MALTA: a CMOS pixel sensor with asynchronous readout for the ATLAS High-Luminosity upgrade

Ivan Berdalovic, Lluis Simon Argemi, Roberto Cardella, Florian Dachs, Valerio Dao, Leyre Flores Sanz de Acedo, Tomasz Hemperek, Bojan Hiti, Thanushan Kugathasan, Cesar Augusto Marin Tobon, Konstantinos Moustakas, Heinz Pernegger, Francesco Piro, Petra Riedler, Enrico Junior Schioppa, Abhishek Sharma, Walter Snoeys, Carlos Solans Sanchez, Tomislav Suligoj, Tianyang Wang, Piotr Rymaszewski, Ignacio Asensi Tortajada

***Abstract*—Radiation hard silicon sensors are required for the upgrade of the ATLAS tracking detector for the High-Luminosity Large Hadron Collider (HL-LHC) at CERN. A process modification in a standard 0.18 μm CMOS imaging technology combines small, low-capacitance electrodes (∼2 fF for the sensor) with a fully depleted active sensor volume. This results in a radiation hardness promising to meet the requirements of the ATLAS ITk outer pixel layers (1.5×10^{15} n_{eq}/cm^2), and allows to achieve a high signal-to-noise ratio and fast signal response, as required by the HL-LHC 25 ns bunch crossing structure.**

The radiation hardness of the charge collection to Non-Ionizing Energy Loss (NIEL) has been previously characterised on prototypes for different pixel sensor cell designs. The encouraging results enabled the design of full-size monolithic CMOS sensors for the ATLAS ITk outermost pixel layer, which comprises ∼1.8 m^2 of pixel sensor active area. In the MALTA sensor, we implement a fast, low-power analogue front-end together with a novel high-speed matrix readout architecture capable of meeting the challenging hit-rate requirements of up to 2 MHz/mm^2 in the outer layers of the ITk pixel tracker. The front-end was optimized for the low sensor capacitance to achieve low noise (ENC $<$ 20 e-) and low power operation ($<$ 1 μW/pixel), with timing that meets the 25 ns requirement. The small size (∼2 μm) of the collection electrode also allows better shielding to prevent crosstalk from the full swing digital signals in the 36.4×36.4 μm^2 pixel. MALTA features a 512×512 pixel matrix with a fully asynchronous readout architecture, without clock distribution over the matrix. This approach combines low digital power consumption with fast signal response and high hit-rate capability.

This paper describes the implementation of this novel depleted monolithic sensor based on a low-capacitance analogue design with asynchronous readout, together with first test results from lab tests, radioactive source tests and X-ray measurements.

I. Berdalovic is with the CERN Experimental Physics Department, CH-1211 Geneva, Switzerland and with the Faculty of Electrical Engineering and Computing, University of Zagreb, 10000 Zagreb, Croatia (e-mail: ivan.berdalovic@cern.ch).

R. Cardella, F. Dachs, V. Dao, L. Flores Sanz de Acedo, T. Kugathasan, C. A. Marin Tobon, H. Pernegger, F. Piro, P. Riedler, E. J. Schioppa, W. Snoeys and C. Solans Sanchez are with the CERN Experimental Physics Department.

L. Simon Argemi is with the SUPA School of Physics and Astronomy, University of Glasgow, Glasgow G128QQ, United Kingdom.

T. Hemperek, K. Moustakas, T. Wang and P. Rymaszewski are with the Physikalisches Institut, Universität Bonn, D-53115 Bonn, Germany.

B. Hiti is with the Jožef Stefan Institute, SI-1000 Ljubljana, Slovenia.

A. Sharma is with the University of Oxford, Oxford OX13PA, United Kingdom.

T. Suligoj is with the Faculty of Electrical Engineering and Computing, University of Zagreb.

I. Asensi Tortajada is with the University of Valencia, 46010 Valencia, Spain.

***Index Terms*—Active pixel sensors, CMOS integrated circuits, position sensitive particle detectors, radiation effects, radiation hardening (electronics), semiconductor detectors, solid state circuit design**

I. Introduction

MONOLITHIC CMOS pixel sensors are being considered for the outer pixel layer of the ATLAS tracker upgrade for the High-Luminosity upgrade [1], [2]. After a process modification [3] in a standard 0.18 μm CMOS imaging technology [4] provided full depletion of the sensitive layer and an improvement of the radiation tolerance by at least an order of magnitude, up to about 10^{15} 1 MeV n_{eq}/cm^2 required for the ATLAS application [5], [6], a significant development effort in this modified technology was started. Several test chips and two large-scale prototypes, TJ-Monopix and MALTA, were fabricated in a first submission, and the chips became available for testing after thinning and dicing earlier this year. Both large scale prototypes use the same front-end design further described below. TJ-Monopix measures 1×2 cm^2 and uses the traditional synchronous column drain architecture requiring clock or timing distribution over the matrix [7], [8]. MALTA, further presented in this paper, is the first full scale (2×2 cm^2) CMOS sensor for this application and is equipped with a novel asynchronous readout, avoiding the distribution of the clock over the matrix to reduce power consumption.

II. Analogue front-end circuit

The schematic of the MALTA/TJ-Monopix front-end is shown in fig. 1. It is an evolution of the front-end [9] designed in the same 0.18 μm technology for the ALPIDE CMOS sensor for the ALICE ITS upgrade [10], optimised for a fast response: the current in the main branch (M0-M3) has been increased to 500 nA to be compatible with the 25 ns timing required for the ATLAS experiment. This yields a total analogue power consumption of about 0.9 μW/pixel. For MALTA, a signal clipping mechanism using transistor M6 has been included to ensure a return to baseline after $<$ 200 ns to minimise dead time. Another important change is the inclusion of an enclosed layout transistor (M5) in the design for better tolerance against ionising radiation (TID). The front-end has been optimised for a threshold of ∼200 e$^-$ and an in-time threshold of ∼300 e$^-$ to obtain full in-time efficiency, considering that a minimum ionising particle

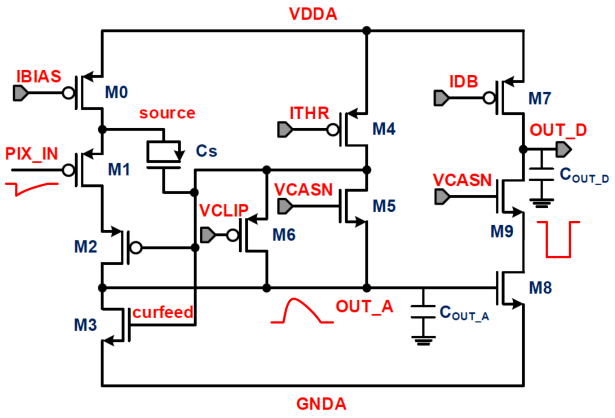


Fig. 1: Schematic of the analogue front-end circuit [9]

releases ∼1500 e$^-$ of charge in the 25 µm thick epitaxial layer. Analogue simulations with parasitic extraction give an expected noise level of less than 10 e$^-$ ENC, with an RMS channel-to-channel threshold variation of about 10 e$^-$. 8 different pixel flavours with different reset mechanisms and different collection electrode/well geometries have been implemented in this prototype.

MALTA also contains several pixels with access to the analogue output node of the front-end before discrimination (OUT_A on fig. 1), which gives the possibility of monitoring the sensor and front-end response before and after irradiation. Fig. 2 shows the spectra of an ^{55}Fe source obtained using these analogue test outputs. The measurement has been done both on an unirradiated chip as well as a chip irradiated with neutrons up to 10^{15} n_{eq}/cm^2 (this chip has also received a background TID dose of 1 Mrad). In both cases the samples were cooled down to −30 °C. The spectra clearly illustrate the functionality and resolution of the sensor. Before irradiation, both characteristic peaks of the source are clearly visible, and the energy resolution obtained from the K_α peak is around 70 eV FWHM. This degrades somewhat after irradiation due to an increase in noise, but the sensor and front-end are still fully functional and both peaks can still be distinguished. Note that the slight increase in the amplitude of the peaks after irradiation is likely due to a slight increase in the gain of the front-end because of TID, which is not compensated for in this measurement.

Fig. 3 shows a time walk measurement performed by collecting a number of waveforms from the same analogue pixels on a chip exposed to a ^{90}Sr radioactive source. The charge threshold setting of this unirradiated sample during the measurement was 210 e$^-$. The $y-$axis shows the time it takes for the analogue signal to reach the discriminator threshold, while the $x-$axis gives the amplitude of the analogue pulse. Signals above ∼130 mV of amplitude arrive within a window of 25 ns, which corresponds to an in-time threshold of about 300 e$^-$ for this setting. Around 5% of the hits in this plot are out of time ($>$ 25 ns), but these are mostly hits caused by charge sharing, where the neighbouring pixels receive the majority of the charge, and therefore do not degrade the in-time efficiency.

The in-time threshold strongly depends on the threshold setting of the front-end. This is demonstrated in the simulation in fig. 4. The curves plotted are the same time walk curves measured in fig. 3, plotted versus input charge instead of voltage, and with the front-end bias settings adjusted for three different thresholds: 100, 200 and 300 e$^-$. The in-time thresholds for these settings, calculated as the minimum charge for which the delay of the discriminator signal is still within 25 ns of the signal delay for a high charge, are 190, 288 and 340 e$^-$ respectively. The simulation also includes the delay of the discrimination stage, but since this delay is negligible already a few tens of e$^-$ above threshold, the simulated number for the in-time threshold at a threshold of 200 e$^-$ matches the measurement results quite well. As evidenced here, the lower one can go with the threshold setting, the lower the in-time threshold will be, resulting in a higher in-time efficiency. Ultimately, it is the noise and threshold dispersion that limit the lowest achievable threshold and thus the in-time efficiency.

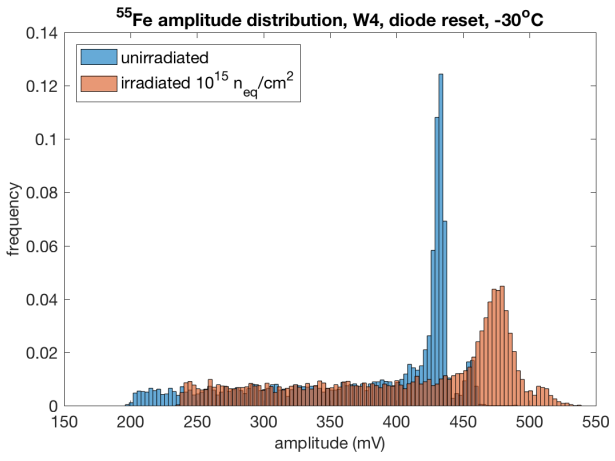


Fig. 2: ^{55}Fe source spectra obtained from analogue test pixels before and after irradiation

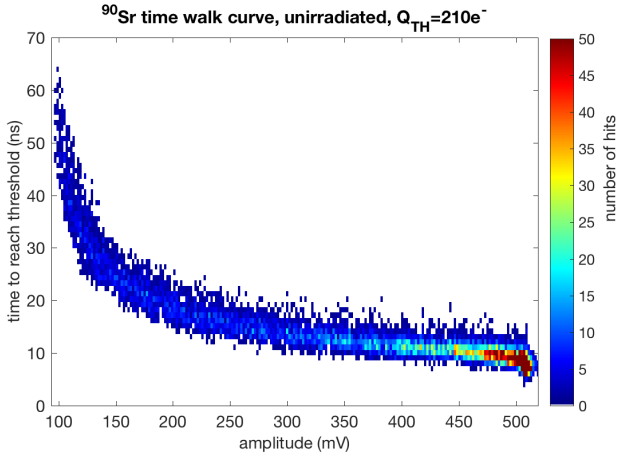


Fig. 3: Time walk measurement obtained from analogue test pixels with a ^{90}Sr source

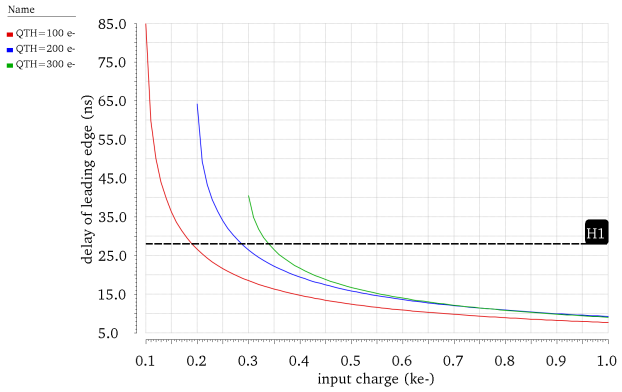


Fig. 4: Simulated time walk curves for thresholds of 100, 200 and 300 e^-. Intersections with the horizontal marker give the in-time thresholds of 190, 288 and 340 e^- respectively.

III. Asynchronous Readout Architecture

The 512×512 pixels in MALTA are organised in double columns, and within a double column in sets of 2 (columns) by 8 (rows). Alternating sets of 16 pixels are connected to two output buses per double column, as depicted in fig. 5 [11]. When a pixel within a set of 16 pixels detects a hit exceeding the charge threshold, it will activate a reference or hit signal, the one line out of 16 corresponding to the hit pixel, and the 5-bit group address corresponding to the set of 16 pixels where the hit was detected. A pulse with a length programmable to 0.5, 1 or 2 ns is transmitted in parallel on every line which needs to transmit a logic one. The delay in the transmission line is matched between the different lines on the bus, and the total number of lines per bus is 22 (1+16+5). If two pixels within one set react simultaneously, the two corresponding lines are activated and only one word is transmitted on the bus. If pixels within one set receive a sufficiently different amount of charge and react one after the other, two words are transmitted sequentially over the bus, guaranteeing sufficient separation of the pulses on the bus for proper transmission. An example of three sequential 2 ns reference pulses at the output of the chip, together with the analogue output waveform of one fired pixel is shown in fig. 6. The sets of 16 pixels in a double column are alternated between the two output buses to prevent data collisions if a particle hits at or near the boundary between two different sets. Data is transmitted almost instantaneously, and is therefore available at the periphery only a few nanoseconds after the hit took place. At the periphery, the hit information is merged and finally provided on a 40-bit wide parallel LVDS output port [12], which includes the full pixel address and some timing information. The massive parallelism in the matrix provides the necessary bandwidth for high-rate applications, but since all information is merged into a single bus, the hit rate in this prototype is limited to the bus bandwidth.

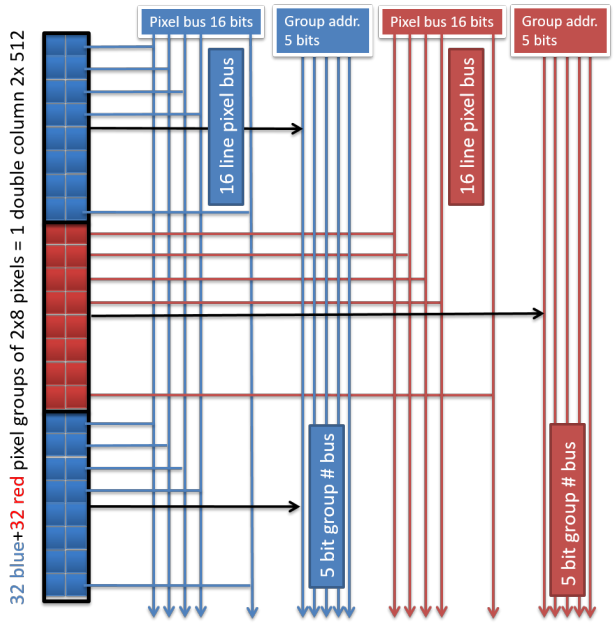


Fig. 5: The MALTA asynchronous readout architecture [11]

During testing, the MALTA chip is read out with an FPGA, which asynchronously oversamples the 40 output signals and performs further processing of the hit data. Using the full readout chain one can obtain a similar time walk measurement as from the analogue monitoring pixels, but this time including the propagation delay of the signals down the column. By measuring the delay of the MALTA output signals with respect to a fast trigger signal (in this case provided by a scintillator), one can obtain the timing distribution of hits coming from the matrix and accurately measure the in-time efficiency. Such a timing distribution during a ^{90}Sr source test for an unirradiated chip with a threshold of ∼300 e^- is shown in fig. 7a. Only the leading signals of clusters are included in this calculation. By checking the fraction of hits within any given 25 ns window and finding the maximum of this fraction, we obtain the maximum in-time efficiency of our sensor. As shown in fig. 7b, this number reaches 98% for this threshold setting. Note that this number is reached without any correction for the few nanoseconds of signal propagation delay down the column,

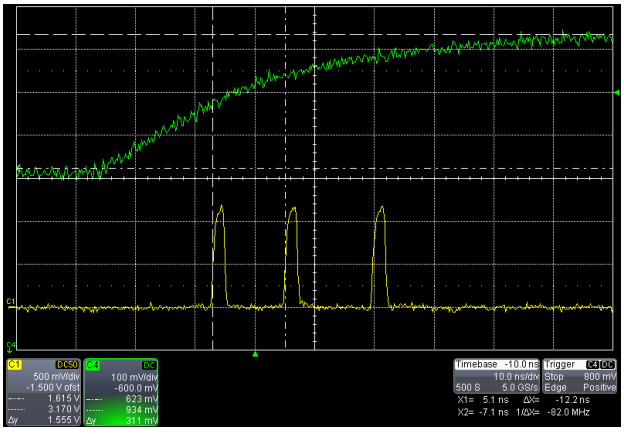
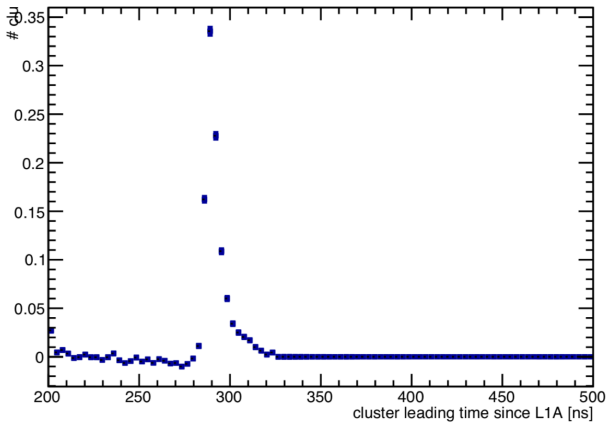
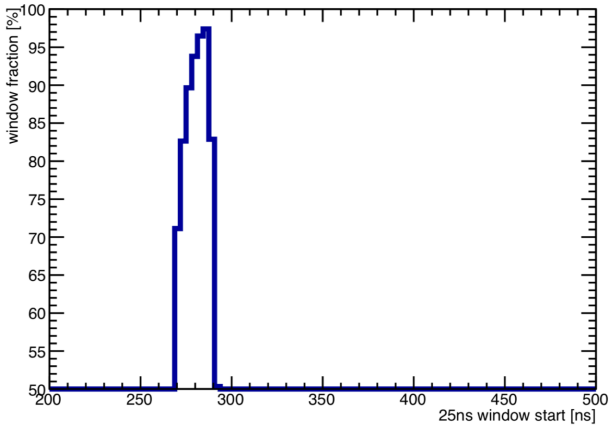


Fig. 6: An example of MALTA output pulses: the analogue output of a front-end (top) and a digital chip output (bottom)



(a)



(b)

Fig. 7: Timing distribution of hits obtained with the FPGA readout during a ^{90}Sr source test with a threshold of 300 e$^-$ (a) and the fraction of hits within a 25 ns window (b).

and that if one were to use the 5-bit group address to correct for this delay, one could achieve full in-time efficiency. Similar measurements have been performed on irradiated devices as well, with negligible degradation in timing after irradiation.

IV. Conclusion

After a process modification of a standard 0.18 µm CMOS imaging process yielded a promising improvement in radiation tolerance for monolithic CMOS sensors, an ATLAS specific development aimed at the outer pixel layer of the inner tracker upgrade for HL-LHC was started. MALTA is a full-scale CMOS sensor with a low-power analogue front-end taking advantage of the low-capacitance collection electrode, and a fully asynchronous readout architecture to minimise digital power consumption and activity in the matrix [13]. The chips have recently returned from fabrication, and results are becoming available both for unirradiated and irradiated devices.

Measurements on both the analogue output signals and the digital output data of the chip in radioactive source tests show promising results in terms of timing. Based on these results, the front-end is expected to be fully in-time efficient within the ATLAS environment. Lab tests on neutron irradiated devices show little change in the front-end signal and timing after irradiation. Further characterisation of the sensor and front-end in beam tests after irradiation is ongoing [14], [15].

Acknowledgement

This paper has been produced as part of the Smart Sensor Technologies and Training for Radiation Enhanced Applications and Measurements (STREAM) project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 675587.

References

[1] The ATLAS Collaboration, "The ATLAS experiment at the CERN Large Hadron Collider", 2008 JINST **3** S08003.

[2] The ATLAS Collaboration, "ATLAS Phase-II upgrade scoping document", CERN-LHCC-2015-020.

[3] W. Snoeys et al., "A process modification for CMOS Monolithic Active Pixel Sensors for enhanced depletion, timing performance and radiation tolerance", Nucl. Instr. Meth. A, vol. 871, Nov. 2017, pp. 90-96.

[4] S. Senyukov et al., "Charged particle detection performances of CMOS pixel sensors produced in 0.18µm process with a high resistivity epitaxial layer", Nucl. Instr. Meth. A, vol. 730, Dec. 2013, pp. 115-158.

[5] H. Pernegger et al., "First tests of a novel radiation hard CMOS sensor process for Depleted Monolithic Active Pixel Sensors", 2017 JINST **12** P06008.

[6] C. Riegel et al., "Radiation hardness and timing studies of a monolithic TowerJazz pixel design for the new ATLAS Inner Tracker", Proceedings of the 18th International Workshop on Radiation Imaging Detectors, 2017 JINST **12** C01015.

[7] T. Wang et al., "Depleted fully monolithic CMOS pixel detectors using a column based readout architecture for the ATLAS Inner Tracker upgrade", 2018 JINST **13** C03039.

[8] K. Moustakas et al., "CMOS monolithic pixel sensors based on the column-drain architecture for the HL-LHC upgrade", Nucl. Instr. Meth. A, in press.

[9] D. Kim et al., "Front end optimization for the monolithic active pixel sensor of the ALICE Inner Tracking System upgrade", 2016 JINST **11** C02042.

[10] G. Aglieri et al., "Monolithic Active Pixel Sensor development for the upgrade of the Inner Tracking System of the ALICE experiment at CERN", 2013 JINST **8** C12041.

[11] I. Berdalovic et al., "Monolithic pixel development in TowerJazz 180 nm CMOS for the outer pixel layers in the ATLAS experiment", 2018 JINST **13** C01023.

[12] R. Cardella et al., "LAPA, a 5 Gb/s modular pseudo-LVDS driver in 180 nm CMOS with capacitively coupled pre-emphasis", Proceedings of the Topical Workshop on Electronics for Particle Physics (TWEPP 2017).

[13] T. Kugathasan et al., "Monolithic pixel development in 180 nm CMOS for the outer pixel layers in the ATLAS experiment", Proceedings of the Topical Workshop on Electronics for Particle Physics (TWEPP 2017).

[14] B. Hiti et al., "Development of the monolithic MALTA CMOS sensor for the ATLAS ITk outer pixel layer", Proceedings of the Topical Workshop on Electronics for Particle Physics (TWEPP 2018), in press.

[15] R. Cardella et al., "MALTA: an asynchronous readout CMOS monolithic pixel detector for the ATLAS High- Luminosity upgrade", Accepted for presentation at the 9th International Workshop on Semiconductor Pixel Detectors for Particles and Imaging (PIXEL 2018).