Nuclear Instruments and Methods in Physics Research A 731 (2013) 103–108

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

3D integration of Geiger-mode avalanche photodiodes aimed to very high fill-factor pixels for future linear colliders

E. Vilella *, O. Alonso, A. Diéguez

Department of Electronics, University of Barcelona (UB), C/ Martí i Franquès 1, 08028–Barcelona, Spain

ARTICLE INFO

Available online 14 May 2013

Keywords:
3D-IC technology
CMOS
Fill-factor
GAPD
Linear colliders
Pixel detector

ABSTRACT

This paper presents an analysis of the maximum achievable fill-factor by a pixel detector of Geiger-mode avalanche photodiodes with the Chartered 130 nm/Tezzaron 3D process. The analysis shows that fill-factors between 66% and 96% can be obtained with different array architectures and a time-gated readout circuit of minimum area. The maximum fill-factor is achieved when the two-layer vertical stack is used to overlap the non-sensitive areas of one layer with the sensitive areas of the other one. Moreover, different sensor areas are used to further increase the fill-factor. A chip containing a pixel detector of the Geiger-mode avalanche photodiodes and aimed to future linear colliders has been designed with the Chartered 130 nm/Tezzaron 3D process to increase the fill-factor.

1. Introduction

Vertically integrated circuits (3D-ICs) are a very promising alternative for silicon pixel detectors aimed to High Energy Physics (HEP) experiments at future particle colliders. The potential of 3D technologies lies in the fact that they allow to achieve higher densities of integration without using nanometer technologies, which complicate the design of analog circuits and suffer from high technologic dispersions. Another advantage of 3D-ICs is that they also allow the separation of functional blocks (sensor, analog and digital) into different layers [1]. Moreover, there exist groups working with the traditional HEP hybrid pixel approach in 3D (ATLAS effort for 3D integration) with the goal of reducing the pixel size while keeping the 130 nm feature size. 3D technologies can also be used to increase the fill-factor of GAPD (Geiger-mode Avalanche PhotoDiodes) detectors, which rarely exceeds 10% when fabricated in conventional 2D technologies. Amongst other challenging requirements, a 100% fill-factor is demanded by future linear colliders on detector systems [2].

In this paper, a GAPD pixel array aimed to particle tracking/vertexing in future linear colliders and fabricated in a 3D process to maximize the fill-factor is presented. The paper is organized as follows. Section 2 describes the principle of operation of GAPDs, their main problems and known solutions. Section 3 introduces the Chartered 130 nm/Tezzaron 3D process. Section 4 presents the detector architecture as well as an analysis of the achievable fill-factor with the Chartered 130 nm/Tezzaron 3D process. Section 5 summarizes the conclusions.

2. Geiger-mode avalanche photodiodes

In conventional CMOS technologies, GAPDs are generally implemented by means of a simple p^+/n-well junction and biased above the breakdown voltage (V_{BD}) to operate in the Geiger-mode [3,4]. At this polarization, an electric field of the order of 10^6 V/cm is created in the depletion region. Upon absorption of impinging radiation within the depletion region, e^-–h^+ pairs are generated. These e^-–h^+ pairs can be accelerated by the high electric field up to the point at which they can in turn generate other e^-–h^+ pairs by impact ionization. The new pairs can be accelerated as well and generate more pairs also by impact ionization, thus starting an avalanche multiplication process that gives rise to a macroscopic current pulse in picoseconds. This current pulse can be detected and digitized by the readout electronics. However, since the avalanche is self-sustaining, the current continues to flow and it needs to be stopped in order to avoid self-heating or even burning the sensor. This operation is performed by the quenching electronics by lowering the reverse bias voltage down to or below V_{BD}. Once the avalanche has been quenched, the GAPD returns to its original state (i.e. it is polarized again) so that the sensor is made sensitive again for upcoming particles.

In contrast with other sensor technologies that are also being developed for future linear colliders, GAPDs enable single hit detection at each BX (Bunch Crossing) thanks to their extraordinary sensitivity and picosecond rise times. In spite of these advantages, GAPDs suffer from two main problems. First, there

* Corresponding author. Tel.: +34 934 039 157; fax: +34 934 021 148.
E-mail address: evilella@el.ub.edu (E. Vilella).

0168-9002/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.nima.2013.05.022

exist noise pulses, which cannot be distinguished from radiation triggered events and thus worsen the detector occupancy and SNR (Signal-to-Noise Ratio). These fake pulses, expressed in counts per second (Hz), are generated by dark counts, after-pulses and cross-talks. Dark counts are spurious pulses due to thermal and band-to-band tunneling carriers within the junction. After-pulses and cross-talks are correlated noise pulses. After-pulses are generated when carriers captured by trapping centers during an avalanche flow are released after the sensor has recovered its polarization condition. Cross-talks are fake pulses which appear when an avalanche in one pixel triggers a secondary avalanche in a neighboring pixel. Usual noise frequencies range from a few kHz to tens of kHz. Second, the fill-factor of GAPDs is too low. It rarely exceeds the 10%, which results in a low detection efficiency. Such a reduced fill-factor is due to two aspects of the design of the pixel, which are the non-sensitive areas of the sensor and the readout electronics. The non-sensitive areas of the sensor include the guard ring surrounding the p^+/n-well junction. Additionally, in deep-submicrometer technologies, the masks that the designers introduce in the layout to block the Shallow Trench Isolation (STI) also create non-sensitive areas. In a conventional CMOS process, the diode geometry creates a higher electric field at the edges, which leads to premature edge breakdown. To avoid this unwanted effect, the junction of the diode is surrounded by a guard ring with a lower doping profile. However, the guard ring usually is non-sensitive. In addition, for those technologies below the 0.25 μm node, a SiO_2 STI is compulsorily constructed in the fabrication process to prevent punch-through and latch-up. Punch-through is the existence of a parasitic current path located below the gate which shorts the drain and source terminals of CMOS transistors. Latch-up is the inadvertent creation of a low-impedance path between the high and the low power supply terminals of CMOS circuits. Both phenomena increase the power consumption and therefore they must be avoided. Nevertheless, the presence of the STI near the GAPD depletion region may induce extremely high levels of noise at frequencies above several MHz [5]. Fortunately, there exist several design techniques at the layout level to force the physical separation of the STI interface from the GAPD depletion region and obtain a beneficial impact on the noise, but at the expense of reducing the fill-factor [6–8]. The readout electronics is also monolithically integrated with the sensor on the same die to improve the dynamic response. Even though readout circuits based on a simple voltage comparator and a memory element are typically used, and thus with a small number of transistors, the area occupied by the transistors is still too large when compared to the sensor area. As a result, the non-sensitive area of the pixel chip is quite large compared with the sensitive area.

The noise issue can be handled with using advanced techniques, such as the time-gated operation or particle sampling at various layers. The GAPD detector presented in this work is operated in time-gated mode [9]. In this regime of operation, the sensor is periodically activated and deactivated under the command of a trigger signal to reduce the probability of detecting noise pulses that interfere with the radiation triggered pulses. Moreover, thanks to the trigger command, the active periods of the sensor can be synchronized with the BXs of the accelerator in order to avoid missing any real hits. As a result, the SNR and Dynamic Range (DR) of the sensor can be increased by several orders of magnitude [10]. Nevertheless, it is difficult to increase the fill-factor with standard CMOS technologies. In this work, the 3D vertical integration of a two-layer chip is proposed as a solution to overcome the fill-factor limitation of GAPD devices.

3. 3D vertical integration with the Chartered 130 nm/Tezzaron process

The 3D GAPD detector described in this work is in a 130 nm CMOS process fabricated by Chartered Semiconductor and vertically integrated by Tezzaron, available in MPW (Multi-Project Wafer) runs. 3D-ICs manufactured in the Chartered 130 nm/Tezzaron 3D process, typically consist of two layers of logic dies fabricated by Chartered Semiconductor and two or (if possible) three layers of memories supplied by Tezzaron. However, it is also possible to build a two-layer stack with no memories attached (no-DRAM option), which is the case with this work.

In this option, the 3D-ICs are manufactured by independently fabricating the 2D logic dies corresponding to the two different layers (called tiers) on separate wafers. Then, the two wafers are stacked face-to-face, bonded together, thinned, and finally diced [11]. During the bonding process, the top of the WTOP wafer is flipped onto the top of the WBOTTOM wafer in a right-to-left orientation. Hence, the two logic dies are connected face-to-face (wafer-to-wafer). The connection between tiers for relaying signals is made through Metal 6, which is the highest metal of the technology process. This 3D process also uses via-first Through Silicon Vias (TSVs) for connection between the logic circuitry and the I/O bond pads, which are placed on the back of the WTOP tier. TSVs are also used to control thinning. As a consequence, it is necessary to maintain a minimum TSV density throughout both tiers, which forces the utilization of dummy TSVs. The recommended TSV pitch is 100 μm. TSVs are arranged in an hexagonal shape and covered with Metal 1. After bonding, the WTOP wafer is thinned down to about 12 μm until the bottom ends of the TSVs are exposed. The WBOTTOM wafer can also be thinned, however this incurs additional costs. Back metal for bonding pads is applied to the thinned WTOP wafer. When all this processing is done, the wafer stack is diced.

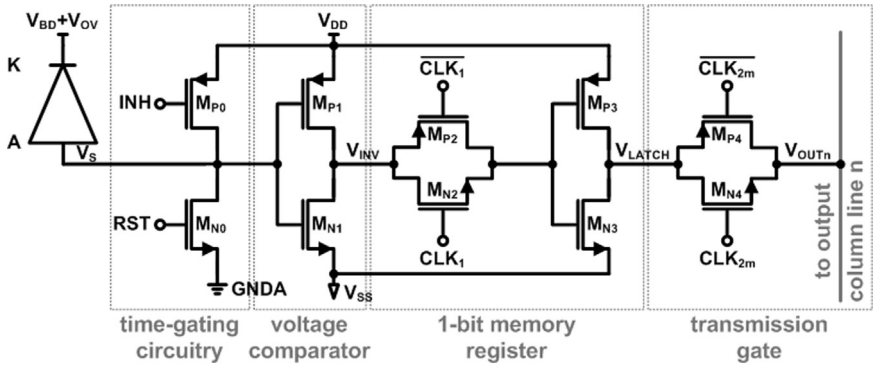

Fig. 1. Schematic diagram of the time-gated digital pixel.

4. Array design

The proposed 3D GAPD detector consists of an array of 48 × 48 pixels. As shown in Fig. 1, each pixel is comprised of a GAPD, active inhibition (M_{P0}) and active reset (M_{N0}) switches to perform the time-gated operation, and a readout circuit based on a voltage comparator (M_{P1}–M_{N1}), a 1-bit memory register (M_{P2}–M_{N2}–M_{P3}–M_{N3}) and a transmission-gate (M_{P4}–M_{N4}). The number of transistors per pixel is the lowest possible for a time-gated GAPD array and the size of the transistors is the minimum allowed by the technology.

4.1. Sensor and mode of operation

The sensor diode is implemented by means of a p^+ anode within an n-well cathode. The junction is surrounded by a low doped p-well guard ring to achieve a planar multiplication region and hence avoid the premature edge breakdown. Moreover, a buried n-type isolation layer or deep n-well, available in this technology, is used to achieve full isolation of the p-well guard ring from the p-substrate. This layer is also used to prevent the punch-trough of the p-well to the p-substrate. The n-well cathode is biased at a positive $V_{BD}+V_{OV}$, V_{OV} being the reverse bias over-voltage, in order to operate the Geiger-mode. The avalanches are sensed at the p^+ anode due to its lower intrinsic capacitance to ground. The electronics is located within the p-substrate, which is connected to ground (V_{SS}).

The Chartered 130 nm technology requires the utilization of the STI. This isolation layer is etched in all regions not covered by a heavy implant or polysilicon to make sure that it surrounds all the p^+ and n^+ implantations for an isolation improvement. In order to avoid contact between the STI and the depletion region of the GAPD, and thereby have an acceptable dark count rate, a poly-silicon gate (polysilicon, oxide, diffusion and p^+ layers) is drawn around the p^+ anode. The GAPD cross-section is shown in Fig. 2. The polysilicon gate is biased at the same potential as the p^+ layer [6]. The separation between two neighboring GAPDs is filled with n-well, which is short-circuited to the n-well cathode through the deep n-well layer. Ohmic contacts for bias to $V_{BD}+V_{OV}$ are placed on top of the n-well separation to ensure a robust cathode biasing throughout all the pixels of the array. The introduction of the p-well guard ring, together with the polysilicon for an STI-free GAPD and the cathode ohmic contacts, generates a minimum separation between two neighboring GAPDs of 2.24 μm.

The sensor time-gating is controlled by means of two external signals (RST and INH) implemented through MOS transistors (M_{N0}–M_{P0}). When the RST signal is set high, the sensor is activated (gated 'on') by increasing its bias up to $V_{BD}+V_{OV}$. The gated 'on' period is started an instant before the BX takes place. In contrast, when the INH signal is set low, the polarization of the sensor is reduced to $V_{BD}+V_{OV}-V_{DD}$, with $V_{OV} < V_{DD}$ and $V_{DD}=1.2$ V in this technology. The sensor is then deactivated (gated 'off') and it remains in this state until the next rising of the RST signal. The gated 'off' period is coincident with the inter-BX. When an avalanche is triggered, the self-sustained current that flows through the junction discharges the sensor capacitance and charges the parasitic capacitance of the sensing node (V_S in Fig. 1). As a result, an analog voltage pulse is generated in the V_S node in picoseconds. When the voltage pulse reaches an amplitude equal to V_{OV}, the polarization of the sensor drops down to V_{BD} and the avalanche quenches.

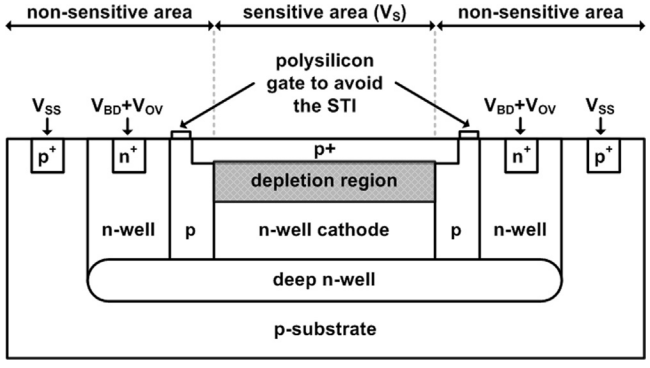

Fig. 2. Cross-section of one GAPD sensor with the Chartered 130 nm/Tezzaron 3D process.

4.2. Readout circuit

The analog voltage pulse generated by the avalanche is detected and transformed into a digital signal by the voltage comparator (M_{P1}–M_{N1}). The comparator is based on a simple CMOS inverter, which was designed to have the threshold voltage of $V_{DD}/2$ and a propagation delay of 150 ps. As a solution to allow the detection of low V_{OV} below $V_{DD}/2$, the sensor and the readout

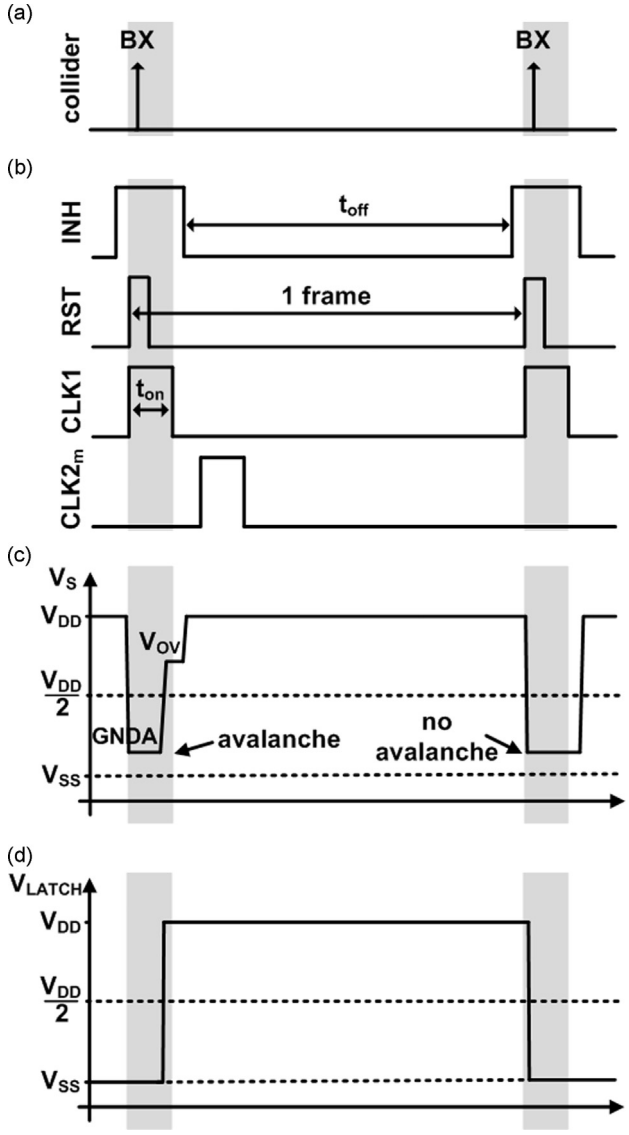

Fig. 3. Temporal diagram for the BXs at the particle accelerator (a), electronics waveforms for the proposed mode of operation (b), response of the GAPD sensing node (c) and response of node V_{LATCH} (d).

electronics have different ground nodes (GNDA for the sensor and V_{SS} for the readout). The output of the inverter (V_{INV}) is fed into a 1-bit dynamic latch (M_{P2}–M_{N2}–M_{P3}–M_{N3}), which performs a 2-stage synchronous operation controlled by the CLK1 signal (M_{P2}–M_{N2}). The CLK1 signal is set high at the beginning of the gated 'on' period to sample V_S. When the CLK1 signal is set low (a few nanoseconds after the BX), the current value of node V_S is stored in node V_{LATCH} ('0' for no avalanche, '1' for avalanche) during the gated 'off' period. A transmission-gate (M_{P4}–M_{N4}) governed by the $CLK2_m$ signal, with $m=[1,48]$, is used to sequentially read the 48 rows of the GAPD array during the gated 'off' intervals. Thereby, the whole array can be read in less than 200 ns during the inter-BXs. The signals RST, INH and CLK1 reach all the pixels simultaneously. Yet, whereas the $CLK2_m$ signal is common to all the pixels of the same row, an output column line (V_{OUTn}) is common to all the pixels of the same column. The waveforms for the pixel operation described here and the pixel response are depicted in Fig. 3.

4.3. Array architecture and layout

Prior to the final layout of the 48×48 GAPD array, the maximum achievable fill-factor with the Chartered 130 nm/Tezzaron 3D process and the readout electronics proposed before was investigated with several array architectures (drawn in Fig. 4). A first scheme implements the sensors in one tier and the readout electronics in the other one (Fig. 4a). With a sensor area of $18\ \mu m \times 18\ \mu m$, a fill-factor of 66% is achieved with this configuration. The other structures studied benefit from the two-layer vertical stacking to overlap the non-sensitive areas of one tier with the sensitive areas of the other tier. In addition, different sensor areas were used to maximize the overlap between tiers. Thus, the second approach is based on clusters of four pixels and two sensor areas of $18\ \mu m \times 18\ \mu m$ and $30\ \mu m \times 30\ \mu m$. Three $18\ \mu m \times 18\ \mu m$ sensors together with the readout electronics of the four sensors are placed in one tier, whereas the $30\ \mu m \times 30\ \mu m$ GAPD is strategically placed in the other tier to overlap the readout electronics and most of the non-sensitive areas of the $18\ \mu m \times 18\ \mu m$ sensors (Fig. 4b). This approach generates a 92% fill-factor. A similar idea is implemented in the structure depicted in Fig. 4c, however clusters of five pixels are used here. A 96% fill-factor is achieved in this structure with two sensor areas of $8\ \mu m \times 8\ \mu m$ and $20\ \mu m \times 30\ \mu m$, however the solution is bizarre and risky. The last strategy explored is shown in Fig. 4d. It uses clusters of four pixels and both tiers have sensors and readout electronics. The sensor areas are $18\ \mu m \times 15\ \mu m$ and $23\ \mu m \times 20\ \mu m$. An 85% fill-factor is achieved in this case.

Out of the four array architectures proposed, the structures represented in Fig. 4a and b were selected to be implemented in the final layout and study their performance. The first structure was chosen for its simplicity and the second one because it provides the maximum fill-factor with the lowest risks. Thus, the 48×48 GAPD array is composed of two sub-arrays of 48×24 pixels each. The GAPDs of the first array plus the $18\ \mu m \times 18\ \mu m$ GAPDs and the readout electronics of the second array are placed in the WTOP tier. In contrast, the readout electronics of the first array plus the $30\ \mu m \times 30\ \mu m$ GAPDs of the second array are implemented in the WBOTTOM tier. Hence, there is interconnection between tiers from the node V_S to the readout circuits in both sub-arrays. The sensors were distributed in the two tiers bearing in mind to place the maximum possible number of sensors in the WTOP tier, which is 12 μm thick. The TSVs are located together with the readout circuits when possible. In those regions were sensors only are present, the TSVs are placed in between two sensors while maintaining the minimum recommended pitch. The 48×48 GAPD detector is sequentially read by rows during the gated 'off' periods. As the chip is equipped with only 6 output pads so as not to increase the occupied area, 6 8-bit shift-registers are placed between the output column lines and the output pads. WrEn and EnOut are the protocol signals to enable the readout of

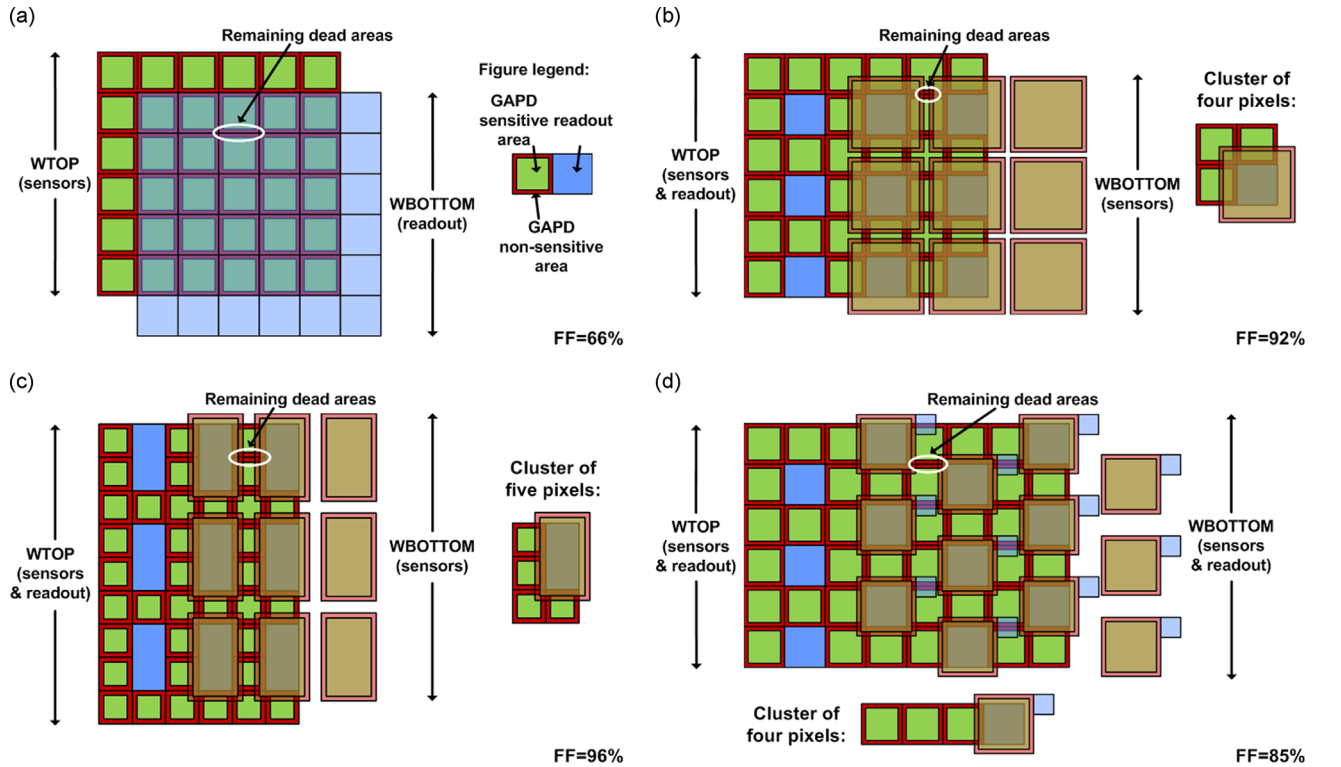

Fig. 4. Analysis of the achievable fill-factor with several array architectures of GAPDs and considering the technology design rules of the Chartered 130 nm/Tezzaron 3D process. The sensors and readout electronics are not drawn to scale.

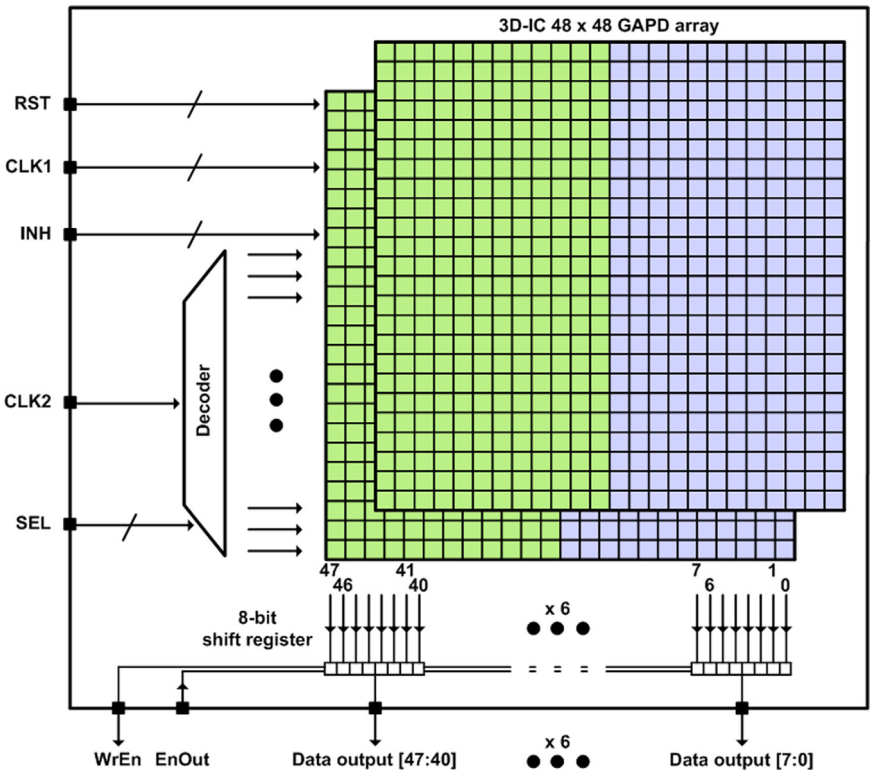

Fig. 5. Functional diagram of the detector readout.

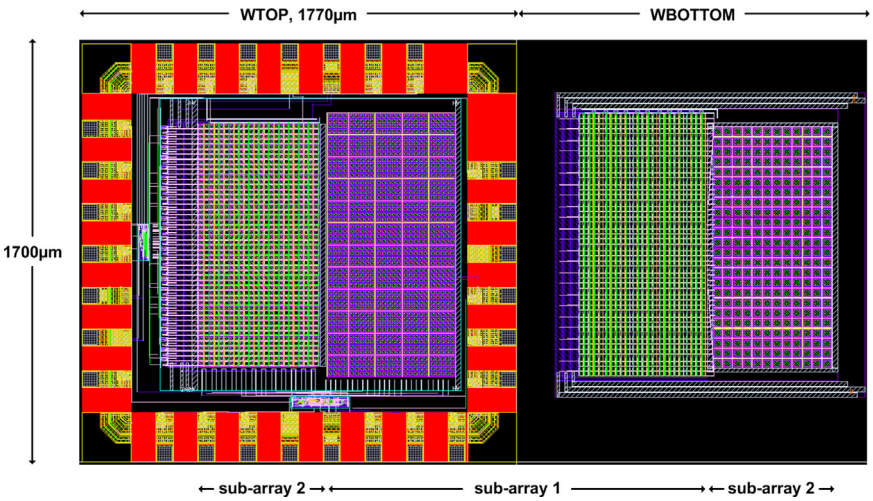

Fig. 6. Layout of the two GAPD sub-arrays designed with the Chartered 130 nm/Tezzaron 3D process.

the detector. The chip also contains a decoder to distribute the $CLK2_m$ control signals (see Fig. 5 for a functional diagram). The final layout of both arrays is shown in Fig. 6.

5. Conclusion

A 3D pixel detector based on low-voltage time-gated GAPDs for future linear colliders has been designed with the Chartered 130 nm/Tezzaron 3D process, which allows the vertical stacking of two layers and thus improve the detector fill-factor. Several array architectures were studied to determine the maximum achievable fill-factor with the Chartered 130 nm/Tezzaron 3D process and a time-gated readout circuit of minimum area. The final design is composed of two sub-arrays of 48×24 pixels each. The first sub-array, where the sensor and the readout electronics are split into the two layers, presents a 66% fill-factor. In contrast, in the second sub-array the sensors are implemented in both tiers to overlap the non-sensitive areas due to the sensors and the readout electronics. In this case, a 92% fill-factor is achieved. It has been proven that the typical low fill-factor of GAPD detectors can be increased up to values close to 100% with 3D technologies.

Acknowledgments

This work has received fundings from the National Program for Particle Physics through the project “Desarrollo de nuevos detectores para los Futuros colisionadores en Física de Partículas”, coded FPA2010-21549-C04-01, and the European Commission within the

Framework Programme 7 Capacities (Grant Agreement 262025) through the project "Advanced European Infrastructures for Detectors and Accelerators—AIDA".

References

[1] M. Demarteau, Y. Arai, H.G. Moser, V. Re, Developments of novel vertically integrated pixel sensors in the high energy phyisics community, in: Proceedings of IEEE Internationl Conference on 3D System Integration, San Francisco, USA (2009) pp. 1–7.
[2] T. Abe, et al., ILD Concept Group—Linear Collider Collaboration, arXiv: 1006.3396[hep-ex].
[3] F. Zappa, S. Tisa, A. Tosi, S. Cova, Sensors and Actuators A 140 (2007) 103.
[4] S. Tisa, F. Zappa, A. Tosi, S. Cova, Sensors and Actuators A 140 (2007) 113.
[5] A. Arbat, Towards a forward tracker detector based on Geiger mode avalanche photodiodes for future linear colliders, Ph.D. Dissertation, University of Barcelona, (2010).
[6] C. Niclass, M. Gersbach, R. Henderson, L. Grant, E. Charbon, IEEE J. Sel. Top. Quantum Electron. 13 (2007) 863.
[7] R.K. Henderson, J. Richardson, L. Grant, Reduction of band-to-band tunneling in deep-submicron CMOS single photon avalanche photodiodes, in: Proceedings of the International Image Sensor Workshop, Bergen, Norway, (2009).
[8] J.A. Richardson, L.A. Grant, R.H. Henderson, IEEE Photonics Technology Letters 21 (2009) 1020.
[9] E. Vilella, A. Diéguez, Sensors and Actuators A 186 (2012) 163.
[10] E. Vilella, A. Comerma, O. Alonso, A. Diéguez, Electronics Letters 47 (2011) 395.
[11] G.W. Deptuch, M. Demarteau, J.R. Hoff, R. Lipton, A. Shenai, M. Trimpl, et al., IEEE Transactions on Nuclear Science 57 (2010) 2178.