First prototypes of two-tier avalanche pixel sensors for particle detection

L. Pancheri [a,b,*], P. Brogi [c,d], G. Collazuol [e,f], G.-F. Dalla Betta [a,b], A. Ficorella [a,b], P.S. Marrocchesi [c,d], F. Morsani [d], L. Ratti [g,h], A. Savoy-Navarro [d,i]

[a] Università di Trento, Dipartimento di Ingegneria Industriale, I-38123 Trento, Italy
[b] TIFPA INFN, I-38123 Trento, Italy
[c] Università di Siena, DSFTA, I-53100 Siena, Italy
[d] INFN, Sezione di Pisa, I-56127 Pisa, Italy
[e] Università di Padova, Dipartimento di Fisica e Astronomia, I-35131 Padova, Italy
[f] INFN, Sezione di Padova, I-35131 Padova, Italy
[g] Università di Pavia, Dipartimento di Ingegneria Industriale e dell'Informazione, I-27100 Pavia, Italy
[h] INFN Sezione di Pavia, I-27100 Pavia, Italy
[i] University Paris-Diderot/CNRS-IN2P3, 75205 Paris Cedex 13, France

A B S T R A C T

In this paper, we present the implementation and preliminary evaluation of a new type of silicon sensor for charged particle detection operated in Geiger-mode. The proposed device, formed by two vertically-aligned pixel arrays, exploits the coincidence between two simultaneous avalanche events to discriminate between particle-triggered detections and dark counts. A proof-of-concept two-layer sensor with per-pixel coincidence circuits was designed and fabricated in a 150 nm CMOS process and vertically integrated through bump bonding. The sensor includes a 48 × 16 pixel array with 50 μm × 75 μm pixels. This work describes the sensor architecture and reports a selection of results from the characterization of the avalanche detectors in the two layers. Detectors with an active area of 43 × 45 μ m^2 have a median dark count rate of 3 kHz at 3.3 V excess bias and a breakdown voltage non-uniformity lower than 20 mV.

1. Introduction

Several charged particle tracking applications call for the development of improved pixel detectors with low material budget and high spatial and temporal resolution, in addition to low power consumption and radiation hardness [1]. In order to meet these specifications, aggressive detector thinning has to be performed, but this operation strongly reduces the available signal, imposing severe constraints on the noise and uniformity of readout electronics.

To cope with this challenge, on the one hand the research community is striving to optimize silicon detectors and readout electronics, working both on the monolithic and hybrid approaches. From this point of view, progress in CMOS process technologies is beneficial, as device shrinking is continuously progressing and new features, such as 3D integration, are becoming more reliable. On the other hand, new detectors and new materials are being actively investigated with the goal of finding alternative and better performing solutions.

The introduction of avalanche gain is seen as an appealing perspective to recover the signal lost by device thinning and improving the timing resolution [2]. Low Gain Avalanche Detectors (LGAD), operating in sub-Geiger regime, could in principle reach a timing resolution in the order of tens of picoseconds.

Geiger-mode avalanche diodes have also been recently proposed as particle detectors [3,4]. These devices are currently well established for photon detection applications, and constitute the basic cells of Silicon Photomultipliers (SiPMs). However, their efficient use as particle detectors calls for a dramatic reduction of the dark count rate (DCR), which for SiPMs is typically of the order of 100 kHz/mm^2. To achieve this goal, the coincidence between two vertically-aligned avalanche detectors can be exploited, as illustrated in Fig. 1 [3]. Thanks to cell-to-cell coincidence, the dark count rate DCR_C of a two-layer pixel will be reduced to

$$DCR_C = DCR_1 \cdot DCR_2 \cdot 2\Delta T \quad (1)$$

where DCR_1 and DCR_2 are the dark count rates of the two cells, respectively, while ΔT is the coincidence time resolution.

This approach offers several advantages in applications requiring low material budget and fine detector segmentation as, for instance, for tracking and vertex reconstruction in particle physics

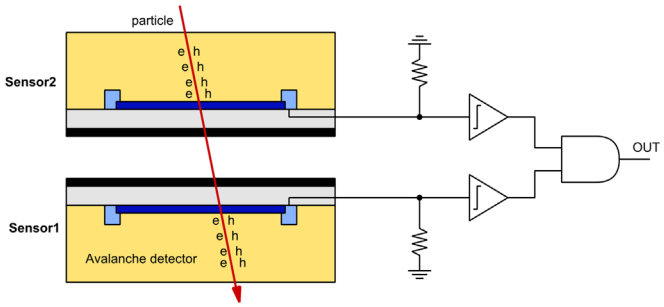

Fig. 1. Schematic illustration of coincidence detection principle.

experiments and charged particle imaging in medicine and biology. In addition, a timing resolution in the order of tens of picoseconds can potentially be achieved thanks to the fast onset of avalanche multiplication in Geiger-mode regime [5].

This detection mode was first studied exposing SiPM detectors to high energy particles of charge $Z=1$ and to ion beams with $Z>1$ [3,6]. The measurements have shown promising results, demonstrating a detection efficiency significantly larger than the geometrical fill factor of SiPMs.

To fully exploit the potential of this idea, however, it is necessary to monolithically integrate the readout electronics with avalanche detectors. Starting from the early 2000s, the co-integration of Geiger-mode avalanche diodes with readout electronics has been demonstrated in several CMOS technology nodes [7]. This has led to the realization of pixelated sensors for optical detection, a.k.a. Single-Photon Avalanche Diode (SPAD) arrays, for time-resolved sensing applicatons such as Time-of-Flight ranging and Fluorescence Lifetime Imaging [8].

The pixel particle sensor presented in this paper is based on the experience gained by the authors in the design of CMOS SPAD arrays. In our design, two layers of avalanche pixels are vertically integrated, with the coincidence detection performed at the pixel level.

The paper is organized as follows. Section 2 introduces the pixel concept and sensor design, while Section 3 presents a characterization of the avalanche detectors included in the two-layered structure. The main results are summarized in Section 4.

2. Chip design

A two-tier sensor assembly was designed and fabricated in a commercial 150 nm CMOS process. Two different types of Geiger-mode avalanche detectors were used in this design; their cross section is shown in Fig. 2. Type-1 detector has a shallow p+/nwell junction, while type-2 is based on a deeper pwell/deep nwell junction. Both devices are described in detail elsewhere [9,10].

In both cases, the active volume is less than 2 μm thick, and both detectors are isolated from the substrate thanks to a deep-nwell. On one hand, this can be an advantage if the sensors need to be thinned, since in principle it is possible to thin the dies down to a few micron without compromising its functionality. On the other hand, the fluctuations of the generated charge in such a small

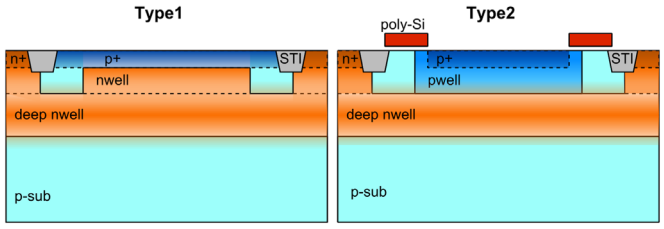

Fig. 2. Cross section of the two detector types.

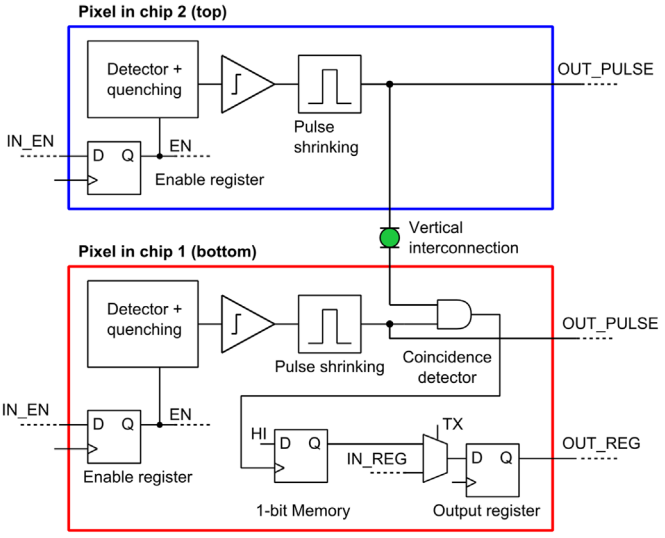

Fig. 3. Schematic diagram of two-layer pixel.

volume would be very large, and could lead to a loss of efficiency. An extensive experimental campaign will thus be necessary to accurately estimate the efficiency obtainable with this approach.

A simplified schematic diagram of the pixels, showing both layers, is sketched in Fig. 3. In the pixels, the detectors are passively quenched and their output signals are digitized by means of a low-threshold comparator. The resulting pulses are shortened by a programmable-length monostable circuit, providing a minimum pulse width of 750 ps. The pixels can be independently enabled or disabled with an arbitrary pattern, defined by a configuration register. The output of the monostable in the top half-pixel feeds a coincidence detector located in the bottom layer, and the coincidence output is stored in a 1-bit memory. Data from all the pixels can be transferred in parallel to an output register for readout and sent off-chip through 8 output pads. In this way, signal detection and data readout can be done simultaneously, thereby avoiding any dead time in the data acquisition process.

The monostable output signals are also sent to a row-wise OR gate, combining the outputs of all the active pixels in a row to a single signal stream. This feature was included to improve the testability of the chip and allows a wide range of tests if combined with proper setting of the configuration register, for example to map the dark count rate (DCR) between different pixels.

In addition, a row-wise coincidence detection circuit has also been included. The output of two given rows can be connected to the row coincidence detector, as shown in Fig. 4, allowing for the study of the coincidence among any arbitrary configuration of pixels between two rows. This feature can be used to measure the cross-talk between different pixels or groups of pixels in different rows.

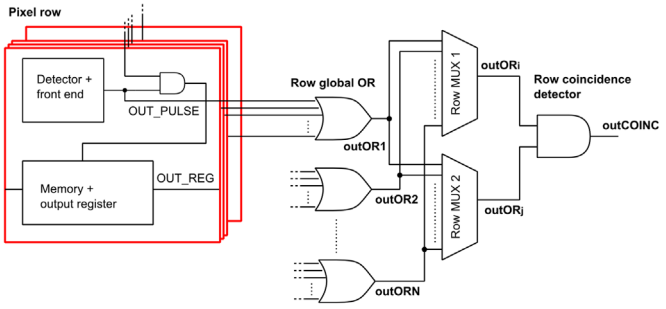

Fig. 4. Schematic diagram illustrating row-wise coincidence detection circuit.

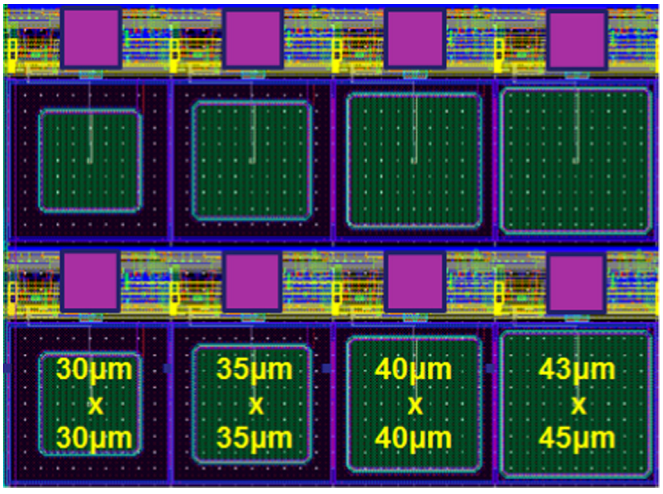

Fig. 5. Layout of pixels with different detector sizes.

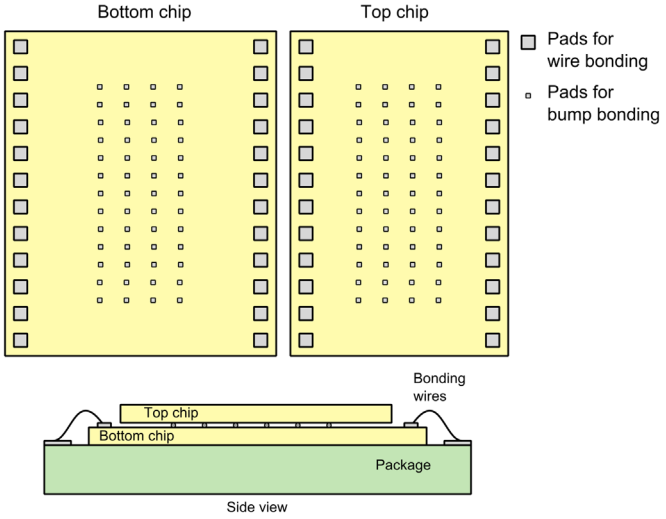

Fig. 6. Floorplan of the chips and cross-section of two-layer sensor assembly.

The core of the sensor consists of a 48×16 pixel array, partitioned in different regions. Each pixel, with a $50\ \mu m \times 75\ \mu m$ area, includes detectors and electronics on both layers, with the top-layer signal transmitted to the bottom layer using a vertical interconnection. Both types of detectors shown in Fig. 2 were included in the array, in order to choose the best one for the application. In addition, the array contains detectors with different active areas, starting from $30 \times 30\ \mu m^2$ to $43 \times 45\ \mu m^2$ (Fig. 5). The main goal of these splittings is to characterize the detector efficiency as a function of the geometrical fill factor.

A bump bonding technique using 12-µm solder bumps was chosen for the vertical integration, due to the accessibility and good yield of the process. Most of the pixels were covered with a metal shield to avoid inter-layer optical cross-talk. A few pixels were left unshielded to study vertical crosstalk and allow for optical measurements. A floorplan and a cross section of the two-layer sensor are shown in Fig. 6.

3. Experimental results

This section shows a selection of results from the characterization of the detector arrays. A few samples of top and bottom

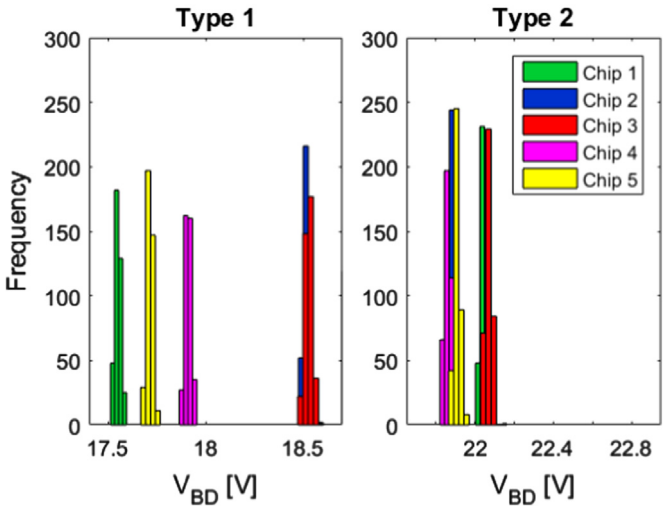

Fig. 7. Avalanche detector breakdown voltage distribution measured on 5 different chips.

chips were wire-bonded for testing before proceeding to vertical integration. Electrical tests confirmed the correct functionality of the avalanche detectors and readout electronics in both chips.

Since the avalanche detectors are co-integrated with the readout circuits, the anodes are not accessible and the cathodes are in common to all the devices. Therefore, current–voltage curves for the extraction of breakdown voltage cannot be measured directly. The breakdown voltage was extracted for each device from DCR vs. bias voltage curves, as discussed in [11]. The breakdown voltage histograms measured on 5 different dies are shown in Fig. 7 for both detector types. A very good breakdown voltage uniformity was found for detectors of the same type in each chip, with a standard deviation lower than 20 mV and a peak-to-peak variation in the order of 100 mV. The variability between devices belonging to different chips, however, is much larger, especially for Type-1 detectors.

The DCR was measured at different bias voltages. For devices of the same type, it spans several orders of magnitude, as observed in different processes [7]. Fig. 8 shows the DCR cumulative distributions of $43\ \mu m \times 45\ \mu m$ detectors measured on 5 different chips. The distribution has a good uniformity between different chips, with a median DCR of the order of 3 kHz for both device types at 3.3 V excess bias. This rate corresponds to 1.5 Hz/µm^2, an order of magnitude larger than typical SiPMs. This result is not

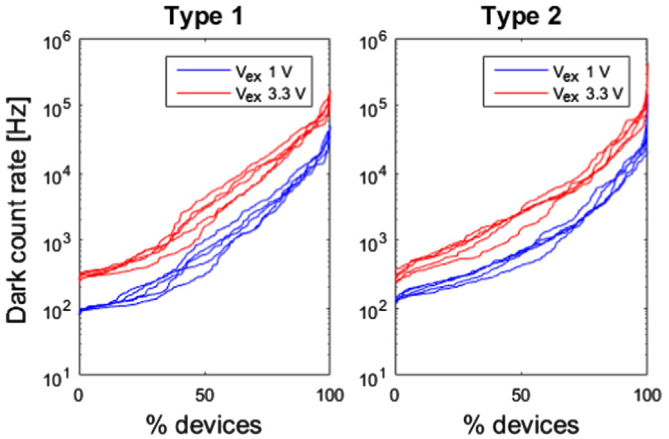

Fig. 8. Cumulative distributions of dark count rate measured on 5 different chips, for both avalanche detector types. Measurements include only $45 \times 43\ \mu m^2$ detectors.

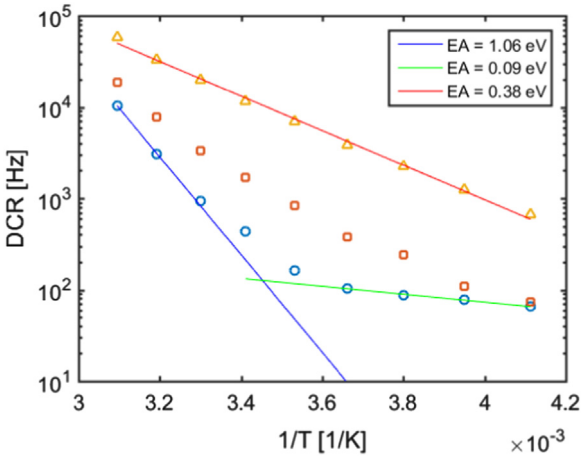

Fig. 9. Temperature dependence of dark count rate for three nominally equal detectors (Type-1, 43 × 45 μm^2, V_{ex}=3.3 V).

surprising, since the technology adopted in this work is a standard CMOS without any dedicated processing steps. For the best devices in the distributon, however, DCR is in the order of 300 Hz (0.15 Hz/μm^2), comparable with commercial SiPM cells of the same size.

A characterization of DCR as a function of temperature has been conducted to gain some insight into the dominant DCR-generating physical mechanisms in this technology. Fig. 9 shows the temperature dependence of three devices representative for the whole distribution in the range −30 °C + 50 °C. Although the devices have the same area, it can be observed that different activation energies can be inferred from the data. Devices with the largest DCR show an activation energy E_A smaller than $E_G/2$, indicating that DCR is dominated by trap-assisted tunneling. For the detector with the lowest DCR, a transition region between two slopes is oberved. At low temperatures, E_A is very small, indicating that probably band-to-band tunneling is the dominant mechanism. At high temperatures, E_A approaches the silicon band gap, and the DCR is likely to be dominated by injection from the neutral regions.

The row-wise coincidence detector was used to test the cross-talk between different pixels in the same column, exploiting the sensor programmability to enable only two pixels at a time. In the measurement, the coincidence rate was measured as a function of the distance between pairs of pixels. A normalized coincidence rate CR_n was defined as:

$$CR_n = \frac{CR_{meas}}{CR_1 \cdot CR_2 \cdot 2\Delta T} \quad (2)$$

where CR_{meas} is the measured coincidence rate, CR_1 and CR_2 are the count rates of the two pixels, ΔT is the width of the output pulses. The measurements were done on unshielded pixels and the detectors were illuminated to obtain a uniform count rate. CR_n is plotted in Fig. 10 as a function of the distance beween the centers of the detectors, expressed in units of row pitch (75 μm). The measured values are shown for both detector types and for two different voltages. For distant detectors, CR_n=1, while nearest neighbor pixels feature $CR_n > 1$ due to optical cross-talk. As

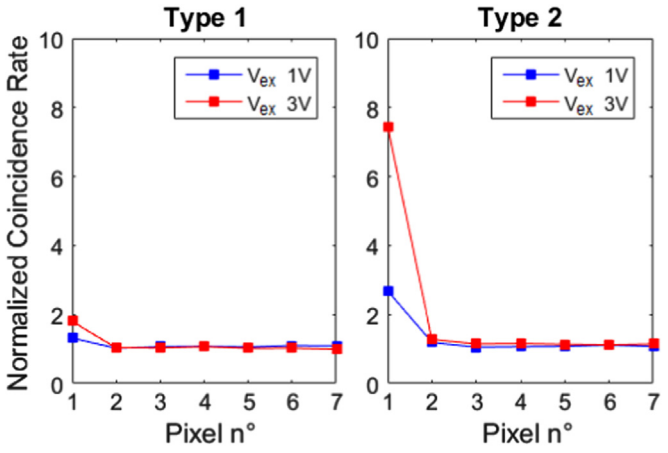

Fig. 10. Normalized coincidence rate as a function of pixel distance.

expected, this effect is larger in Type-2 detectors, that have a deeper junction, and increases with the breakdown voltage.

4. Conclusion

A pixelated sensor based on two vertically-integrated layers of avalanche detectors has been designed and fabricated in a commercial CMOS process. The functionality of the two layers has been verified and the integrated detectors have been separately characterized on several dies. A measurement campaign has been conducted to characterize the breakdown voltage uniformity and the DCR of the detectors. Although the standard CMOS process used for this design is not optimized for this application, the detector characteristics are suitable for a proof-of-concept validation. Currently, the vertical integration of the sensors has been completed and packaging is under way. Measurements on two-tier sensors will be presented in a future paper.

Acknowledgments

The research activity presented in this paper has been carried out in the framework of the APiX2 experiment funded by Istituto Nazionale di Fisica Nucleare (INFN), CSN5.

References

[1] G. Casse, Nucl. Instrum. Methods Phys. Res. Sect. A 732 (2013) 16.
[2] N. Cartiglia, et al., J. Instrum. 9 (2014) C02001.
[3] N. DAscenzo, et al., J. Instrum. 9 (2014) C03027.
[4] E. Vilella, et al., Nucl. Instrum. Methods Phys. Res. Sect. A 731 (2013) 103.
[5] S. Cova, et al., Appl. Opt. 35 (1996) 1956.
[6] P.S. Marrocchesi, et al., IEEE Trans. Nucl. Sci. 61 (2014) 2786.
[7] G.-F. Dalla Betta, et al., Avalanche photodiodes in submicron CMOS technologies for high-sensitivity imaging, in: Advances in Photodiodes, Intech House, 2011.
[8] E. Charbon, Philos. Trans. R. Soc. A 372 (2014) 20130100.
[9] L. Pancheri, D. Stoppa, Proceedings of IEEE ESSDERC, 2011, p. 179.
[10] L. Pancheri, et al., IEEE J. Sel. Top. Quantum Electron 20 (2014) 328.
[11] L. Pancheri, et al., Proceedings of IEEE ICMTS, 2014, p. 161.