the avalanche onset, sensed by a fast discriminator, marks very precisely the photon arrival time. The avalanche current flows until an external circuit quenches the diode, lowering the bias voltage below V_b. After a dead-time, the bias is restored and the photodiode is ready to detect the arrival of another photon. By using an Active Quenching Circuit (AQC) [7,8], SPADs can work in accurately controlled bias conditions, and the dead time following each pulse can be made very short.

The device internal noise consists in a random dark-counting rate arising from free carriers thermally generated or emitted by deep levels. During an avalanche discharge a charge carrier can be trapped by deep levels in the depletion region, and if this carrier is released after the diode has been reset to the operative voltage, it produces a correlated after-pulse. To operate properly as a SPAD, the device must fulfil three requirements: (i) the breakdown voltage must be uniform over the entire active area; (ii) the dark-count rate must be low, or at least moderate (1 kcounts/second); (iii) the afterpulsing must be negligible or fast enough to be minimal after the dead time [7,9].

The time response of a single-photon detection is the statistical distribution of the delays from the arrival time of the photon to the onset of the electrical pulse of the detector. The intrinsic SPAD time-response curve has a fast peak followed by a slow tail [10]. The peak is due to photons absorbed in the depletion layer; its full width at half maximum (FWHM) gives the resolution of the detector and is related to the fluctuations of the avalanche build-up. The tail is due to minority carriers photogenerated in the neutral region beneath the junction that anyway succeed in reaching the depletion region by diffusion, and triggering the avalanche.

We present a new SPAD fabrication recipe, designed using processes compatible with a standard CMOS technology of ST Microelectronics, Catania (Italy). We developed SPADs with active area diameter from 10μm

* and Department of Physics, University of Catania (Italy).

to 75μm, 27.3V breakdown voltage, dark-counting rates as low as 10 counts/second (at room temperature), timing resolution of 80ps FWHM (at room temperature) and 50% quantum efficiency at 514nm wavelength.

2. SPAD structure and processes

Fig. 1 shows the cross-section of the SPAD structure. It's a n+p junction where the n+ shallow diffusion is obtained by depositing a n+ doped polysilicon layer.

The substrate is a low defect, low impurity density n-silicon wafer. The first processing step is the growth of a 2μm epitaxial layer highly doped with boron, to be used as a buried-layer. A lightly boron-doped epilayer is then grown with 7μm thickness [10]. The reasons to form a buried p-n junction are twofold. First, the time-response of the detector is improved by reducing the number of photogenerated free carriers in undepleted regions [10]. Second, isolation with the substrate is introduced in order to monolithically integrate many SPADs. The p++ buried-layer is necessary to reduce the series resistance of the device. The p- epilayer should be thin enough to limit the diffusion tail of the SPAD timing response.

Then p++ sinkers are created through an high-dose boron implantation step, so to reduce the contact resistance of the anode and provide a low resistance path to the avalanche current.

The next is a key step in the process, since it has a great impact on the device performance. We realized a heavy $POCl_3$ diffusion through an oxide mask. Heavy phosphorus diffusions are well known to be responsible for transition metal gettering [11]. Unfortunately the well known phosphorus predeposition on the backside of the wafer is not able to getter the distant device active area because metal diffusers diffuse too slowly during the final annealing. For this reason if gettering sites are created suitably close to the active region a major improvement is observed. For this reasons we performed the gettering process not on the back of the wafer but on the top and not farther than 20μm from the photodiodes as is shown in Fig. 2.

The next step is the p+ enrichment diffusion obtained with a low energy boron implantation, producing a peak concentration higher than 10^{16} cm^{-3}, followed by a high temperature annealing and drive-in [10]. In order to avoid edge breakdown we deposited a polysilicon layer, that was arsenic implanted or in situ doped, in a window that is larger than the enrichment zone. By increasing the boron concentration, the breakdown voltage of the abrupt n+p junction is reduced. Therefore, the breakdown voltage is higher in the outer zone than in the central of the photodiode [7]. Anyway, in order to avoid Zener breakdown at room temperature, the peak concentration of the enrichment must be suitably limited.

The polysilicon deposition is another important process for producing low dark-counting rate SPADs. The first devices were fabricated with a deposited polysilicon, doped by arsenic implantation. Better devices were then produced using a deposition of an "in situ" doped polysilicon. But the best results (in terms of dark-counting rate) were obtained using the polysilicon as a source for doping the substrate, instead of using it as the n+ part of our device [12]. An accurately engineered Rapid Thermal Anneal (RTA) at a low temperature was specifically designed in order to produce the required doping profile for the As in the silicon (a peak concentration of around $1x10^{19}$ at/cm^{-3}) and to shift the junction away from defects near the silicon-polisylicon interface.

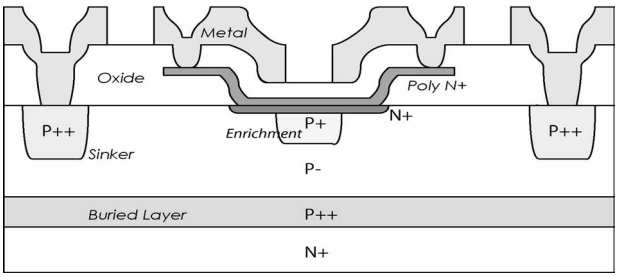

Figure 1. SPAD cross-section.

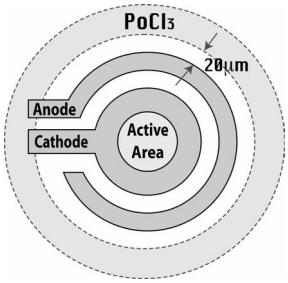

Figure 2. SPAD top view.

Devices with active-area diameters from 10μm to 75μm were designed; small devices provide higher timing resolution, but larger devices are required in many applications.

3. Experimental Characterization

A fairly uniform breakdown voltage of about 27.3V is obtained. The series resistance is few hundred Ohms for larger devices and increases up to 17kΩ for the 10μm devices. This is due to the high resistivity of the neutral p- epilayer below the active area and can be reduced to about 1kΩ by reducing to 3-4μm the epilayer thickness. The dark-counting rate was measured at different values of the overvoltage (difference between the actual bias voltage and V_b). By increasing the overvoltage the jitter of the time response is reduced, the avalanche triggering probability is increased (thus increasing the quantum efficiency), but the dark-counting rate is increased. A suitable trade off must therefore be decided for the overvoltage, according to the application requirements.

3.1 Dark-counting rate

Table 1 shows the average dark-counting rates measured for devices with various active area diameters at two different overvoltages.

Table 1. Dark-counting rates for different active area diameters at two different overvoltages.

Active Area Diameter	Dark Counts @ 5V overvoltage	Dark Counts @ 10V overvoltage
10 μm	23 c/s	40 c/s
20 μm	267 c/s	386 c/s
30 μm	750 c/s	1.1 kc/s
40 μm	2.2 kc/s	3.3 kc/s
50 μm	3.6 kc/s	5 kc/s
75 μm	7 –56 kc/s	10 – 84 kc/s

Table 2. Highest and lowest dark-counting rates measured at 5V overvoltage for different SPADs, fabricated with previous recipes.

Active Area Diameter	Dark Counts (Lowest)	Dark Counts (highest)
10 μm	510 c/s	20 kc/s
20 μm	5 kc/s	30kc/s
40 μm	13 kc/s	350 kc/s
50 μm	16 kc/s	360 kc/s

Small area (10μm and 20μm) photodiodes have very low dark-counting rate even at high overvoltage and larger devices have reasonably low dark-counting rate (<5kc/s). Smaller photodiodes present the same characteristics across the entire wafer, whereas larger ones have to be selected (specially the 50μm or the 75μm).

For comparison, Table 2 shows the dark-counting rate measured for SPADs fabricated with a backside gettering step and a shorter RTA.

3.2 Timing resolution

The SPAD time resolution curve was obtained by using a time-correlated photon counting (TCPC) apparatus [13]. The experimental setup, as shown in Fig. 3, includes a pulsed (10kHz) picosecond (~10ps FWHM) laser at 820nm, which triggers the avalanche on the detector and provides a synchronization electrical pulse. A time to amplitude converter (TAC) and a multichannel analyzer (MCA) collect the photon arrival times, measured with respect to the synchronization signal from the laser. By collecting data over many laser pulses, the histogram of the photon arrival time is collected. The shape of the histograms results from the convolution of the laser pulse shape with the instrumental response.

Since the FWHM of the laser pulse is around 10ps and the measured jitter of the experimental setup is 12ps, a histogram with FWHM larger than 30ps reasonably represents the intrinsic SPAD impulse response.

Fig. 4 shows the typical histogram obtained, which is characterized by a narrow peak and a slow tail.

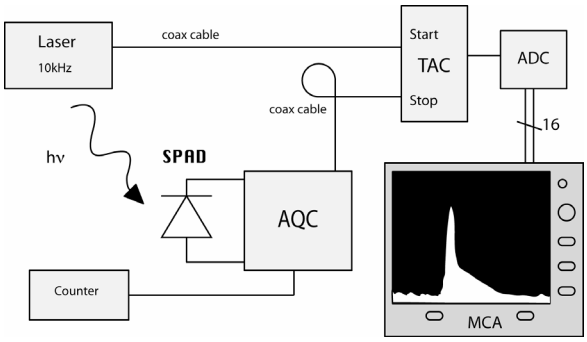

Figure 3. Time correlated experimental setup.

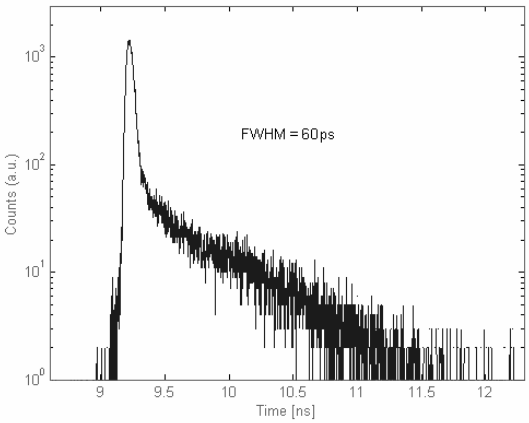

Figure 4. Time response of a 10μm-SPAD biased 10V above breakdown.

The peak is due to the photons absorbed in the depletion region and its width depends on the statistical fluctuations of the avalanche build-up time. The tail is due to minority carriers (in our case electrons) photogenerated in the neutral region below the space charge zone, that diffuse slowly toward the latter and trigger the avalanche. It can be demonstrated [10] that the time constant of the tail is equal to

$$\tau = \frac{W^2}{\pi^2 D_n} \qquad (1)$$

where W is the thickness of the neutral undepleted region and D_n is the diffusion coefficient for electrons.

The equivalent neutral region thickness of the fabricated SPAD can be calculated by substituting in Eq.1 the time constant τ measured from the time resolution curves (about τ = 580ps). The calculated thicknesses ~2.4μm is in agreement with the process simulations.

Table 3 shows the typical SPAD timing resolution as a function of the overvoltage for SPADs of different active areas. The jitter increases with the active area and decreases at higher overvoltages (higher electric field across the junction) [14]. For the smaller devices a resolution under 80ps is currently achieved at room temperature.

Table 3. Typical timing resolution for SPADs with different active areas measured at three different overvoltages.

SPAD Active Area	FWHM @ 5V	FWHM @10V	FWHM @ 13V
10 μm	161 ps	77 ps	74 ps
20 μm	184 ps	144 ps	113 ps
30 μm	236 ps	181 ps	165 ps
40 μm	285 ps	232 ps	188 ps
50 μm	330 ps	250 ps	-

By cooling the photodetector time resolution can be improved. As shown in Fig. 5, a FWHM of 95ps is achieved for a 20μm SPAD at –20°C, with a 21% FWHM reduction compared to room temperature. By using commercially available Peltier coolers it is easy to operate in the range of temperatures shown in Fig. 5 (room temperature/-30°C). The reduction of time jitter is due to a reduced jitter in the avalanche buildup. At lower temperatures the mean free path is longer and carriers accelerated by the electric field have higher probability to acquire enough energy to impact-ionize lattice atoms, thus speeding up the statistical process of the avalanche buildup.

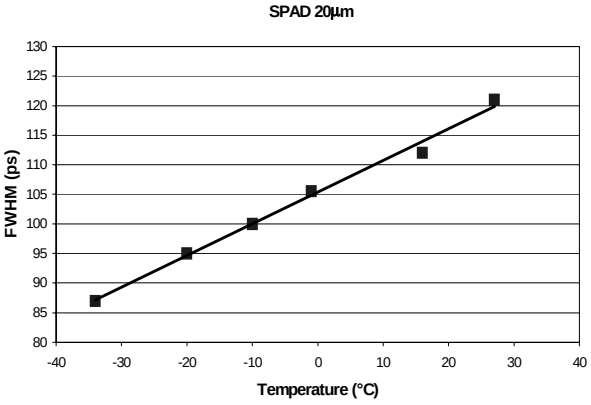

Figure 5. FWHM for a 20μm diameter SPAD as a function of the temperature at 10V overvoltage.

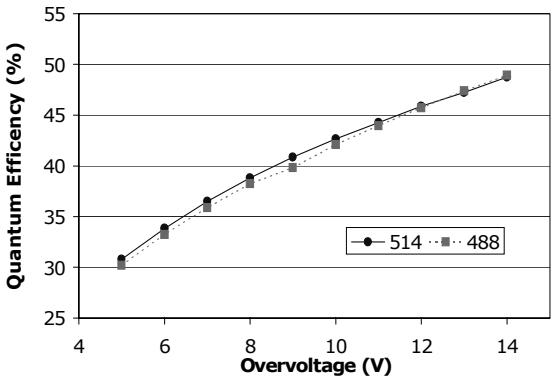

Figure 6. Measured quantum efficiency at two wavelengths, as a function of the overvoltage.

3.3 Quantum efficiency

Fig. 6 reports the quantum efficiencies measured at two different wavelengths (488nm and 514nm). The quantum efficiency increases with the overvoltage because of the higher ignition probability (i.e. that the photogenerated carrier triggers the avalanche).

References

[1] T. A. Louis, G. Ripamonti and A. Lacaita, "Photoluminescence lifetime microscope spectrometer based in time correlated single photon counting with an avalanche diode detector", *Rev. Sci. Intrum.*, vol.63, pp. 2994-2998, 1992.

[2] S. Cova, A. Longoni, A. Andreoni, R. Cubeddu, "A semiconductor detector for measuring ultra-weak fluorescence decays with 70ps FWHM resolution", *IEEE J. Quantum Electron.,* vol. QE-19, pp. 630-634, 1983.

[3] A. Lacaita, P. A. Francese, S. Cova and G. Ripamonti, "Single-photon optical-time-domain reflectometer at 1.3 μm with 5 cm resolution and high sensitivity", *Op. Lett.*, vol. 18, pp. 1110-1112, 1993.

[4] F. Stellari, F.Zappa, S. Cova and L. Vendrame, "Tools for noninvasive optical characterization of CMOS circuits", in *Proc. IEDM '99,* Washington, DC, Dec. 5-8,1999.

[5] G. Ocvirk, T. Tang and D. J. Harrison, "Optimization of confocal epifluorescence microscopy for microchip-based miniaturized total analysis systems", *Analyst*, vol. 123, pp. 1429-1434. July 1998.

[6] H. Kume, K. Koyama, N. Nakatsugawa, S. Suzuki and D. Fatlowitz, "Ultrafast microchannel plate photomultipliers", *Appl. Opt.,* vol. 27, pp. 1170-1178, 1988.

[7] S. Cova, M. Ghioni, A. Lacaita, C. Samori, F. Zappa, "Avalanche photodiodes and quenching circuits for single- photon detection", *Appl. Optics*,vol. 35, pp 1956-1976, 1996.

[8] F.Zappa, M. Ghioni, S. Cova, C.Samori, A. C. Giudice, "An Integrated Active-Quenching Circuit for Single-Photon Avalanche Diodes", *IEEE Trans. On Instrumentation and measurements,* vol. 49, pp 1167-1175, December 2000.

[9] S. Cova, A. Lacaita, G. Ripamonti, "Trapping Phenomena in Avalanche Photodiodes on Nanoseconds Scale", *IEEE Electron Device Letters,* vol. 12, no. 12, pp. 685-687, December 1991.

[10] A. Lacaita, M. Ghioni, S. Cova, "Double epitaxy improves single-photon avalanche diode performance", *Electronics Letters,* vol. 25, no. 13, 22 June 1989.

[11] H. Hieslmair, S. A. McHugo, A. A. Istratov and E. R. Weber, "Gettering of transition metals in silicon", *EMIS Dtareview series INSPEC,* No. 20, 1999.

[12] W. J. Kindt and H. W. van Zeijl, "Fabrication of Geiger mode Avalanche Photodiodes", *Nuclear Science Symposium*, vol. 1, pp. 334-338, 1997.

[13] V. O'Connor and D. Phillips, "Time-correlated Single Photon Counting", *Academic Press*, London, 1984.

[14] A. Lacaita, M. Mastrapasqua, "Strong dependence of time resolution on detector diameter in single photon avalanche diodes", *Electronic Letters,* vol. 26, no. 24, 22 November 1990.