MAPS with pixel level sparsified readout: from standard CMOS to vertical integration

Luigi Gaioni[*a,c], **Alessia Manazza**[a,c], **Massimo Manghisoni**[b,c], **Lodovico Ratti**[a,c],
Valerio Re[b,c] **and Gianluca Traversi**[b,c]

[a]*Università degli Studi di Pavia*
Via Ferrata 1, 27100 Pavia, Italy.

[b]*Università degli Studi di Bergamo*
Via Marconi 5, 24044 Dalmine (BG), Italy.

[c]*INFN Pavia*
Via Bassi 6, 27100 Pavia, Italy.
E-mail: luigi.gaioni@unipv.it

In recent years, scaled CMOS technologies have led to an increased cell functional density and a better spatial resolution in monolithic active pixel sensors (MAPS) for vertexing applications. Advanced MAPS may also take advantage of CMOS process features such as the deep n-well (DNW) structure, which can be used as the charge collecting element of the detector. A new kind of DNW-MAPS, namely SDR1, which exploits the capabilities of vertical integration (3D) processes, is presented and discussed in this paper. 3D processes may offer significant advantages in terms of detection efficiency, pixel cell size and immunity from cross-talk, therefore complying with the severe constraints set by future HEP experiments. This work includes a description of the analog and digital circuits integrated in the SDR1 chip, together with circuit simulations. Device simulation results concerning detection efficiency will be also discussed.

---

[*]Speaker.

---

1. Introduction

CMOS monolithic active pixel sensors (MAPS) designed in standard VLSI CMOS technology have recently been proposed as compact pixel detectors for vertexing/tracking applications at the next generation colliders like the International Linear Collider (ILC) and SuperB-Factory [1], [2]. MAPS are already extensively used in visible light applications: with respect to other competing imaging technologies, MAPS sensors have several potential advantages in terms of cost, power, noise and functionality.
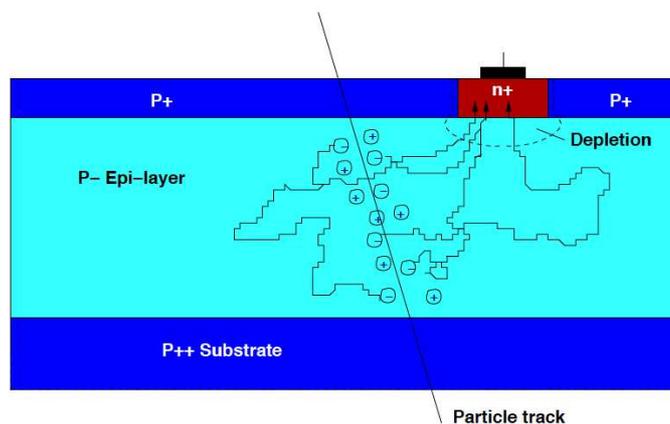

Figure 1: simplified cross-section of a CMOS pixel with epitaxial layer.

The cross-sectional view of Figure 1 shows the basic principles underlying CMOS sensor operation. In most modern CMOS processes, n- and p-wells are fabricated on top of a thin p-doped epitaxial layer, with resistivity of the order of 1-10 Ω·cm. A p-n junction exists between the n-well and the p-epilayer and can be used as the collecting element for the charge released by the radiation. In the extremely simple configuration of typical MAPS, only three transistors (3T) are integrated in the pixel cell, thereby providing a fine spatial resolution. The use of large area electrodes is strongly discouraged in 3T-MAPS design: indeed, increased capacitance would unacceptably degrade the noise figure and the charge sensitivity at the same time. Moreover, use of PMOS devices in the design of the front-end electronics is avoided as the n-well they are integrated in might subtract charge from the collecting electrode, leading to potentially serious efficiency loss. The lack of complementary devices represents a significant limitation to the design of stages with satisfactory properties, thereby restricting the set of available readout solutions. An innovative design solution for MAPS, leading to the so called deep n-well (DNW) MAPS, was proposed a few years ago [3] in order to deal with the large amount of data produced in the readout of large matrices of pixels. This solution relies upon the use of a deep n-well/p-substrate junction, provided by triple-well CMOS technologies, as the collecting element. DNW-MAPS allow designers to implement more complex readout circuits, taking advantage of fully CMOS architectures. However, although this approach is attractive because of the advantages of a full CMOS implementation, the overall detection efficiency is adversely affected by the use of PMOS devices. In order to address this issue and to further improve such DNW MAPS properties as spatial resolution, new technology options have been recently considered: in particular, the interest of several research groups has shifted towards

vertical integration processes (usually referred to as 3D) [4]. 3D processes, by stacking two or more layers one on top of the other, make it possible to improve DNW MAPS collection efficiency (see section 5) since PMOS can be placed in a different layer with respect to the sensor. Moreover, 3D processes may enhance functional density of the elementary cell and provide physical separation of the analog front-end from the digital blocks, thereby preventing undesirable cross-talk phenomena. The 3DIC Consortium [5] was promoted by Fermilab to explore the 3D integration. In 3DIC, as a first step, European and US institutions submitted a multi-project wafer run in the Tezzaron/Chartered fabrication process described in section 3. The SDR1 chip, designed in the 3DIC framework, is the first generation of deep n-well CMOS monolithic sensors in a 130 nm vertical integration technology. The main design features and the expected performance of the chip will be presented in section 4.

2. 2D deep n-well MAPS

The DNW MAPS are based on the same sensing principle as standard MAPS, where radiation induced charge carriers diffuse and are collected by n-type electrodes. In a DNW MAPS sensor, whose structure is shown in Figure 2, an n-well with a deep junction acts as the collecting element for the charge released in the substrate.
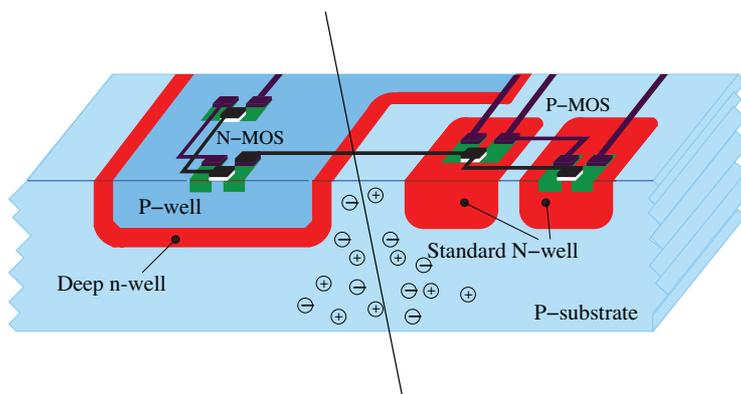
Figure 2: simplified structure of a DNW-MAPS.

The collected charge is read out by a classical optimum chain for capacitive detectors, including a fully CMOS charge preamplifier whose closed loop gain is independent of the detector capacitance. The deep n-well, which in modern, triple-well CMOS processes is used to shield NMOS devices from substrate coupled noise in mixed signal circuits, may host n-channel devices, thus relaxing the constraints set by the readout circuits on the sensor area and geometry. Moreover, designers may implement more complex readout circuits, taking advantage of fully CMOS architectures, laying out large area DNW sensors. Digital blocks involved in data sparsification and time stamping are also integrated in the elementary pixel cell.

The so called APSEL series is the first generation of DNW MAPS with on-pixel data sparsification and time stamping, successfully tested in a beam for the first time in September 2008 [6]. The APSEL series features a continuous readout architecture suitable for application to the SuperB Layer 0. Another DNW MAPS prototype is the SDR0 chip [7], including a number of different
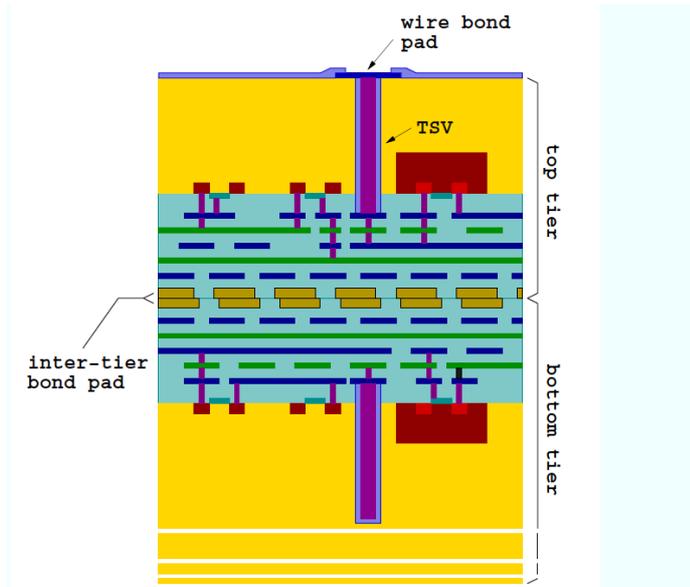
Figure 3: cross-sectional view of the double-layer 3D process provided by Tezzaron Semiconductor

structures, among which a 16x16 DNW MAPS matrix. SDR0 was designed in view of vertexing applications to the ILC. Its operation is based on the ILC beam structure, featuring two different processing phases: a detection phase, corresponding to the bunch train period, and a readout phase, corresponding to the intertrain period, during which the matrix is read out. SDR0 analog processor consists of a shaperless version of a classical readout chain for capacitive detectors, combined with a threshold discriminator which enables binary readout of the pixel cell. The digital front-end enables single-hit storage and includes a 5-bit time stamp register and sparsification blocks, based on a token passing architecture. Two major issues, namely collection efficiency and storing capability, affect the overall performance of 2D DNW-MAPS. Collection efficiency may be inadequate because of the fully CMOS architecture requiring a non negligible amount of n-well area for the integration of PMOSFETs. Moreover, SDR0 and APSEL prototypes enable only single-hit detection: improvements in detection efficiency, which depends on the occupancy and on the expected probability for multiple hits, may result as a direct consequence of enhancing the storing capability. Vertical integration processes can be effectively used to solve, or at least to significantly alleviate, the above mentioned issues.

3. 3D processes

3D technologies rely on layering tiers of active circuitry interconnected to each other. A vertically interconnected wafer, in addition to having increased overall circuit density, reduces the overall length of the device interconnections, increasing the speed by reducing resistance, inductance and parasitic capacitance. Power consumption is also decreased due to the reduced wire length and to the smaller capacitance. In addition, the layers may be fabricated in different technologies, each optimized for a specific function. A variety of methods for the 3D integration have been demonstrated [8].

Figure 3, in particular, refers to the main features of the process provided by Tezzaron Semiconductor [9], which can be used to vertically integrate two or more 130 nm CMOS layers specifically processed by Chartered Semiconductor. In the Tezzaron/Chartered process, wafers are face-to-face bonded by means of thermo-compression techniques. Bond pads on each wafer are laid out on the copper top metal layer and provide both the mechanical connection of the wafers and the electrical contacts between devices integrated in the two layers. After bonding, the top wafer is thinned to the bottom of through silicon vias (TSV). These are electrically isolated metal via penetrating the silicon substrate which makes the connection to the buried circuits possible. The features of the Tezzaron/Chartered process have been exploited in the design of the SDR1 chip, which inherits the intertrain sparsified readout architecture of the SDR0 monolithic sensor, whose characteristics will be taken into account in this work for comparison purposes. Although the digital readout architecture of the SDR1 chip has been conceived for vertexing applications at the ILC, this concept is also compatible with other applications, such as X-ray imaging at XFEL [10].

4. The SDR1 chip

SDR1 features two vertically integrated layers each fabricated in a 130 nm CMOS process, containing the analog and the digital front-end respectively. SDR1 consists of a 240x256 DNW-MAPS matrix with a pixel pitch of 20 μm, token passing binary readout architecture and the capability for storing two hits and the relevant 5-bit time stamps.
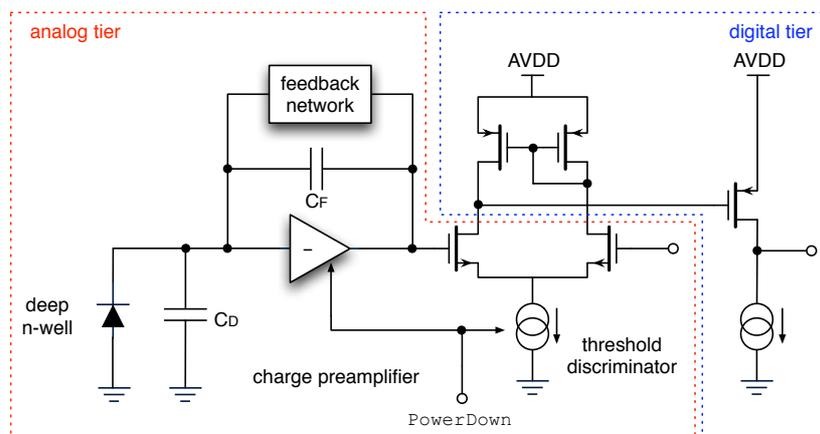
Figure 4: schematic diagram of the analog front-end circuit processing the signal from the deep n-well sensor.

4.1 Analog front-end

In the SDR1 design, the pixel-level front-end processor, whose schematic diagram is shown in Figure 4, includes a charge sensitive amplifier and a threshold discriminator, which has been only partially integrated in the bottom (analog) tier, as highlighted in the figure. Large area PMOS devices (needed to minimize threshold dispersion) belonging to the threshold discriminator were laid out on the top (digital) tier of the chip, thus remarkably reducing the competitive n-well areas integrated in the sensor layer. The charge preamplifier input NMOS device (whose dimensions
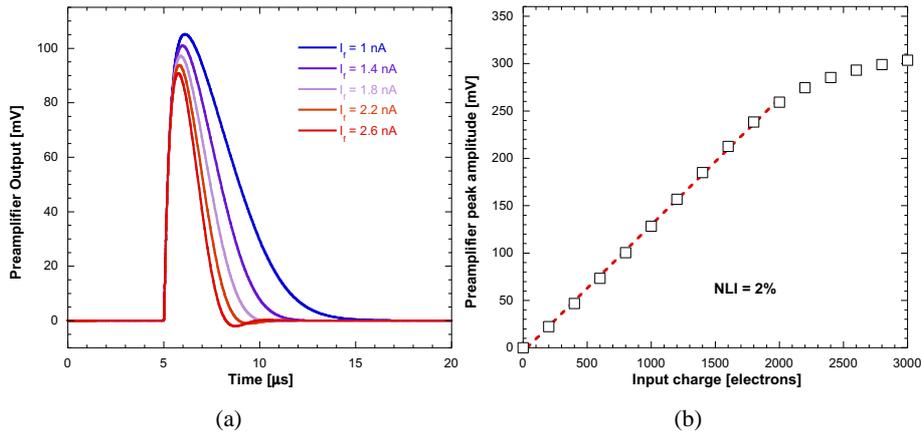
Figure 5: preamplifier output in response to an 800 electrons pulse for different values of the current I_F biasing the feedback current mirror (a); preamplifier output peak value as a function of the injected charge (b).

were chosen based on criteria for optimum detection efficiency in multichannel systems with binary readout under noise hit rate constraints [11]) features a W/L=20/0.18 and a drain current of 1.4 μA. Charge restoration in the preamplifier feedback network is obtained through a PMOS current mirror stage, providing a linear discharge of the metal-oxide-metal capacitor C_F (about 1 fF). Charge sensitivity in the preamplifier is designed to be about 800 mV/fC. In the design of the circuit, the high frequency noise contribution has been reduced by purposely limiting the preamplifier bandwidth. Figure 5 (a) shows the simulated preamplifier output in response to an 800 electrons pulse, for different values of the current I_F biasing the feedback current mirror: the peak amplitude is close to 100 mV. The charge sensitivity is defined as the slope of the interpolating straight line of Figure 5 (b), showing the preamplifier output peak value as a function of the injected charge. An integral non-linearity of about 2% has been obtained over an input dynamic range of 2000 electrons.

For a detector capacitance C_D of 200 fF an equivalent noise charge (ENC) of 35 e$^-_{rms}$ was obtained from circuit simulations. An overall input referred threshold dispersion of 36 e$^-_{rms}$ was computed from Monte-Carlo simulations: the main contributions arise from the preamplifier input device and from NMOS and PMOS pairs in the discriminator. Power consumption of the elementary cell is about 5 μW. This is compatible with the power constraints set by the ILC vertex detector specifications, which require power dissipation less than 10 mW/cm^2 with power cycling operation (1% duty-cycle), controlled by the `PowerDown` command shown in Figure 4.

4.2 Digital front-end

Beside the PMOS load and the gain stage of the threshold discriminator, the top layer of the SDR1 elementary cell also includes digital blocks providing double-hit storing and time stamping, data sparsification and pixel masking. The digital front-end of the SDR1 chip is shown in Figure 6. The 5 time stamp bits are fed to the registers by a Gray counter located in the chip periphery and allow the bunch train interval to be subdivided into 32 time slots. During the bunch train period, when a pixel is hit for the first time, its discriminator fires and sets the first-hit latch (FFSRK). The
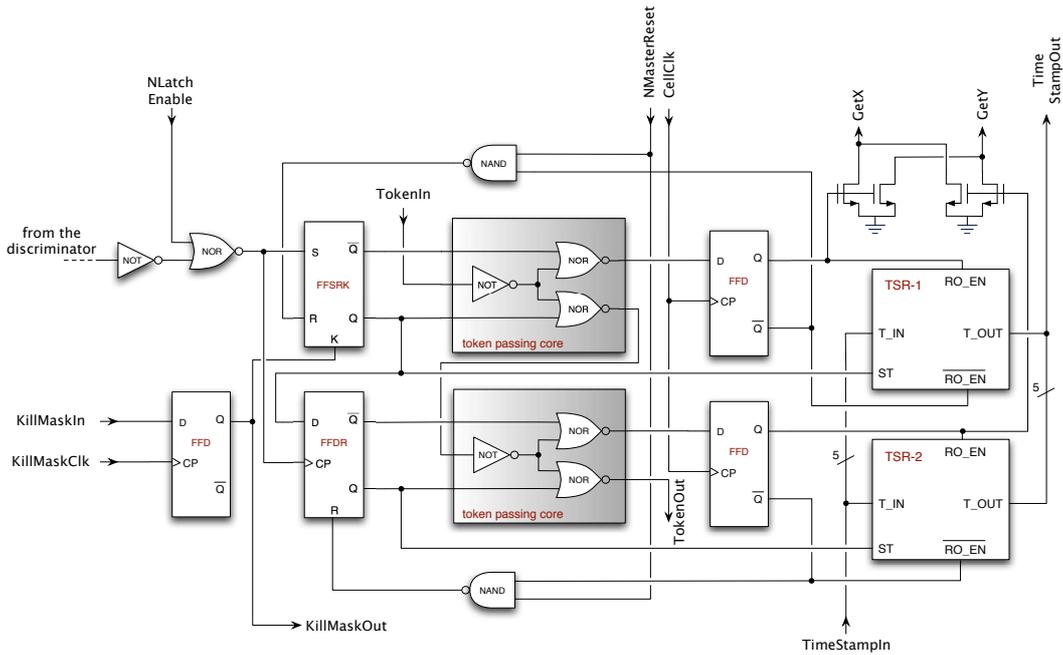
Figure 6: block diagram of the digital front-end electronics integrated in the elementary cell of SDR1.

latch output is used to freeze the content of the time stamp register (through the ST input in the time stamp register), therefore providing the arrival time of the hit with about 30 μs resolution in the case of the ILC beam structure. If a second hit occurs, the output of the flip-flop FFDR is set and the content of the second time stamp register gets frozen. At the end of the bunch train period, when the readout phase begins, the NLatchEnable signal is switched on in order to prevent the FFSRK and FFDR flip-flop from accidentally firing in those pixels which were not hit during the detection period.

4.3 Digital back-end

The periphery of the 240x256 matrix integrated in the SDR1 chip includes a serializer, X and Y coordinate registers, time stamp buffers and a Gray counter used to provide the current time stamp value to the time stamp registers in all the cells during the detection phase, as schematically shown in Figure 7.

At the end of the detection phase, a token is launched through the MAPS matrix by setting the FirstTokenIn signal. Each hit pixel, after receiving the token (TokenIn signal), gets hold of the column and row buses (GetX and GetY signals are pulled down) at the next cell clock (CellCLK) rising edge, and releases position and time stamp information (by acting on the ROEN input of the time stamp register) within a cell clock period. Within a very short time interval, the first-hit latch is reset (this is possible if the NMasterReset signal is high, which is the common operating condition in both detection and readout phases). Almost immediately, the token scans the second time stamp register and if there is not a second detected hit, it is released and sent out (TokenOut signal) to the next hit pixel or to the matrix output (LastTokenOut signal). The token scans two time stamp registers per cell: two cell clock periods are thus needed to perform the
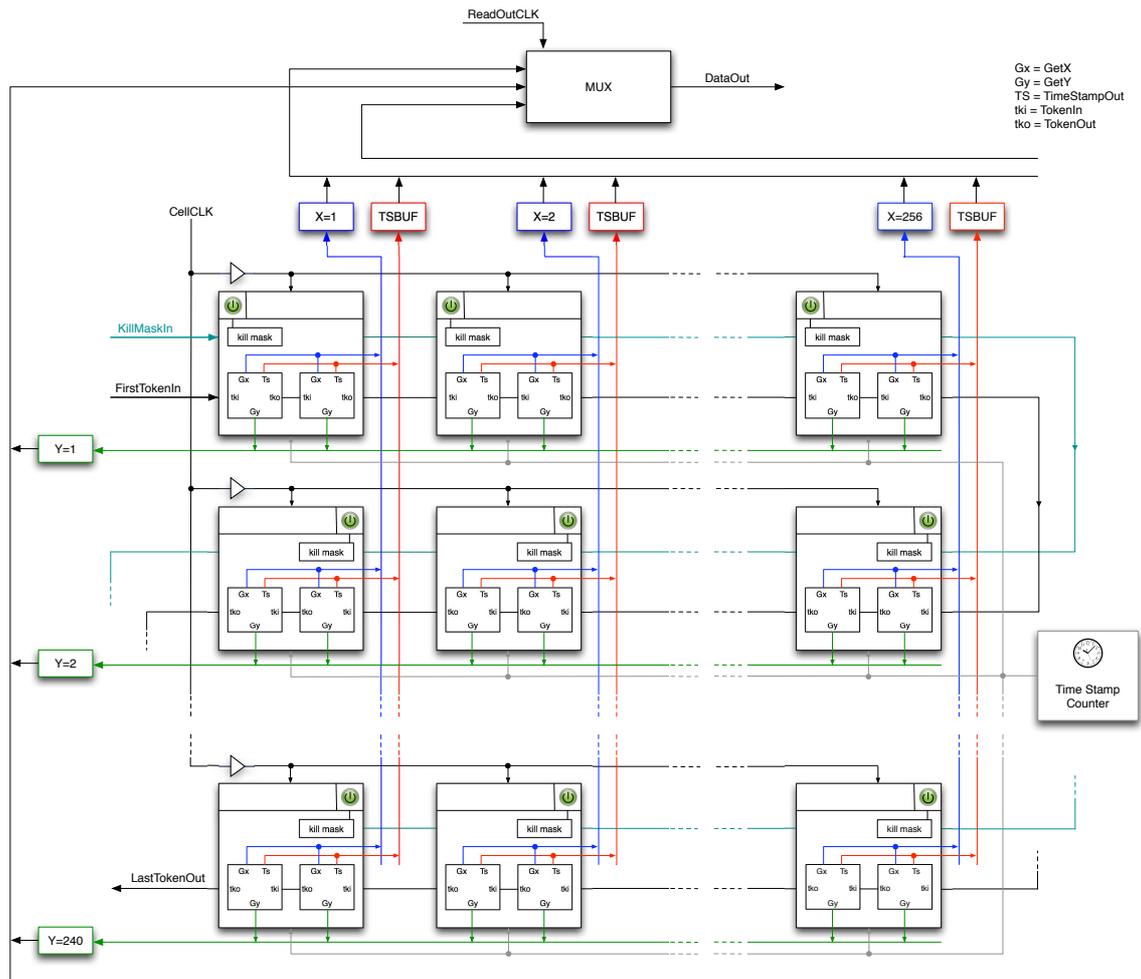
Figure 7: digital readout architecture of the SDR1 MAPS detector.

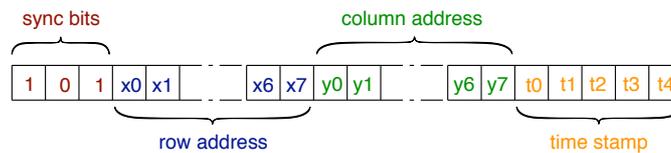readout of a pixel which has detected two hits.


Figure 8: output word data format.

Output data are transmitted off the chip by means of a serializer controlled by the readout clock (`ReadOutCLK`), whose frequency is an integer multiple of the cell clock. Output data are 24 bit long words with the format shown in Figure 8 (3 sync bits, 8 bits for the X and the Y coordinates and 5 bits for the time stamp). The achievable bit rate is of the order of 100 Mbit/s. The kill mask block consists in a shift register which scans all the cells of the matrix: the kill mask can be loaded in the matrix during the system initialization to disable noisy pixels.

5. 3D sensor detection efficiency

In order to evaluate the collection efficiency, Monte Carlo simulations have been performed on clusters of 3x3 DNW MAPS featuring the layout of SDR0 and SDR1 cells, shown in Figure 9.
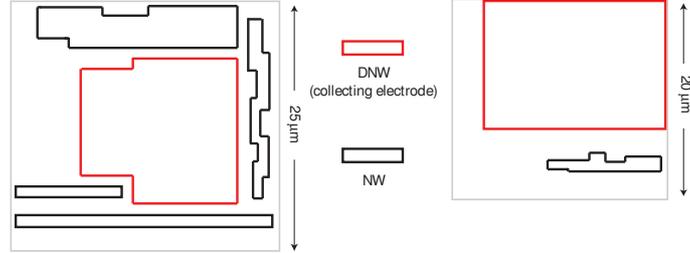
Figure 9: layout of the DNW sensor and of the n-wells in the SDR0 (left) and SDR1(right) MAPS detectors.
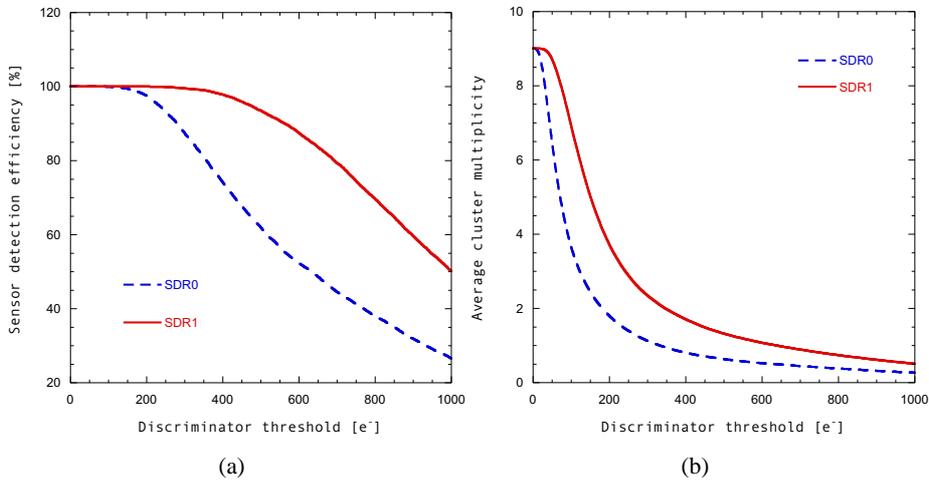
Figure 10: sensor detection efficiency (a) and average cluster multiplicity (b) as a function of the discriminator threshold in the case of the SDR0 and of the SDR1 chip.

These simulations are based on a random walk algorithm modeling the carrier motion in the undepleted substrate of monolithic pixel detectors [12] and consist of a set of 10000 particles randomly hitting the central pixel of the cluster. The resulting sensor detection efficiency is displayed in Figure 10 (a) as a function of the discriminator threshold. Sensor detection efficiency is still well over 99% at a discriminator threshold of 300 electrons in the SDR1 MAPS. Figure 10 (b) shows the average cluster size as a function of the discriminator threshold. It is worth noticing that the SDR0 cluster size is always smaller than in the SDR1 case. This may be due to the larger area covered by standard n-wells in the SDR0 chip, which in turn reduces the available charge for the main DNW electrode, therefore reducing, at each threshold level, the average number of pixels over threshold.

6. Conclusion

In this paper, the main features of the first generation of deep n-well CMOS monolithic sensors

in a 130 nm vertical integration technology has been discussed. The use of a double-layer process in the development of the proposed DNW MAPS, the SDR1 chip, may address the main issues related to planar 130 nm CMOS technology. In particular, the 3D approach can cope with the low charge collection efficiency (due to the non negligible area covered by charge-stealing n-wells) and with the low detection efficiency (due to limitations in hit-storing capability) exhibited by the 2D detectors. Moreover, beside the higher functional density, better point resolution and reduction of cross-talk phenomena between digital blocks and the analog section can be achieved by means of the 3D technology. Several test structures have been integrated in the SDR1 prototype to evaluate the suitability of the Tezzaron/Chartered technology for the fabrication of DNW MAPS. In particular, a 240x256 MAPS matrix with sparse readout and time stamping capabilities should be ready together with the needed setup for a test beam in the third quarter of 2010. SDR1 also includes smaller matrices among which a 16x16 MAPS matrix chiefly conceived for the test of the sparsified readout architecture and a 8x8 MAPS matrix into which the outputs of the charge preamplifiers in the 64 pixels of the matrix can be accessed with the purpose of testing the cluster properties of the sensor by means of laser sources.

Acknowledgments

References

[1] G. Deptuch, *"Design and testing of monolithic active pixel sensors for charged particle tracking"*, IEEE Trans. Nucl. Sci., vol. 49, no. 2, pp. 601-610, Apr. 2002.

[2] M. Barbero et al., *"Development of a B-Factory Monolithic Active Pixel Detector -The Continuous Acquisition Pixel Prototypes"*, IEEE Trans. Nucl. Sci., vol. 52, no. 4, pp. 1187-1191, Aug. 2005.

[3] L. Ratti et al., *"Monolithic pixel detectors in a 0.13 µm CMOS technology with sensor level continuous time charge amplification and shaping"*, Nucl. Instrum. Methods, vol. A568, pp. 159-166, 2006.

[4] R. Yarema et al., "3D IC Pixel Electronics, the next challenge", Proceedings of the Topical Workshop on Electronics and Particle Physics, Naxos, Greece, September 15-19, 2008, pp.183-187.

[5] The 3DIC Consortium, http://3dic.fnal.gov

[6] G. Rizzo et al, *"Development of deep N-well MAPS in a 130nm CMOS technology and beam test results on a 4k-pixel matrix with digital sparsified readout"*, IEEE Nuclear Science Symposium Conference Record, Oct. 2008.

[7] G. Traversi, M. Manghisoni, L. Ratti, V. Re, V. Speziali, *"CMOS MAPS with pixel level sparsification and time stamping capabilities for applications at the ILC"*, Nucl. Instrum. Methods, vol. A581, pp. 291-294, 2007.

[8] *Handbook of 3D Integration: Technology and Applications of 3D Integrated Circuits.* Edited by P. Garrou, C. Bower and P. Ramm. Wiley-VCH, Weinheim, 2008.

[9] R.S. Patti, *"Three-Dimensional Integrated Circuits and the Future of System-on-Chip Design"*, Proceedings of the IEEE, vol. 94, No. 6, June 2006.

[10] The European X-Ray Laser Project XFEL, http://xfel.desy.de

[11] L. Ratti, M. Manghisoni, V. Re, G. Traversi, *"Design of time invariant analog front-end circuits for deep n-well CMOS MAPS"*, IEEE Trans. Nucl. Sci., vol. 56, no. 4, pp. 2360-2373, Aug. 2009.

[12] E. Pozzati et al., *"MAPS in 130 nm triple well CMOS technology for HEP applications"*, Proceedings of the Topical Workshop on Electronics and Particle Physics, Prague, Czech Republic, 3-7 September 2007, pp. 492-497.