EPFL

Thèse n° 10 693

Front-End Circuits for Radiation-Hard Monolithic CMOS Sensors targeting High-Energy Physics Applications

Présentée le 13 octobre 2023

Faculté des sciences et techniques de l'ingénieur
Laboratoire d'architecture quantique
Programme doctoral en microsystèmes et microélectronique

pour l'obtention du grade de Docteur ès Sciences

par

Francesco PIRO

Acceptée sur proposition du jury

Prof. I.-C. Benea-Chelmus, présidente du jury
Prof. E. Charbon, Dr W. J. Snoeys, directeurs de thèse
Dr R. Dinapoli, rapporteur
Dr F. Hartmann, rapporteur
Prof. K. Choo, rapporteur

École polytechnique fédérale de Lausanne

2023

Non chi comincia
ma quel che
persevera.
– Leonardo Da Vinci

Acknowledgements

As I approach the end of my PhD, there are countless people to whom I feel grateful, even though a few lines are not enough to express all my gratitude.
First and foremost, I owe a great deal of gratitude to Dr. Walter Snoeys who, after already working together for almost two years, put trust in me and allowed me to continue working with him as a doctoral student. I particularly admire his creativity and positive attitude. I also feel grateful to Prof. Edoardo Charbon, who gave me the privilege to be part of his laboratory, for the fruitful discussions and his exceptional availability over these years.
Special thanks go also to Thanushan Kugathasan for his continuous technical guidance and, in particular, moral support.
Throughout these years, I had the fortune of sharing the office with a lot of amazing people: Ivan Berdalovic, Roberto Cardella, Leyre Flores, Wenjing Deng, Geun Hee Hong, Leonardo Cecconi, Dominik Dobrievic. Thank you all for creating the best work environment imaginable and for all the laughs that made the loooong workdays more enjoyable.
Thanks also to my current office mates who kept me sane during the final stretch of my PhD: João for his "constructive criticism", the "French ministry of Baguettes" Corentin Lemoine for brightening our breaks with his cakes, and Simone Emiliani who actually kindly read this manuscript and helped me to ensure good quality risking severe neurological damages... I really appreciated.
I would also like to thank my former colleagues from UNIGE: Lorenzo Paolozzi, Pierpaolo Valerio, Emanuele Ripiccini, and Fulvio Martinelli, with whom this amazing experience in Geneva started and who continue to provide valuable advice.
Un ringraziamento speciale va ai miei amici di sempre: Emiliano, Dario, Toni, Gianluigi e Marilena su cui posso sempre contare, e ogni qualvolta ci incontriamo, riescono sempre a farmi sentire a casa, come se il tempo non fosse mai trascorso, nonostante le enormi distanze che ci separano.
Grazie a Mamma, Papà e alle mie sorelle, per essermi stato accanto ad ogni passo e farmi sentire amato ogni giorno.

Geneva, 5 September 2023 Francesco Piro

Abstract

Monolithic pixel sensors integrate the sensor matrix and readout in the same silicon die, and therefore present several advantages over the more largely used hybrid detectors in high-energy physics. They offer an easier detector assembly, lower cost, lower material budget and lower power consumption. This work has been carried out in the EP-R&D program at CERN and its main goal is the development of analog readout solutions for monolithic sensors able to cope with the requirements of future high-energy physics experiments.

The optimization of an analog front-end amplifier integrated in a large-scale monolithic sensor targeting the requirements of the outer layer of the ATLAS Inner Tracker upgrade is shown. The chip is manufactured in the TowerJazz 180 nm imaging process. The sensor is designed with a small collection electrode offering therefore a small capacitance (< 5 fF), key feature to achieve high analog performance. Furthermore, it implements a process modification which enhances the charge collection properties of the sensor and its radiation tolerance. The chip has been extensively tested and the characterization results relevant to the front-end circuit are also presented.

To cope with the demand of higher granularity and lower material budget of future high-energy physics experiments, the possibility of moving future monolithic developments in a smaller node technology has been explored in the framework of the EP-R&D program. One of the main targeted applications is the ALICE Inner Tracking System upgrade. The 65 nm imaging technology from the Tower Partners Semiconductor Co. was considered as a possible candidate. Several test structures have been therefore developed to validate this technology for HEP applications. Amongst these, an analog test structure allows to monitor the analog behavior of the sensor and its charge collection properties. Another prototype, instead, allows to characterize the sensor with a fully-featured readout, composed of an analog front-end amplifier and discriminator followed by a digital logic. This work presents the design of these structures and discusses their main characterization results.

Another important target of the R&D effort is to prove the possibility of realizing wafer-scale monolithic sensors with the stitching technique offered by the aforementioned 65 nm process. Large-scale sensors facilitate the coverage of large sensitive areas and, if able to cover the entire sensitive area of a detector, eliminate the need of tiling multiple chips achieving a significant reduction of the material budget. The upgrade of the ALICE Inner Tracking System is based on

this idea. Two wafer-scale sensors have been developed to gain fundamental knowledge and experience on the stitching technique for particle detection. A primary concern in designing such large systems is to obtain a high yield. The two prototypes feature different readout architectures and cope differently with this issue. An overview on the main design aspects of these two structures is given in this manuscript.

Key words: High-energy physics, vertex detectors, Monolithic Active Pixel Sensors, radiation-hardness, front-end electronics.

Résumé

Les capteurs monolithiques à pixels intègrent la matrice sensible et l'électronique de lecture dans la même puce en silicium. De ce fait, ils possèdent plusieurs caractéristiques avantageuses pour leur utilisation en physique des hautes énergies en comparaison aux plus répandus capteurs hybrides. Ils sont plus faciles à assembler, moins couteux, ont un plus faible budget matière et une consommation plus faible. Ce travail, réalisé dans le cadre du programme EP-R&D du CERN, a pour objectif principal le développement d'une électronique de lecture analogique pour les capteurs monolithiques capable de répondre aux besoins des futures expériences en physique des hautes énergies.

L'optimisation du front-end analogique intégré dans un capteur monolithique de grande taille, répondant aux exigences de la couche externe de la mise à niveau du tracker interne d'ATLAS est présentée. Cette puce est fabriquée dans le processus d'imagerie TowerJazz 180 nm. Le capteur est conçu avec une petite électrode de collection présentant donc une faible capacité (< 5 fF), une caractéristique clé pour atteindre des performances analogiques élevées. En outre, elle inclut une modification du processus de fabrication qui améliore à la fois les propriétés de collection de charge du capteur mais aussi sa radio-tolérance. La puce a fait l'objet de mesures approfondies et les résultats de la caractérisation du front-end sont également présentés.

Pour répondre aux besoins des futures expériences de physique des hautes énergies tels que la réduction de la taille des pixels ou la diminution du budget matière, la possibilité d'utiliser un nœud technologique plus fin a été explorée dans le cadre du programme EP-R&D. L'une des principales applications visées est la mise à niveau du tracker interne d'ALICE. La technologie d'imagerie 65 nm de Tower Partners Semiconductor Co. a été étudié en tant que candidat. Plusieurs structures de test ont donc été développées et fabriquées pour valider cette technologie pour les applications en physique des hautes énergies. Parmi celles-ci, une structure de test analogique permet de caractériser le comportement analogique du capteur et ses propriétés de collecte de charges. Un autre prototype permet quant à lui de caractériser le capteur via une chaine de lecture complète, composée de l'amplificateur et du discriminateur analogiques, mais également de la logique numérique. Ce travail présente la conception de ces structures de test et les résultats de leur caractérisation.

Un autre objectif important de la R&D est de prouver la possibilité de réaliser des capteurs monolithiques de très grande taille avec la technique du stitching offerte par la technologie 65nm susmentionné. Ces capteurs de très grande taille facilitent la couverture de larges

zones et, si leur taille permet de couvrir la totalité de la zone sensible d'un détecteur, il n'est plus nécessaire d'assembler plusieurs puces, ce qui permet une réduction significative du budget matière. La mise à niveau du tracker interne d'ALICE est basée sur cette idée. Deux prototypes de capteur très grande taille ont été mis au point pour acquérir des connaissances et de l'expérience sur la technique de stitching et sur son utilisation pour la détection des particules. L'une des principales préoccupations lors de la conception de systèmes de cette taille est d'obtenir un taux de succès élevé lors de la fabrication. Les deux prototypes sont dotés d'architectures de lecture différentes et répondent différemment à ce problème. Une vue d'ensemble des principaux aspects de la conception de ces deux structures est présentée dans ce manuscrit.

Mots clefs : Physique des hautes énergies, détecteur de vertex, détecteur monolithique à pixels actifs, radio-tolérance, électronique front-end.

Contents

List of Figures

List of Tables

Glossary

ALICE A Large Ion Collider Experiment. iii, xi, xvii, 7, 8, 11–15, 37, 39, 59, 76, 80, 82, 85, 93, 101, 111, 115–117
ALPIDE ALice PIxel DEtector. 37, 39, 59
APTS Analog Pixel Test Structure. xiii, xvii, 15, 59, 61–66, 71, 90, 115
ATLAS A Toroidal LHC ApparatuS. iii, xi, xvii, 7–10, 13–15, 27, 37, 48–50, 57, 115

CCD Charge-Couple Device. 1, 14
CERN Conseil Européen pour la Recherche Nucléaire. iii, 7, 55, 59, 88, 90, 93, 115
CMS Compact Muon Solenoid. 7, 8, 27
CSA Charge-Sensitive Amplifier. 28, 31–33, 41, 72

DPTS Digital Pixel Test Structure. xiv, xvii, 15, 59, 72, 79, 80, 82, 86, 88, 90, 91, 95–97, 101, 104, 105, 113, 116

ELT Enclosed Layout Transistor. 35, 36, 48
ENC Equivalent Noise Charge. xii, xiv, 29, 33, 44, 46, 47, 54, 55, 66, 78, 86–88, 97, 105, 115, 116
EP Experimental Physics. 59, 90, 91, 93, 115

FHR Fake-Hit Rate. xiv, 86, 88, 89

HEP High-Energy Physics. iii, 1–3, 7, 11, 13–15, 18, 27–29, 38, 59, 90, 91, 93, 115–117

IBL Insertable B-Layer. 9
ITk ATLAS Inner Tracker. xi, xvii, 9, 10, 15, 37, 48, 49, 57, 115
ITS Inner Tracking System. xi, xvii, 11–15, 37, 59, 76, 80, 82, 93, 101, 111, 115–117

LHC Large Hadron Collider. 7–9, 11, 85
LHCb LHC beauty. 7, 8

MALTA Monolithic from ALice To Atlas. xii, 37, 38, 49, 50, 55, 57, 59, 72, 115
MAPS Monolithic Active Pixel Sensor. xi, 11, 17, 27, 37, 61
MIP Minimum Ionizing Particle. 18, 20, 50, 85, 101, 111
MOSS MOnolithic Stitched Sensor. xiv, xvii, 15, 93, 95–104, 106, 109, 110, 112, 113, 116, 117
MOST MOnolithic Stitched sensor with Timing. xiv, xv, xvii, 15, 93, 103–105, 107–113, 116, 117

NIEL Non-Ionizing Energy Loss. 25–27, 30, 35, 37–39, 48, 53, 57, 59, 71, 87–91, 115

PS Proton Synchrotron. 88
PVT Process, Voltage and Temperature. 84

QGP Quark-Gluon Plasma. 8

RINCE Radiation-Induced Narrow Channel Effects. 35
RISCE Radiation-Induced Short Channel Effects. 35
RMS Root Mean Square. 3, 4, 34, 51, 52, 66, 85
RTS Random Telegraph Signal. 37, 46, 54, 55, 97

SCT SemiConductor Tracker. 9

SEE Single-Event Effects. 36
SEL Single-Event Latchup. 36
SEU Single-Event Upset. 36, 83
SNR Signal-to-Noise Ratio. 4, 28, 33, 34, 44, 47
SPS Super Proton Synchrotron. 55
STI Shallow Trench Isolation. 34

TDC Time-to-Digital Converter. 51
TID Total Ionizing Dose. xii, 34–36, 40, 48, 53, 55, 59, 71, 87–89
ToT Time over Threshold. xiv, 46, 78, 80, 81, 84, 91, 101, 106, 111, 116
TPSCo Tower Partners Semiconductor Company. xiii, 59, 60, 90, 91, 93, 112, 115, 116
TRT Transition Radiation Tracker. 9
TTS Transistors Test Structure. 59

1 Introduction

The search for the smallest building blocks of matter is an eternal ongoing quest. High-energy physics (HEP) pursues this search seeking to understand the forces that govern matter and therefore the universe. Important steps forward in this field are often related to the development of novel and revolutionary detector systems, which give insights into the properties of particles and radiation interacting with them. After the discovery of the radioactivity by H. Becquerel, photographic emulsion, and later the invention of the cloud chamber [1] and of the bubble chamber [2], enabled the visualization of particles through the observation of their tracks. These systems led to the discovery of many particles [3] [4]. They recorded the interactions with the exposure to photographic films and their main limitation was the inability to handle high interaction rates. To overcome this limitation, the use of electronic readouts was introduced with the gas-based multi-wire proportional chambers [5], which allowed computer-aided reconstruction of the tracks. Advancements in the fabrication of semiconductor processes led gas-based detectors to be superseded by semiconductor detectors, as silicon microstrips [6], due to their higher channel density and thus spatial resolution. Only a one-dimensional information on the particle track is provided by these types of detectors. To observe particle tracks with a two-dimensional spatial information, Charge Coupled Devices (CCDs), developed as memories at first and as light detectors later, have been employed also in HEP experiments [7]. The idea of a pixel sensor with a bi-dimensional array of elements integrating electronic circuits for an "intelligent" collection of data on incident particles was proposed in [8]. Progress in microelectronic technologies with the miniaturization of electronic devices rendered pixel sensors suitable to cope with the needs of high granularities and interaction rates, essential to probe rare physics phenomena. Pixel sensors are nowadays one of the instruments of choice for HEP experiments.

This chapter aims to highlight the importance and the role of pixel sensors. Section 1.1 gives an overview on HEP experiments. Section 1.2 explains the main requirements of pixel sensors, and Section 1.3 their challenges for the next generation of HEP experiments. Section 1.4 reports the objectives and the contributions of this work. Finally, Section 1.5 describes how this manuscript is organized.

1.1 Motivation and background

In HEP experiments, to probe deep into the subatomic structure, particles are accelerated to increase their energy and forced to collide either with one another or with a stationary target. When a high-energy particle collides, part of its energy may convert into mass ($E = mc^2$), generating numerous other particles. Some of them are of particular interest but cannot be directly observed as they quickly decay into a few secondary particles. The existence of these short-lived particles can only be inferred by detecting and reconstructing the point of origin of their decay products. The location of the initial collision is called primary vertex, whereas the points where the secondary particles originate are referred to as secondary vertices. In this context, sophisticated systems made of several detector planes are used to image the trajectories of particles as they emerge from the collision point. These systems are indeed named tracking detectors or, when topologically optimized to reconstruct primary vertices of interactions, vertex detectors. A schematic representation of the decay of a short-lived particle measured with three detector planes is depicted in Figure 1.1.

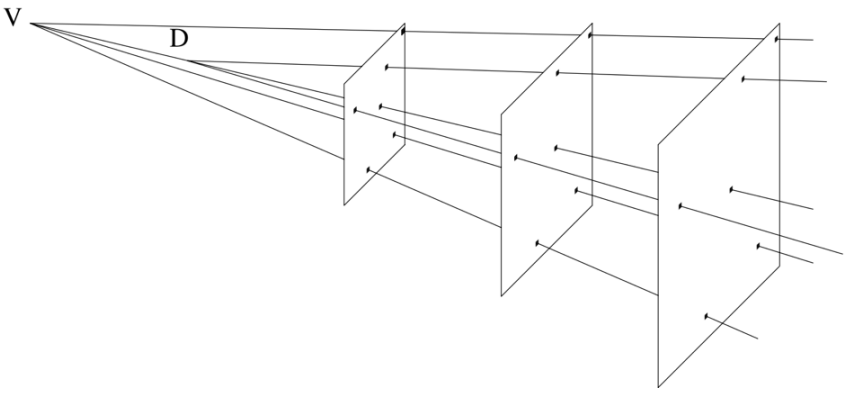


Figure 1.1: Topology of a short-lived particle decay. V indicates the primary vertex whereas D the decay or secondary vertex. Three detector planes are used for track reconstruction [9].

In HEP experiments at accelerators, the particles are accelerated in bunches and confined into a pipe where they are made to collide. The vertex detector is placed closest to the beam pipe and is typically immersed in a magnetic field which bends the tracks of charged particles and enables their momenta to be measured. For the generation of the magnetic field, a solenoidal magnet typically surrounds the vertex detector with the colliding beams of particles crossing its center. To cover the highest possible percentage of the full solid angle, the sensitive layers are usually arranged in concentric cylinders around the collision point and supplemented by endcap disks. A calorimeter is placed outside the vertex detector and its purpose is to measure the energy of the generated particles by absorbing them. Typically, only a small fraction (less than 0.01 % [10]) of the events are of interest, therefore, processing of the data is performed to decipher them and to decide whether they are interesting enough to be stored or, alternatively, to be discarded. This decision takes place in real-time to keep the amount of stored information within acceptable limits. A schematic overview of a high-energy physics detector is shown in Fig. 1.2.

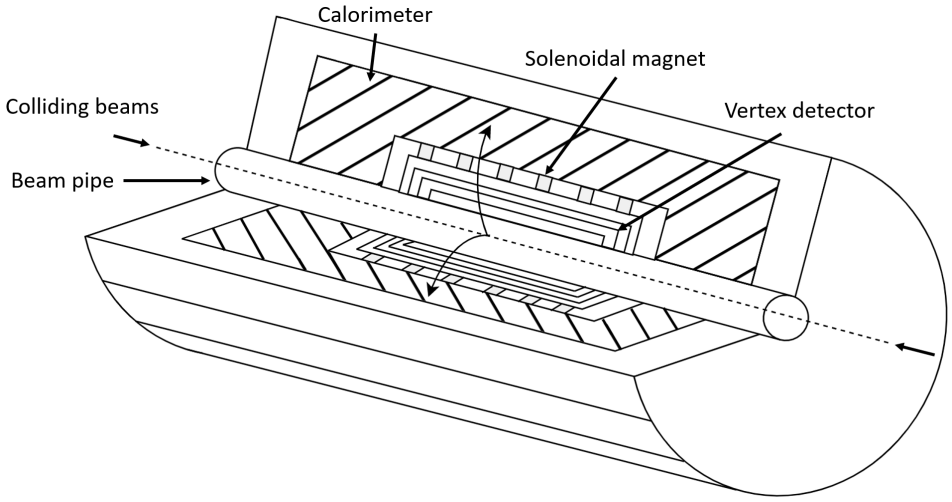


Figure 1.2: Schematic structure of a high-energy physics detector.

To return 3D space points for the reconstruction of the tracks, the sensitive layers are "pixellated", i.e. bi-dimensionally segmented into sensing elements, the pixels. These collect the charge generated into a sensitive layer upon interaction with a particle and incorporate circuital functionality to process the signal and subsequently transmit the data. To probe deeper into the sub-atomic structure, HEP experiments aim at increasing the energy of the colliding particles. Additionally, to search new physics phenomena, increasingly rare processes must be probed and higher collision rates are also desired. To cope with the demand of increasing density of tracks, more and more accurate vertex detectors are required. The performance of pixel sensors have a direct impact on the tracking precision of the vertex detector. The success of future HEP experiments therefore highly relies on the progress of pixel sensors.

1.2 Requirements of pixel sensors for particle tracking

The basic design parameters of pixel sensors that influence the tracking performance of vertex detectors are herewith presented.

1.2.1 Position resolution

Pixel sensors make use of a matrix of electrodes to measure the space coordinates of a particle passage through a reference plane. The spatial resolution of the measurement is primarily determined by the segmentation width (pixel pitch) of the sensitive area and it is defined as the standard deviation, i.e. root mean square (RMS), of the measurement error distribution due to the non-infinitesimal size of the pixels. In some applications, only binary information on the event coordinates is obtained, typically by comparing the charge generated in the sensitive layer against a threshold value. Assuming a uniform particle occupancy, the spatial resolution across one dimension can be estimated as the standard deviation of a uniform distribution

($f(x) = 1$) across the pixel pitch p, which is

$$\sigma = \frac{\int_{-p/2}^{p/2} x^2 f(x) dx}{\int_{-p/2}^{p/2} f(x) dx} \rightarrow \sigma = \frac{p}{\sqrt{12}}. \tag{1.1}$$

The extension to two orthogonal dimensions (x and y) can be obtained by considering the geometrical mean $\sigma = \sqrt{\sigma_x \sigma_y}$ [9]. The spatial resolution improves if the particle travels close to the border of a pixel and the generated charge gets collected by more than one element. If information on the collected charge is also provided, it is possible to use a center of gravity algorithm to determine the impact position, which leads to an even more precise estimate. The precision of the interpolation improves as a function of the signal-to-noise (SNR) ratio of the estimated charge [9]. The increased complexity of the pixel functionality to provide the charge information might lead to a larger pixel area reducing the benefits on the spatial resolution.

1.2.2 Multiple scattering

When a particle travels through the detector, it is deflected by many small-angle scatters. This process is called multiple scattering and is mainly caused by the Coulomb interaction between the particle and the nuclei. The scattering angle of the particle when leaving the material after a large number of interactions follows a Gaussian distribution with an RMS value of [11]:

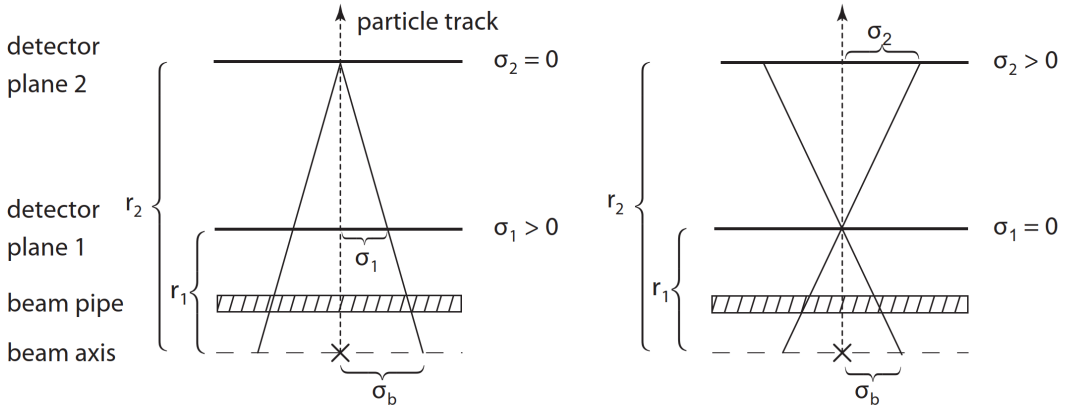
$$\theta_{RMS} = \frac{13.6 \frac{MeV}{c}}{p\beta} |z| \sqrt{\frac{x}{X_0}}, \tag{1.2}$$

where p, β and z are respectively momentum, normalized velocity and charge of the particle. x/X_0 is instead the thickness of the medium in units of the radiation length X_0, which is the mean path length over which a high-energy electron loses all but $1/e$ of its energy. A thick detector material can therefore cause a significant deviation of the particle direction, deteriorating the tracking resolution of the vertex detector. As suggested from eq. 1.2, the uncertainty due to multiple scattering becomes more relevant for low-momentum particles.

1.2.3 Impact parameter resolution

The primary figure of merit of a vertex detector is the resolution of the impact parameter which is defined as the perpendicular distance of the closest approach of a reconstructed track to the primary vertex. To assess which parameters are crucial for a low impact parameter resolution, a simple but indicative case of a vertex detector with two layers can be studied. The geometry under consideration is shown in Figure 1.3. For simplicity, the sensor modules are assumed planar and transversed by straight tracks.

The impact parameter resolution of the track σ_b can be determined by the spatial resolution of each layer and their distance to the interaction point. Summing quadratically the con-



(a) Detector plane 2 assumed to be perfect. (b) Detector plane 1 assumed to be perfect.

Figure 1.3: Simplified two-layer vertex detector. The layers are cylindrically arranged at distances r1 and r2 and have position resolutions σ_1 and σ_2 in the plane perpendicular to the beam [12].

tributions of each plane considering the other perfectly accurate, it is possible to write [12]:

$$\sigma_b^2 = (\frac{r_2}{r_2 - r_1}\sigma_1)^2 + (\frac{r_1}{r_2 - r_1}\sigma_2)^2. \tag{1.3}$$

This result can be expanded for a linear track to a general case of N layers equally distributed over a length L and spaced by $D = L/(N-1)$. Under the assumption that the position resolution σ_{meas} is the same for all the layers, the impact parameter resolution is [12]:

$$\sigma_b = \frac{\sigma_{meas}}{\sqrt{N}}\sqrt{1 + \frac{12(N-1)}{N(N+1)}(\frac{x_0}{L})^2}, \tag{1.4}$$

where x_0 is the extrapolated length of the track. Considering that the layers also feature same thickness d and scattering angle $\theta_{ms,sl}$, the impact parameter resolution due to multiple scattering $\sigma_{b,ms}$ can be calculated by quadratically summing the contributions of each layer [12] and is:

$$\sigma_{b,ms} = \frac{13.6\frac{MeV}{c}}{p\beta}|z|\sqrt{\frac{\frac{d}{\sin\theta}}{X_0}}\sqrt{\frac{N(2N-1)}{6(N-1)}}, \tag{1.5}$$

where θ is the angle between the track and the planes. Considering these equations, for good vertex resolutions, a high intrinsic position resolution of the sensitive layers and a large detector length L are required. Increasing the number of layers in the detector improves the contribution due to the spatial resolution but it increases the contribution of multiple scattering. This number therefore needs to be optimized. Furthermore, the inner layers of the vertex detector should be placed as close as possible to the interaction point to benefit from a shorter extrapolation length and their resolution is the most important.

1.2.4 Momentum resolution

As explained in section 1.1, a magnetic field B is applied to bend the tracks. The momentum of charged particles can thus be inferred from the track curvature. For the aforementioned case of equally spaced detector layers, the resolution of the transverse momentum p_T can be expressed as [12]:

$$(\frac{\sigma_{p_T}}{p_T}) = \frac{p_T}{0.3|z|}\frac{\sigma_{meas}}{L^2 B}\sqrt{\frac{720}{N+4}}. \tag{1.6}$$

The contribution of multiple scattering to the momentum resolution is instead [12]:

$$(\frac{\sigma_{p_T}}{p_T})_{ms} = \frac{13.6}{0.3\beta LB}\sqrt{\frac{(N-1)\frac{d}{\sin\theta}}{X_0}}\sqrt{C_N}, \tag{1.7}$$

where C_N is a factor that describes different detector layouts and depends on the number of layers. Both contributions reduce with the tracker length. To achieve large lengths while optimizing the cost, the tracker is in most cases split in layers with different technologies: high-granularity pixel detectors are typically used for the inner layers, whose resolution is more relevant, whereas less expensive and accurate techniques are used for the outer ones.

1.2.5 Readout electronics requirements

The geometrical requirements of the pixel sensor for high vertex resolutions should be compounded with the basic performance of the readout electronics. In-pixel electronics is necessary to amplify the generated charge, discriminate it and transmit the acquired data. The main specifications are:

- pixel size: as explained in sections 1.2.1 and 1.2.3, a small pixel size is beneficial for good tracking resolutions. Complex readout circuits might require large areas and be a limiting factor to the pixel size. Therefore, area-efficient solutions have to be adopted. The processing of a particle hit requires a minimum time to be performed. If a hit occurs on a pixel which is still processing a previous event, it might be lost for readout. The loss of data due to hits arriving in a too short time frame is referred to as pile-up. For a given hit rate, a smaller pixel is less likely to be hit and is therefore beneficial also to reduce pile-up.
- power consumption: in view of the large number of channels in the vertex detector, low power consumption is crucial. Large mechanical constructions are used to power and cool down the sensors and these increase the detector material, deteriorating its momentum and vertex resolution due to multiple scattering, as explained in sections 1.2.3 and 1.2.4. A lower power consumption reduces the need of cooling and power distributions, allowing for a lighter and thus more accurate detector.
- time response: in accelerators, the particle bunches are made to collide at a specific

frequency. The pixel sensors are therefore required to react fast enough to be able to associate the events to the correct bunch crossing.

- radiation tolerance: the prolonged exposure of pixel sensors to a particle stream slowly degrades its performance [13]. Placing the detector layers closer to the collision point for a better vertex resolution exposes them to a larger particle flux and their behaviour is more severely affected. The detector performance has to be ensured over its entire scheduled lifetime. The radiation effects on pixel sensors are discussed in Chapter 2.

1.3 Challenges for future HEP experiments at the LHC

The current state-of-the art for HEP experiments is represented by the detectors in operation at the Large Hadron Collide (LHC) [14] at the European Organization for Nuclear Research (CERN). The LHC is currently the most advanced particle accelerator. It accelerates particles up to a center-of-mass energy of 14 TeV in a 27 km ring through a complex of different preceding accelerator stages. The complex of accelerators at CERN is shown in Figure 1.4. The beams are not continuous since the particles are bunched together and collide at discrete intervals. The beams collide at four crossing points around which the four largest experiments, ATLAS (A Toroidal LHC ApparatuS) [15], CMS (Compact Muon Solenoid) [16], ALICE (A Large Ion Collider Experiment) [17] and LHCb (LHC beauty) [18] are positioned.

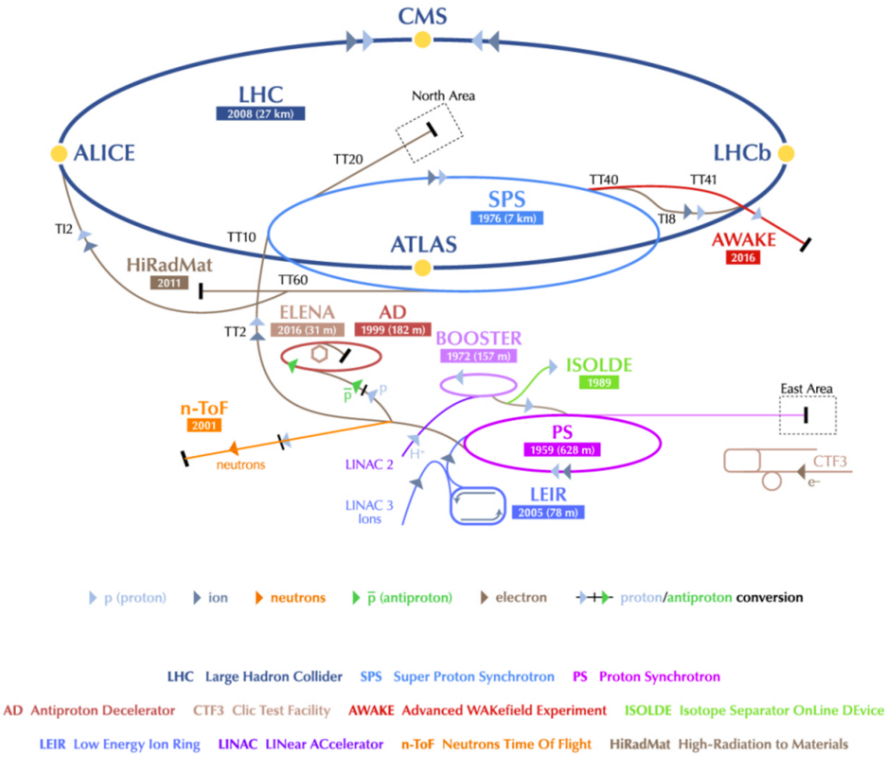


Figure 1.4: An overview of the accelerator complex at CERN and the experiments within it [19].

The quantity that measures the ability of a particle accelerator to produce the required number of interactions is called luminosity L and is defined as the ratio of the number of events per second $\frac{dR}{dt}$ to the production cross-section σ_p [20]:

$$L = \frac{dR}{dt}\frac{1}{\sigma_p}. \tag{1.8}$$

The luminosity can be obtained from [20]:

$$L = \frac{N^2}{t \cdot \sigma_{eff}}, \tag{1.9}$$

where N is the number of particles in each beam (assuming they are equal), t is the time between bunches and σ_{eff} is the effective cross section of the collisions that depends on the beam profile.

ATLAS and CMS [15], [16] are general-purpose experiments which aim to the discovery of new particles or unknown physics phenomena. During their operation, the LHC is filled with beams of protons accelerated to a center-of-mass energy of 14 TeV and made to collide at discrete intervals, 25 ns apart from each other, resulting in a bunch crossing frequency of 40 MHz. The nominal luminosity of the LHC with these conditions is equal to $1 \cdot 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The other two experiments, ALICE and LHCb, focus on specific phenomena. The LHCb experiment [18] specializes in investigating the matter-antimatter asymmetry by studying a type of particle called "b-quark". The ALICE experiment [17] is designed to address the physics of strongly interacting matter and, in particular, the properties of the Quark-Gluon Plasma (QGP), a state of the matter where quarks and gluons are de-confined, i.e. not bound into hadrons. This condition can be recreated with collisions of high-energy heavy ions. During operation of the ALICE experiment, indeed, the LHC is filled with beams of heavy ions, as Pb-nuclei, which are accelerated to a center-of-mass energy of 5.02 TeV and made to collide at a frequency of 50 kHz. In these conditions, the ALICE experiment operates with a nominal luminosity of $1 \cdot 10^{27}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.

One of the major achievements of the LHC was the observation of the Higgs boson with the ATLAS and CMS detectors [21], [22]. The statistical gain in further operating the accelerator without increasing its luminosity is now marginal. To maintain scientific progress and maximize its capabilities, the LHC will therefore undergo a major upgrade during a shutdown planned between 2026 and 2028 [23]. After this upgrade, the peak luminosity of the accelerator will be $5 \cdot 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, a factor of five higher than the current design value. The integrated luminosity, i.e. the amount of collected interactions, is expected to be $3\,\mathrm{ab}^{-1}$ by 2035, approximately 10 times higher than the one attainable with the current machine. To cope with the increased density of tracks, the detectors of the CMS, ATLAS and ALICE experiments will be upgraded in parallel to the LHC. An overview on the upgrade of the ATLAS and ALICE detectors is hereby given as their specifications are taken as a reference for the developments presented in this work.

1.3.1 ATLAS Inner Tracker upgrade

The ATLAS experiment is composed of several sub-detector systems. The Inner Detector [24], the tracking system closest to the interaction point, is shown in Figure 1.5. It is itself divided into three sections with the inner one made out of hybrid pixel sensors, in which the sensitive electrodes and their readout are fabricated in different silicon dies [9]. Initially, the pixel detector was designed as a system with three cylindrical layers around the beam pipe and three disks in each direction along it. The innermost pixel layer is placed at about 5 cm from the beam axis whereas the last one at a radius of 12 cm. The layers cover a total area of $\sim 1.7\,\mathrm{m}^2$ and include over 80 million pixels, about 50 % of the total readout channels of the whole experiment, with a size of 50 µm × 400 µm. A fourth pixel layer with pixels of 50 µm × 250 µm, called Insertable B-Layer (IBL) [25], was subsequently installed at a radius of 3.3 cm with the pretense of high pile-up. The middle section of the Inner Detector is called Semiconductor Tracker (SCT) and is composed of four layers with silicon microstrips [26] with a size of 80 µm × 6 cm. These devices allow for a more practical coverage of larger areas, albeit with one-dimensional accuracy. In the outer section of the Inner Detector, the Transition Radiation Tracker (TRT), gas-filled drift tubes [27] with a diameter of 4 mm and 144 cm long are used to achieve even a larger coverage. The entire structure is surrounded by a superconducting solenoid generating a magnetic field of 2 T.

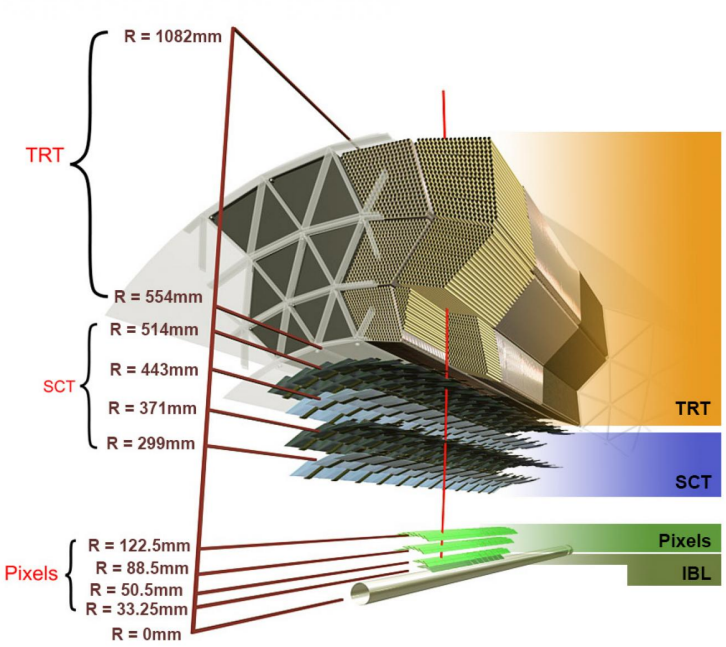


Figure 1.5: Structure of the ATLAS Inner Detector [28].

For the operation of the upgraded LHC, the ATLAS Inner Detector will be replaced with a new tracking system called ATLAS Inner Tracker (ITk), entirely made of silicon sensors [29]. A schematic view of the ATLAS ITk layout is shown in Figure 1.6. It is constituted of two

sub-systems: the one closer to the beam axis is made out of pixel sensors, the outer one is composed of microstrip sensors. The total size of the ATLAS ITk is the same as the ATLAS Inner Detector since the systems that surround it will not be modified. The number of pixel sensor layers has been increased to five with the inner one at a radius of 3.6 cm and the outer one at a radius of 27.3 cm, which is approximately two times higher than in the previous structure. Thus, the total pixel detector area is $\sim 14\,\mathrm{m}^2$, significantly larger compared to the ATLAS Inner Detector. The average hit rate for the inner layer is expected to be as high as $3\,\mathrm{GHz\,cm^{-2}}$ while for the outer layer it drops approximately to $100\,\mathrm{MHz\,cm^{-2}}$, which is similar to the hit rate of the inner layer of the current ATLAS Inner Detector. To cope with the higher hit rate and limit the material budget to 2 % of X_0, the pixel size has been reduced to 50 µm × 50 µm with a power density of $500\,\mathrm{mW\,cm^{-2}}$. A summary of the main requirements of the pixel sensors for the ATLAS ITk comparing the inner and outer layers is given in table 1.1. As a result of the improved geometry and sensor specifications, the transverse impact parameter resolution is expected to improve up to 40 %, reducing from ~ 9 µm to ~ 5 µm for a particle momentum of 100 GeV/c [30].

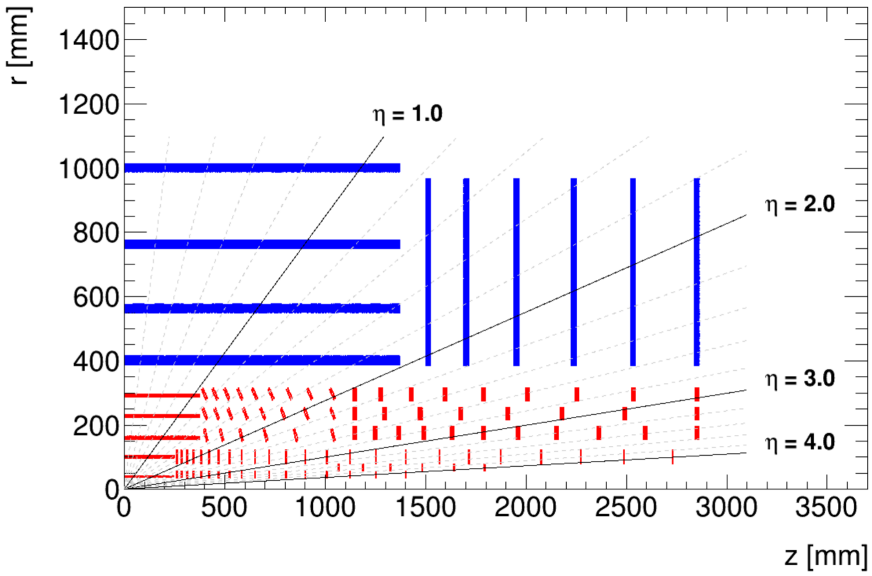


Figure 1.6: Schematic view of an ATLAS ITk quadrant with pixel sensors shown in red and microstrip sensors shown in blue.

Table 1.1: Requirements of the pixel sensors for the inner and outer layers of the ATLAS ITk.

Parameter	Inner layer	Outer layer
Pixel pitch (µm × µm)	50 × 50	50 × 50
Time response (ns)	< 25	< 25
Particle rate ($\mathrm{MHz\,cm^{-2}}$)	3000	100
NIEL fluence ($1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$)	$1 \cdot 10^{16}$	$1 \cdot 10^{15}$
TID (Mrad)	500	80
Power density ($\mathrm{mW\,cm^{-2}}$)	500	500

1.3.2 ALICE Inner Tracking System upgrade

The innermost detector surrounding the beam pipe in the ALICE experiment is called Inner Tracking System (ITS) [31]. The one currently installed, ITS2, is entirely based on Monolithic Active Pixel Sensors (MAPS), in which the sensitive elements and their readout are integrated in the same silicon die [9], with a granularity of $\sim 27\,\mu m \times 29\,\mu m$ [32]. A schematic view of its layout is shown in Figure 1.7. It is composed of seven cylindrical sensitive layers grouped in two separate barrels, called Inner Barrel and Outer Barrel. The Inner Barrel consists of the three innermost layers whereas the Outer Barrel contains the four outermost ones. The seven concentric cylinders cover a radial extension from 22 mm to 430 mm with respect to the beam axis and have a length of 270 mm in the Inner Barrel, 843 mm for the middle layers in the Outer Barrel and 1475 mm for the outermost ones. The sensitive area of the entire tracker is $\sim 10\,m^2$ and it is covered by 12.5 billion pixels, making it the largest scale application of MAPS in a HEP experiment currently in operation.

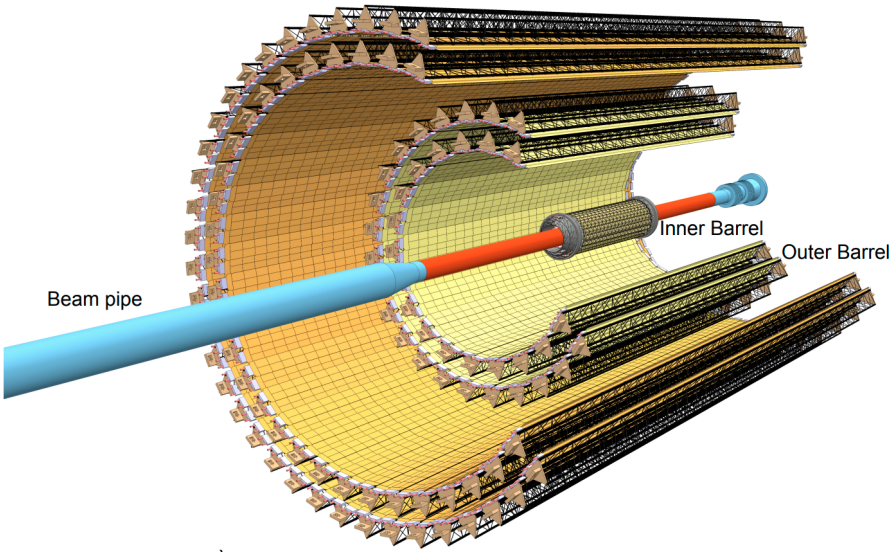
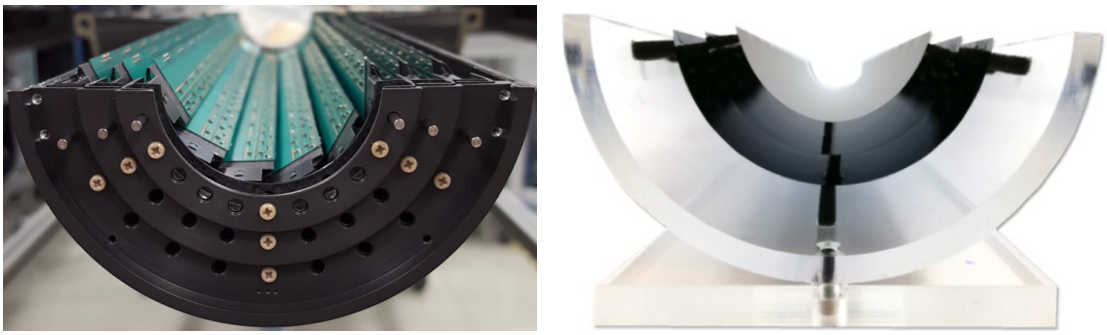


Figure 1.7: Schematic view of an ALICE ITS layout [31].

For the upgrade of the LHC, the ALICE collaboration has expressed interest in upgrading the three innermost layers of the ITS [33]. The goal of the upgrade is to improve the granularity of the sensors and reduce the material budget close to the interaction point. A picture of the Inner Barrel of the current ITS is shown in Figure 1.8a. To cover the surface of the ITS, the pixel sensors are tiled up on mechanical supports which also distribute water cooling pipes, power and electrical signals. The material budget breakdown of a detector layer shows that the pixel sensor, which is ideally the only component needed in the detector acceptance, contributes only to 15 % of the total detector material. In order to reduce the material budget, the electrical, mechanical and cooling material have to be reduced and the idea behind the upgrade is to realize a detector which would completely do away with them. In standard CMOS manufacturing, the maximum size of a chip is limited to the reticle area defined by the



(a) (b)

Figure 1.8: Picture of a) the Inner Barrel of the current ALICE ITS b) the dummy silicon model for the upgrade of the Inner Barrel of the ALICE ITS.

field of view of the photolithographic process, typically a few centimeters in both directions. An established technology, called stitching [34], allows fabricating sensors larger than the reticle size with the only limit given by the wafer size. Thanks to the wafer sizes offered by the modern manufacturing processes, the stitching technology enables the fabrication of pixel sensors with dimensions compatible to the size of an entire layer of the ITS Inner Barrel. By replacing the layers with a single stitched sensor, the distribution of the power and electrical signals could be done entirely inside the silicon chip and the interconnections to the external world confined on the layer edge, outside the detector acceptance, where more material can be used. To remove the material associated with the cooling, it has been experimentally demonstrated that using a low-speed ($<2\,\mathrm{m\,s^{-1}}$) air flow to remove heat by convection is a viable option for power densities below $20\,\mathrm{mW\,cm^{-2}}$ as the vibration of the sensors caused by the airflow are still smaller than their intrinsic spatial resolutions [35]. Furthermore, it has been proved that with thicknesses below 50 µm, the flexible nature of thin silicon can be exploited to bend the sensors without degradation on the performance [36], [37]. The sensors will be therefore thinned down to $\sim 20\,\mu\mathrm{m}$ and curved around the beam pipe. The new Inner Barrel is divided into two halves, named half-barrels, to allow the detector to be mounted around the beam pipe. Each half-barrel consists of three half-layers realized by means of a single stitched sensor, arranged inside the half-barrel as illustrated in Figure 1.8b, which shows a picture of a silicon dummy model. A cylindrical structural shell provides support to the half-layers which are connected to the outer shell and to each other by means of ultra-light spacers which define their radial position. In particular, the three layers are placed at a radial distance of 18 mm, 24 mm and 30 mm from the beam axis and therefore the first layer is also placed closer to the interaction point which exposes it to a larger particle flux. The Inner Barrel is still 270 mm long and, to realize stitched-sensors able to cover this length, processes with wafers at least 300 mm in diameter have to be used. The proposed structure reduces the material budget of each layer by a factor seven, going from 0.35 % of X_0 to an unprecedented low value of 0.05 % of X_0. To further improve the tracking accuracy of the detector, a smaller pixel size in the order of $15\,\mu\mathrm{m} \times 15\,\mu\mathrm{m}$ is targeted. The detector specifications are the result

of an optimization expected to reduce the transverse impact parameter resolution by a factor of two, going for example from ~ 20 µm to ~ 10 µm for a particle momentum of 1 GeV/c [33]. A summary of the main requirements of the pixel sensors for this upgrade is given in Table 1.2.

Table 1.2: Requirements of the pixel sensors for the Inner Barrel of the ALICE ITS upgrade.

Parameter	Value
Pixel pitch (µm × µm)	15 × 15
Time response (µs)	< 1
Particle rate ($MHz\,cm^{-2}$)	2.2
NIEL fluence ($1\,MeV\,n_{eq}\,cm^{-2}$)[a]	$1 \cdot 10^{13}$
TID (Mrad)[a]	1
Power density ($mW\,cm^{-2}$)	20

[a] The absorbed radiation doses are referred to the innermost layer.

1.3.3 Other applications

The development of pixel sensors for particle detection in HEP experiments has spun off in several other applications. In astronomy, pixel sensors are often mounted on satellites for space and earth observation by detecting low-energy X-rays originating from astronomical points [38], [39]. Particle sensors are also largely used as dosimetry devices to measure radiation levels, typically to evaluate the exposure of a body to potentially harmful doses [40], [41]. Another notable field of application is medical imaging where an image of a body is generated for diagnostic purposes by exposing it to a particle beam. A detector is here used to analyze the interactions of the particles with the body (e.g. absorption, scattering). The oldest and most common imaging technique is the radiography, which uses X-rays [42]. Beams of high-energy protons are instead used in the field of proton therapy not only to image a body but also to precisely irradiate a diseased tissue and kill it [43]. In these applications, as in HEP, low-power and highly-granular pixel sensors are highly desirable.

1.4 Thesis contributions

The main goal of this thesis is to propose and realize readout circuits for monolithic pixel sensors that satisfy the requirements of low-mass and highly-granular vertex detectors for future HEP experiments. According to the physics scope of the experiment, different specifications might be required. General-purpose experiments as ATLAS focus on probing new and extremely rare phenomena and therefore aim at increasing the energy of the accelerated particles, which are made to collide at high collision rates to acquire a large amount of data. Fast readouts are therefore needed to be able to associate the events to the correct collision. Other experiments as ALICE study specific physics phenomena which have to be probed with high precision. Due to the beam characteristics, low-momentum particles are generated. As

the uncertainty due to multiple scattering increases on decreasing momenta (see section 1.2.2), it is essential to reduce the detector material as much as possible. On the other hand, the collision rate is much lower and the time constraints are less stringent. The low power consumption of the readout circuits, necessary to keep a low material associated to powering and cooling, can be obtained at the expense of a slower time performance. The main contributions of this thesis are:

- the study and optimization of analog front-end circuits, which amplify and discriminate the generated charge in the sensitive layers. These are typically the most crucial in-pixel components in terms of power and timing. In this respect, several front-end solutions have been designed targeting different specifications. First, the front-end circuit of a pixel sensor in a 180 nm CMOS imaging technology has been optimized to reduce its noise and meet the requirements of the ATLAS experiment upgrade, very demanding in terms of speed and radiation hardness. Subsequently, starting from this circuit, a novel front-end topology has been developed in a 65 nm CMOS imaging technology for the very stringent constraints of the ALICE ITS upgrade focusing more on low power consumption and compactness.

- the development of the analog circuitry required to test structures, which allowed the validation of a sub-100 nm imaging technology for HEP applications. Offering larger integration densities and lower power consumption, fine linewidth technologies help to meet the demands of future HEP experiments for high-resolution and low-mass vertex detectors. The requirements for particle detectors are slightly different from the ones of CMOS imagers. In traditional imaging, the charge is generated within a depth of a few microns. High-energy particles, instead, generate charge across the full thickness of the sensitive layer which has to be collected as efficiently as possible. Furthermore, the pixels have to be sensitive over their entire area. Before a technology can be used for HEP applications, its compatibility with particle detection must first be assessed. This has been successfully done thanks to the work on the test structures described in this thesis.

- the design of monolithic stitched wafer-scale sensors for HEP experiments. To the best of the author's knowledge, the stitching technique has been used so far for HEP applications only on CCDs [44]. It is more commonly used for visible light and X-ray imaging [45], [46]. In these examples, the pixels contain only a few transistors. Pixels in sensors for HEP require much more complex circuitry and contain several hundred transistors. This work reviews the main challenges in the design of such large systems and introduces some possible techniques to cope with them.

1.5 Manuscript organization

The manuscript is structured as follows:

Chapter 2 introduces the fundamentals of pixel sensors, explaining how particles interact with silicon and different sensor approaches. An overview on the front-end circuits as well as on the radiation effects on the sensor and the readout electronics is given.

Chapter 3 presents the MALTA2 chip, a monolithic sensor in the TowerJazz 180 nm imaging technology which targets the specifications of the ATLAS ITk outer pixel layer. The chapter focuses in particular on the design and optimization for low noise of its analog front-end circuit. The prototype was extensively measured and the main results relevant to the front-end are reported.

Chapter 4 shows the sensor in the Tower Partners Semiconductor Co. 65 nm imaging process which is the technology chosen for future monolithic developments, such as the ALICE ITS upgrade. This process is mainly focused on the detection of visible light. A small-scale structure called Analog Pixel Test Structure (APTS) has been designed to verify the possibility of using it for HEP applications as the aforementioned 180 nm technology. A small-scale prototype called Digital Pixel Test Structure (DPTS) has also been designed to validate the sensor with a fully-featured readout. The chapter describes the design of these structures. They have been characterized with laboratory measurements and beam tests. The main characterization results of the structures are also shown.

Chapter 5 gives an overview on the design of two wafer-scale stitched sensors called MOnolithic Stitched Sensor (MOSS) and MOnolithic Stitched sensor with Timing (MOST). One of the main challenges in designing such large sensors is to obtain a high yield, i.e. to prevent that a single manufacturing defect, likely to occur due to the very large area, jeopardizes the entire chip. The two sensors feature different architectures and face this problem with different strategies. The front-end topology of the DPTS has been slightly modified and integrated into these structures by adapting it to their needs.

Chapter 6 concludes the thesis and gives an outlook on future developments.

2 Fundamentals of pixel sensors

The detection of particles is performed through their interaction with matter. The basic detection mechanism of silicon sensors is the generation and movement of mobile charges in a silicon p-n junction upon interaction with a particle. To provide a better understanding of these concepts, the fundamental principles of silicon sensors are discussed in this chapter. Section 2.1 explains the principles of the interactions of particles with matter and their detection with silicon sensors. Sections 2.2 and 2.3 introduce the two main approaches of pixel sensors, i.e. Hybrid Pixel Sensors and Monolithic Active Pixel Sensors (MAPS) respectively. Section 2.4 gives an overview on the readout solutions commonly used to process the charge signals.

2.1 Detection of particles in silicon

Particles passing through a sensor interact with its matter depositing part of their energy. The amount of deposited energy depends on the particle, its energy and the traversed material. The following sections describe the interaction of charged particles and of electromagnetic radiation with matter.

2.1.1 Energy loss of charged particles

The energy lost by a charged particle traversing a medium can be attributed to a sum of ionization, atom excitation, and bremsstrahlung radiation effects. The average energy loss, or stopping power, is characterized by different processes depending on the particle mass and momentum, each of them described by a different theoretical description. As an example, Fig. 2.1 shows the dependence of stopping power for pions in silicon as a function of the normalized momentum $\beta\gamma = p/mc$. This can be described with the Bethe-Bloch formula [47]:

$$< -\frac{dE}{dx} >= Kz^2 \frac{Z}{A}\frac{1}{\beta^2}[\frac{1}{2}ln\frac{2m_e c^2\beta^2\gamma^2 T_{max}}{I^2} - \beta^2 - \frac{\delta}{2}], \tag{2.1}$$

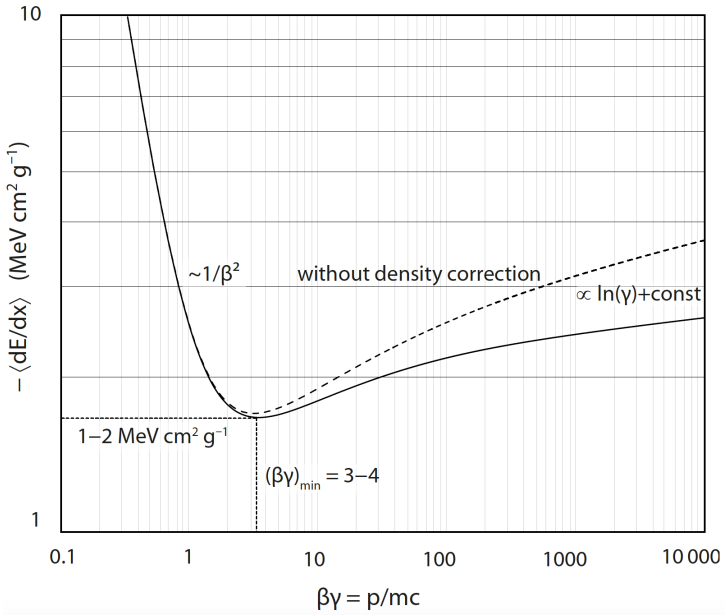


Figure 2.1: Stopping power of pions in silicon as a function of the normalized momentum [12].

where z is the charge of the incident particle, Z the atomic number, A the atomic mass of the absorber, m_e the mass of the electron, c the speed of light, T_{max} the maximum kinetic energy which can be imparted to a free electron in a single collision, I the mean excitation energy, δ the density-effect correction factor described in [48], and β, γ and K are defined as follows:

- $\beta = \frac{v}{c}$.
- $\gamma = \frac{1}{\sqrt{1-(\frac{v}{c})^2}}$.
- $K = 4\pi N_A r_e^2 m_e c^2$.

Here, v is the velocity of the incident particle, N_A the Avogadro number and r_e the classical radius of the electron. At low energies (and therefore momenta), the stopping power is high due to the $\frac{1}{\beta^2}$ term being dominant. This dependence can be explained by the effective interaction time which increases for lower momenta leading to a larger energy loss. For large energies the stopping power increases due to a rise of the maximum energy transfer T_{max} with γ and other relativistic effects [12]. In between these regions there is a minimum at $\beta\gamma \approx 3-4$, whose exact value depends on Z. A particle with energies in this range is called Minimum Ionizing Particle (MIP). Since the energy loss increase for $\beta\gamma > 3-4$ is moderate, also particles with higher energies are referred to as MIPs. In practice, most of the particles in HEP experiments have an energy loss rate close to the minimum. Furthermore, due to the stochastic nature of the involved processes, the energy loss is subject to statistical fluctuations. The probability density function of the energy loss follows a Landau distribution [49]. As an example, the distributions of energy loss per unit of absorber for 500 MeV pions in silicon of

different thicknesses is shown in Fig. 2.2. In the distributions, the average value of the Bethe-Block equation is larger than the most probable value due to long tails, i.e. the presence of few high-loss events. Furthermore, as the silicon thickness decreases, the most probable value also decreases. For very thin layers, the energy loss distribution is not described accurately by the classical Landau function and alternative models are used [50].

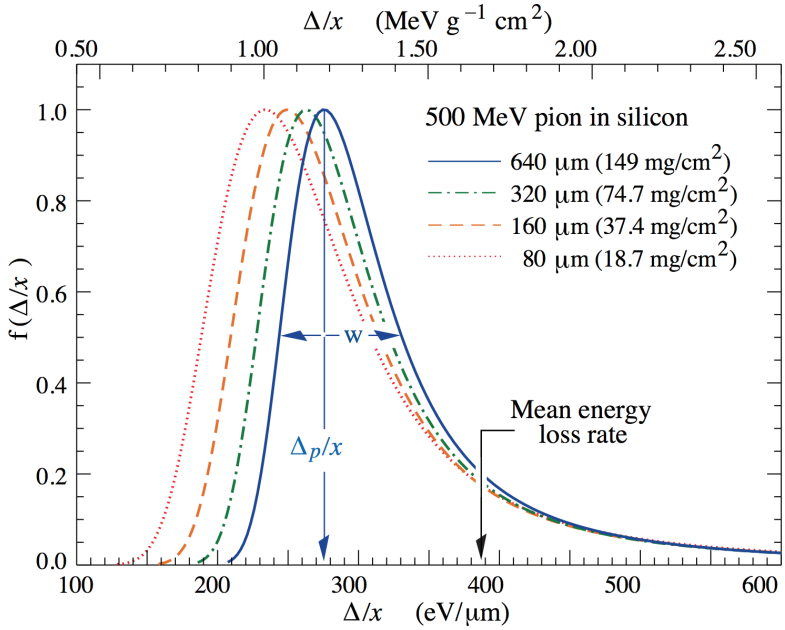


Figure 2.2: Probability density functions $f(\Delta/x)$ of energy loss (Δ) for 500 MeV pions in silicon of different thicknesses (x), normalized to unity at the most probable value [47].

2.1.2 Interaction with electromagnetic radiation

Photons interact with matter mainly via three processes: photoelectric effect, Compton effect and pair production. In the photoelectric effect and pair production, the photon is completely absorbed by the material whereas in the Compton effect, the photon is scattered. The intensity of a monochromatic photon beam penetrating through a material is attenuated as [9]:

$$I(x) = I_0 e^{-\frac{x}{\mu}}, \tag{2.2}$$

with I_0 and $I(x)$ being the initial and final beam intensity after traversing a material of thickness x. The attenuation length μ is function of the material and photon energy. Fig. 2.3 shows the probability of photon absorption in a 300 μm layer of silicon or cadmium telluride. For silicon, the contributions from the different processes are also noted. The photoelectric effect is the dominant process at low photon energies, in silicon below ≈ 100 keV. At higher energies, the cross section of the photoelectric effect drops down and scattering processes become more important. At energies exceeding twice the electron mass, pair production also contributes and becomes the only relevant process at energies exceeding 10 MeV. In silicon, the photon

interaction probability is high up to ≈ 10 keV. Above this value, semiconductors with a higher atomic number Z, as cadmium telluride, are more likely to interact with the photons and are thus preferred.

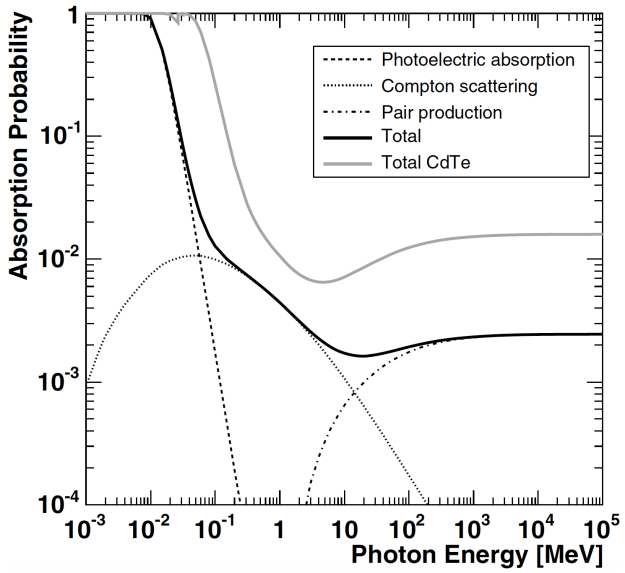


Figure 2.3: Photon absorption probability for a 300 µm layer of silicon or cadmium telluride (CdTe) as function of the photon energy. The contributions from the different processes are indicated for the case of silicon [9].

2.1.3 Charge generation

The energy released in a medium by a particle generates charge carriers in the form of electron-hole (e/h) pairs. In order to create an e/h pair, an energy at least larger than the band-gap energy of the material should be released. Silicon has a band-gap energy of 1.12 eV. Since its band-gap is indirect, however, the average energy required to create an e/h pair is $w_i = 3.6$ eV, which is roughly three times higher than the band-gap energy [9]. The additional energy goes into lattice excitations called phonons, dissipated as thermal energy. For a given amount of deposited energy E, the average number of generated e/h pairs is $N = E/w_i$. In the case of a MIP traversing silicon, the most probable value of energy loss for a layer thickness of tens of µm is ≈ 0.2 keV/µm [51], therefore, the average number of produced e/h pairs is ≈ 60 per µm traversed. The fraction of deposited energy that is used for charge and phonon generation is subject to fluctuations which cause the number of generated e/h pairs to vary by:

$$< N^2 >= FN = F\frac{E}{w_i}. \tag{2.3}$$

F is the so-called Fano factor [52], which is in the order of 0.1 for most semiconductors. It represents the ultimate limit of energy resolution in semiconductor detectors.

2.1.4 Charge transport

The e/h pairs generated through the interaction of particles with matter are free to conduct. The main charge transport mechanisms in a semiconductor are two: diffusion and drift [9]. The diffusion transport takes places when a gradient in the concentration of charge carriers is present: due to thermal motion, the charge carriers move randomly, however, with a concentration gradient they are more likely to arrive in a low-concentration region. For a concentration n and p of electrons and holes respectively, the diffusion current density can be expressed as [53]:

$$J_{diff} = J_{n,diff} + J_{p,diff} = -qD_n\Delta n + qD_p\Delta p, \tag{2.4}$$

where q is the charge of the electron and D_n and D_p are the diffusion coefficients for electrons and holes respectively which depend on the semiconductor material and temperature. As transport by diffusion originates from the thermal random walk, the carrier path is usually long. If an electric field E is applied, the carriers move by drift, i.e. they are accelerated along the field lines while scattering off lattice phonons and crystal defects. The drift current density can be described with [53]:

$$J_{drift} = J_{n,drift} + J_{p,drift} = -qn\mu_n E + qp\mu_p E, \tag{2.5}$$

where μ_n and μ_p are the mobilities of electrons and holes respectively. The mobility is the proportionality factor between the charge speed and the applied electric field and depends on the semiconductor material and the electric field itself. For low fields, the mobility is constant, however, for high fields it gradually degrades leading to the charge velocity to saturate.

2.1.5 P-n junction

In order to construct a particle sensor, the charge released by a traversing particle needs to be generated in an area depleted of free carries which features an electric field to collect it. This is obtained with a reverse biased p-n junction, which is the fundamental building block of practically all silicon particle sensors. Fig. 2.4 shows the structure of an abrupt p-n junction. It is formed by bringing in contact an n-doped and a p-doped silicon crystal. At the boundary between the two regions, the majority carriers diffuse to the opposite part where they recombine. A region depleted of free carriers called space charge region thus forms at the p-n interface. The fixed dopants in the space charge region create an electric field that produces a drift current opposing the majority carrier diffusion until an equilibrium state is reached. The electric field in the structure has a maximum at the p-n interface and is described by [53]:

$$E(x) = \begin{cases} \frac{-qN_A}{\epsilon_{Si}}(x + x_p) & -x_p < x < 0 \\ \frac{qN_D}{\epsilon_{Si}}(x - x_n) & 0 < x < x_n \end{cases}, \tag{2.6}$$

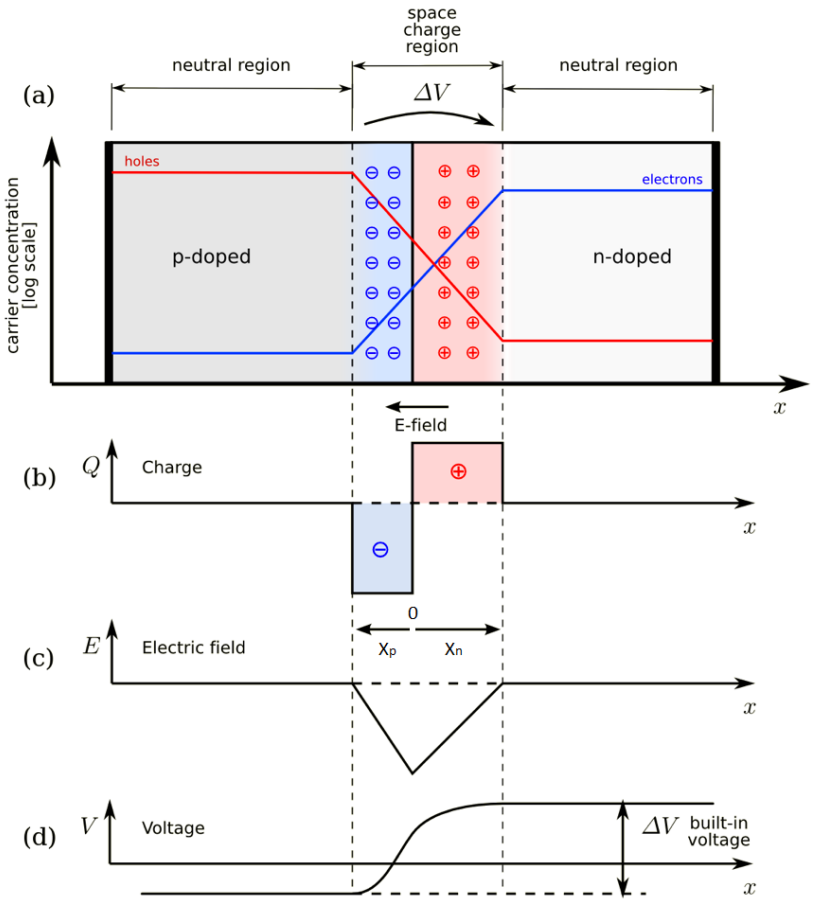


Figure 2.4: The p-n junction: a) structure b) space charge density c) electric field distribution and d) potential distribution [54].

where x is the distance from the junction boundaries, x_p and x_n are the widths of the depleted region in the p-type and n-type regions respectively, ϵ_{Si} is the permittivity of silicon, and N_D and N_A are the donor and acceptor doping concentrations. The potential across the junction, called built-in potential V_{bi}, can be calculated with [53]:

$$V_{bi} = \frac{k_B T}{q} ln(\frac{N_D N_A}{n_i^2}), \tag{2.7}$$

where k_B is the Boltzmann constant, T is the sensor temperature and n_i is the intrinsic carrier concentration. By applying a reverse bias voltage V across the junction in addition to the built-in voltage, the majority carriers on each side can be further removed and the depleted region extended. The width of the depleted region for a planar junction is obtained as the sum of the depletion width in the n-type and p-type region and, by solving the one dimensional

Poisson equation, it can be expressed as [53]:

$$d = x_n + x_p = \sqrt{\frac{2\epsilon_{Si}}{q}(\frac{1}{N_D} + \frac{1}{N_A})(V + V_{bi})}. \tag{2.8}$$

With a reverse bias voltage significantly higher than the built-in voltage and a p-n junction with the n-side much more doped than the p-side, the expression simplifies to [53]:

$$d \approx x_p = \sqrt{\frac{2\epsilon_{Si}}{q}\frac{1}{N_A}V}. \tag{2.9}$$

The resistivity ρ of a semiconductor is roughly inversely proportional to the doping concentration N_A, so from eq. 2.9 one can write [53]:

$$d \propto \sqrt{\rho V}. \tag{2.10}$$

From a circuital standpoint, the reverse-biased p-n junction can be modelled as a capacitance. The capacitance of a planar junction can be estimated with the well-known formula of a parallel-plate capacitor which is [53]:

$$C = \epsilon_{Si}\frac{A}{d} \propto \frac{A}{\sqrt{\rho V}}, \tag{2.11}$$

where A is the area of the junction and d the width of the depleted region of eq. 2.10. The capacitance C is one of the most important properties of the sensor since low capacitance values are highly beneficial on the overall performance of the detector as explained later on. For a low sensor capacitance, large widths d of the depleted regions, i.e. high reverse bias voltages and lowly doped (highly resistive) junctions, are required.

Even in absence of traversing particles, e/h pairs are thermally generated due to impurities in the silicon which act as generation/recombination centers. These carriers produce a leakage current which is proportional to the depleted volume of the sensor and can be calculated as [53]:

$$I_{Leak} \approx -q\frac{n_i}{\tau_g}Ad, \tag{2.12}$$

where τ_g is the carriers' generation lifetime. The intrinsic carrier concentration strongly depends on temperature ($n_i \propto T^{3/2}$), therefore I_{Leak} is also temperature dependent. The leakage current has an impact on the operating point of the readout electronics and increases its input noise. Typically, to minimise the leakage current and limit its influence on the detector, the sensors are operated at low temperatures. Furthermore, a sharp increase of the current can occur if a too large reverse bias voltage is applied to the junction, e.g. to reduce its capacitance. This phenomenon is referred to as breakdown as it can also lead to the destruction of the junction. The breakdown voltage of a junction can be increased with proper selection of its doping levels and geometry.

2.1.6 Signal formation

The movement of the charge generated in the sensor induces a current on the electrodes of the p-n junction where it is collected. The charge generated into the depleted region is collected by drift under the effect of the electric field. The charge generated outside the depleted region moves instead by diffusion until it eventually arrives in the depleted region and is finally collected by drift. According to the Ramo theorem [55], the current induced on an electrode by a charge q which moves with a velocity v (proportional to the electric field through the mobility) is:

$$i = -\frac{dQ}{dt} = qE_w \cdot \vec{v}, \tag{2.13}$$

where E_w is the so-called weighting field, different from the actual electric field in the sensor. It is obtained by applying a unit potential to the electrode under consideration and a zero potential to all the others. The weighting field determines how well the movement of a charge carrier couples with the electrode and depends on the sensor geometry. The charge induced on the electrode by a carrier q drifting in the time interval $[t_1, t_2]$ from position x_1 to x_2 is therefore:

$$Q = \int_{t_1}^{t_2} i(t)\,dt = q[\phi_w(x1) - \phi_w(x2)], \tag{2.14}$$

where ϕ_w is the weighting potential obtained by solving the Poisson equation in the configuration just described. Fig. 2.5 shows the weighting potential under the assumption of an arbitrary sensor thickness of 1 in the y direction, an infinitely wide electrode at $y = 1$ and a collection electrode of different widths at $y = 0$ on which the charge is collected and the unit potential is applied. Fig. 2.5a shows the weighting potential for two infinite parallel plates, while b and c show it for an electrode width of 1/3 and 1/10, respectively. The smaller the

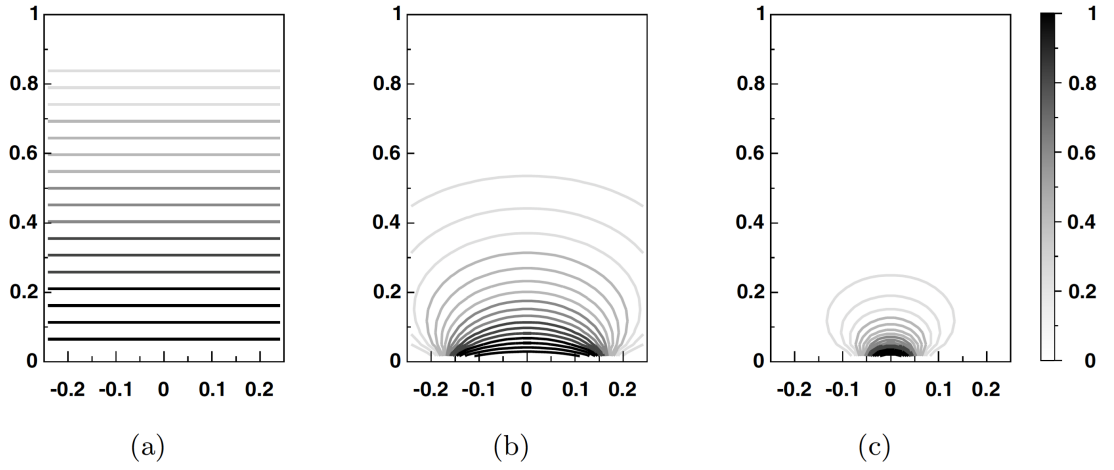


Figure 2.5: Weighting potential for a) two infinite parallel plates and for a collection electrode width of b) 1/3 and c) 1/10 of the sensor thickness, arbitrarily set to 1 [9].

electrode, the larger the gradient of the weighting potential close to the electrode and the area where it approaches zero far from the electrode. For small electrodes, therefore, most of the signal is induced in the last part of the carrier path. Once the carrier reaches the collection electrode, the total charge induced on it is equal to q. In case of a particle passage, the current induced on the electrode has a duration corresponding to the charge collection time and an integral equal to the total charge generated in the sensor.

2.1.7 Radiation effects in the sensor - Non ionizing energy losses

When traversing a silicon sensor, particles that interact with the atomic nuclei of the crystal lattice can displace the nuclei out of their lattice position creating vacancies and interstitials in between the lattice atoms. The primary atoms dislodged by the incident particle can have sufficient energy to kick-off other atoms before they settle. As a result, an area with a high density of defects is created. Some of the lattice defects generated by incident neutrons are shown in Fig. 2.6. This non-ionizing energy loss (NIEL) process leads therefore to the creation of defects in the crystal which change the electrical properties of the material.

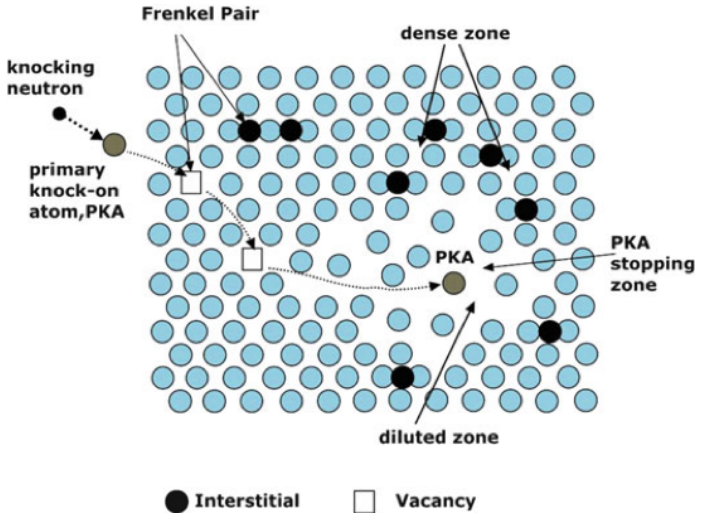


Figure 2.6: Example of defects in silicon generated by impinging high-energy neutrons [56].

Different types of particles interact differently with matter. In order to be able to easily compare the damage caused by particles of different types and energies, the NIEL damage is normalized to the damage caused by 1 MeV neutrons and described in terms of neutron equivalent fluence ($1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$). The defects created in the silicon due to NIEL cause the introduction of energy levels in the bandgap which can result in several negative effects [9]:

- the radiation-induced energy levels near the middle of the bandgap act as generation/recombination centres and can cause a considerable increase in thermal generation rates. The sensor leakage current therefore increases proportionally to the NIEL damage.

- the charge generated by a traversing particle can get trapped by the defect levels and be released after some time or even recombine. The carrier lifetimes and diffusion lengths therefore decrease with NIEL damage and so does the detectable sensor signal. In order to improve the NIEL tolerance of sensors, the generated charge needs to be collected as quickly as possible to reduce the probability for it to be captured by radiation-induced traps and get lost for detection.
- the doping concentration of the sensor bulk is influenced by donor- or acceptor-like states. The sensor doping therefore changes as a function of the radiation fluence and for sufficiently high fluences can even be inverted.

These effects can be partially counteracted with the annealing, i.e. recombination and re-arrangement of defects over time. After annealing, in fact, several material properties, such as carrier lifetime, diffusion length or leakage current can experience a significant recovery. These processes are strongly dependent on the temperature [57].

2.2 Hybrid pixel sensors

In hybrid pixel sensors, the sensitive layer and the readout electronics are in separate entities. The pixellated sensor matrix is produced in a specialized sensor material. The readout is manufactured with a standard CMOS process in a chip with pixel cells arranged in the same bi-dimensional structure as the sensor matrix. Sensor and readout chip are then connected in every pixel using flip-chip and bump-bonding techniques [59], as shown in Fig. 2.7. The main advantage of hybrid pixels is that the sensor and the readout chip can be optimized separately.

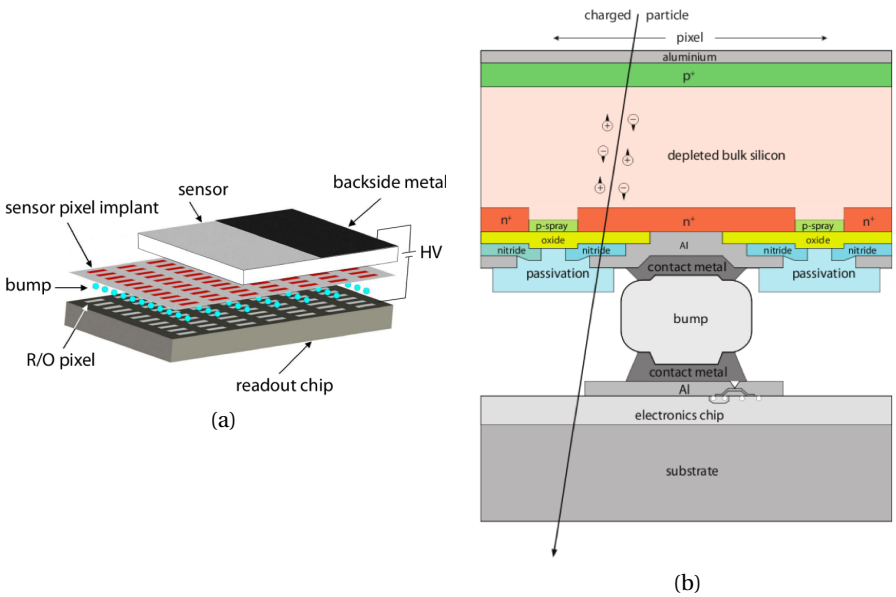
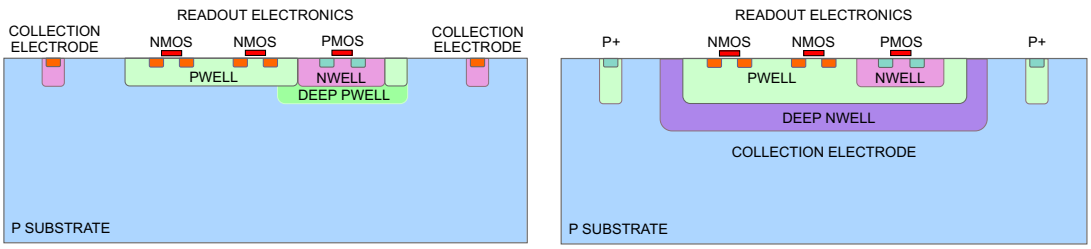


Figure 2.7: Hybrid pixel sensor: a) cut-away view and b) single pixel cell [58].

The sensor is optimized for fast charge collection and thus high radiation tolerance. The readout electronics is also optimized for radiation tolerance and to be fast and low noise. This separate optimization renders hybrid pixel sensors the state-of-the-art technology for rate capability and radiation tolerance. In fact, they are used in the most extreme radiation environments, such as the innermost pixel layers of the ATLAS and CMS experiments [60], [61]. Despite its good performance, the hybrid approach does have its disadvantages. These are mainly related to the bump-bonding technique: the bump bonds represent typically a limit to the minimum pixel size and increase the sensor capacitance, resulting in a larger power consumption for the same performance and therefore in a larger material budget. Furthermore, the module assembly with the bump-bonding and flip-chip techniques is a complex and expensive process for large-area detectors.

2.3 Monolithic active pixel sensors

Monolithic active pixel sensors (MAPS) integrate sensor and readout electronics into the same chip. These are a standard for detection of visible light. In HEP, they represent a cost-effective alternative to hybrid sensors as they do not require the complex and time-consuming bump-bonding and can be produced in commercial CMOS technologies. Thanks to the lack of bump-bonding, they offer the possibility to realize smaller pixels and a lower pixel capacitance, beneficial for a reduction of the power consumption and therefore material budget. A reduction of the material budget is however already obtained thanks to the inherently thinner sensor module which is formed by a single layer. Fig. 2.8a shows the cross-section of a possible MAPS implementation. The collection electrode is an n-well in a p-type substrate. The readout electronics is placed next to the collection electrode and to be able to use full CMOS circuits while avoiding competition in charge collection between the n-well of the PMOS transistors and the one of the collection electrodes, a deep p-well is used to shield the PMOS transistors from the substrate. With visible light the charge is generated within a depth of a few microns whereas high-energy particles generate charge over the full thickness of the sensor which needs to be collected as quickly as possible for high NIEL tolerance. Therefore, collection by drift has to be enhanced with respect to collection by diffusion and a reverse bias is typically applied to the sensor to increase the depletion volume. However, as the collection electrode is



(a) (b)

Figure 2.8: Cross section of a MAPS: a) small and b) large collection electrode design.

small to achieve a low sensor capacitance and complex readout circuits require large areas, it is difficult to obtain full depletion of the p-type substrate over the entire pixel area. An alternative approach is to place the readout electronics inside the collection electrode as in Fig. 2.8b. Full depletion of the p-type substrate and a more uniform field over the entire pixel area, beneficial for the radiation tolerance, can thus be obtained. This approach is called large collection electrode design, as opposed to the one in Fig. 2.8a referred to as small collection electrode design. The large collection electrode features a larger capacitance with values comparable to the ones in hybrid pixel sensors leading to larger power consumption. Furthermore, the bulk of all NMOS transistors is capacitively coupled to the sensor and the noise in their p-well caused by in-pixel digital activity can cause undesired signals on the collection electrode. To reduce the crosstalk to acceptable levels, it is necessary to give special attention to the design of the readout circuitry. On the other hand, the small collection electrode design better decouples the readout circuits from the collection node and the digital crosstalk is drastically lower.

It is worth mentioning that, due to the advancement in assembly technologies which led to techniques such as microbumps, through-silicon vias and wafer stacking, the distinction between hybrid and monolithic pixel sensors is becoming more and more vague.

2.4 Readout electronics

The charge signal induced on the collection electrode by the traversal of a particle is quite small (about 60 e/h pairs per µm of sensitive layer [51]). Signal amplification is therefore typically required. In sensors for HEP, the required functionality is often more complex than for many sensors for visible light, resulting in more in-pixel circuitry. The block diagram of the readout chain typically implemented for HEP applications is shown in Fig. 2.9. The chip is typically divided into an active area, which is the repetitive matrix of identical pixels, each equipped with a dedicated circuitry, and the periphery, which controls the active matrix and processes the hit data it generates. In each pixel, the collection electrode is connected to an amplification stage. The latter is typically a charge-sensitive amplifier (CSA) obtained by a high-gain element with a capacitive feedback which generates a voltage proportional to the collected charge. A shaper, essentially a band-pass filter with tuneable bandwidth, usually follows the amplifier to improve the SNR and reduce the probability of pile-up. The filtered amplifier output is then

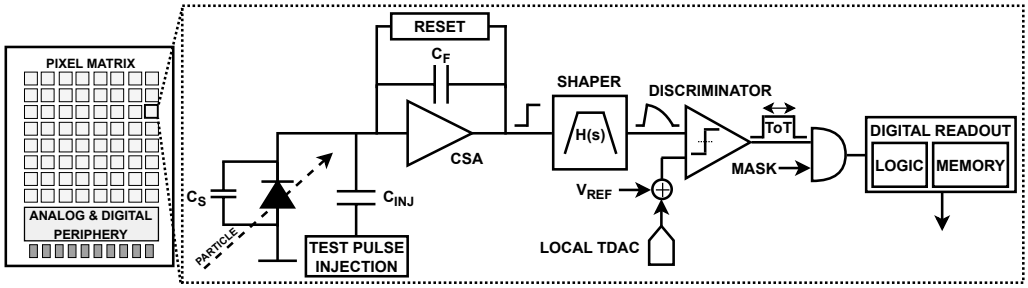


Figure 2.9: Typical readout scheme used in pixel sensors.

compared to a threshold voltage with a discriminator. When the analog signal overcomes the threshold, the discriminator generates a pulse and the pixel is considered to have "fired". In-pixel digital circuitry further processes the data and transmits the addresses of the fired pixels to the periphery where they are sent off-chip. The pixel circuitry frequently includes additional features for testability and ease of operation. The possibility of capacitively injecting a test pulse to the input of the amplifier is typically implemented to test the readout chain without a particle beam. Another common feature is the possibility of masking a pixel, i.e. disabling its output in case it generates an excessive noise hit rate because of e.g. a broken circuit or a defected sensor cell. The simplest way to assign an address to a pixel is to connect the discriminator outputs of each pixel directly to the periphery and assign the address there. For large matrices, this approach is soon limited by the routing resources. To reduce the required connections, pixels can be grouped to implement a parallel transfer of the hits from the pixels to the periphery [62]. In common readout approaches, however, the addresses are assigned in the pixels and transmitted to the periphery according to a priority scheme. The latter can be implemented for example with a column-based priority encoder [63] or a token logic, where a token signal, which allows the pixels to write on a shared bus, passes from pixel to pixel [64], [65]. In HEP experiments, the hits are sparse over the pixel sensors and have a low occupancy (< 1 % of pixels are hit). Furthermore, most of the data do not contain interesting physics events. To reduce the required output bandwidth and the amount of data to process, a selection of potentially relevant events is made by other systems. A trigger signal is then sent to the sensors and only the triggered events are sent off-chip. The readout thus needs to be able to store the hits until a trigger decision selects the ones to read out. The main performance metrics for the readout channel of a pixel sensor are:

- detection efficiency: for a too high charge threshold, particles releasing a small charge might not be able to trigger the discriminator and thus be detected. The fraction of the particles traversing the sensor which are actually detected is the detection efficiency.
- equivalent noise charge: noise from the sensor and the electronic devices result in voltage fluctuations which limit the minimum detectable charge, as explained in the next point. Being the first stage of the readout chain, the charge amplifier is the most crucial element in terms of noise. The noise of a sensor readout channel is quantified with the equivalent noise charge (ENC), defined as the fluctuation of the input charge required to cause the voltage noise observed at the output.
- threshold and threshold dispersion: the discriminator threshold has to be as low as possible to be able to detect hits even with a very small collected charge and thus obtain high detection efficiencies. Due to the components' mismatch, the threshold differs from pixel to pixel. The threshold dispersion and noise determine the minimum reliable threshold setting. Indeed, for a too low threshold, the discriminator might be triggered by noise fluctuations and the detector be overwhelmed by noisy hits. If need be, in addition to the global setting, the possibility to fine tune the threshold per pixel is implemented to improve its uniformity and achieve lower thresholds.

- time response: the sensor and readout are required to be fast enough to be able to associate events to a particular bunch crossing. The time response is made up of different contributions such as charge collection time, amplifier rise time, shaper bandwidth, discriminator speed and digital logic delay to output the data. The time for the amplifier output to reach the discriminator threshold, which is typically the dominating contributor, decreases for larger amplitudes and thus input charges. This dependence is called time walk. The lowest input charge providing the required time response is called in-time threshold.

- hit rate: the time required to process a hit is referred to as dead time, as new hits arriving during this time might be lost. For instance, in the case of the charge amplifier, the dead-time is the recovery time required to restore its DC condition after a hit. The dead-time can be a limiting factor to the maximum hit rate the readout can process.

- power consumption: as mentioned in the first chapter, a lower power consumption reduces the detector material related to powering and cooling and so the particle scattering which deteriorates the detector accuracy. For this reason, the previous figures have to be obtained with the least possible power.

- radiation tolerance: for good tolerance to NIEL, the system needs to be able to cope with a wide range of sensor leakage currents. Furthermore, the readout circuitry has to be resilient to the effects of ionizing radiation which alters the transistors' characteristics as threshold and transconductance (more details in section 2.4.4).

2.4.1 Noise sources

Electronic noise is the result of stochastic fluctuations in the number and velocity of charge carriers. Noise is typically described by means of its power spectral density. Depending on the physical processes, the following noise sources can be identified:

- thermal noise: it comes from the thermal excitation of charge carriers which induces fluctuations in a device current. The power spectral density of a current flowing through a conductor with resistance R and temperature T is [66]:

$$d < i^2 >_{therm} = 4k_B T \frac{1}{R} df, \tag{2.15}$$

  which is independent of the frequency (white noise) and current. The thermal noise of a transistor depends on its biasing and size and can be expressed as a voltage in series with its gate whose power spectral density is [67]:

$$v^2_{n,th} = 4k_B T \frac{2}{3} \frac{1}{g_m} df, \tag{2.16}$$

  where g_m is the transconductance of the transistor.

- shot noise: it is caused by statistical fluctuations of carriers emitted independently of each other over a potential barrier, as in a p-n junction. Its power spectral density is [66]:

$$d < i^2 >_{shot} = 2qI_0 df, \tag{2.17}$$

where I_0 is the mean current flowing through the barrier. The shot noise is therefore directly proportional to the current and has a white noise spectrum.

- flicker noise: the defects at the interface between the silicon and the gate oxide are responsible for the capture/release of charge. The fluctuations of carriers due to these defects lead to flicker noise [68]. The power spectral density of the flicker noise has a $1/f$ behaviour and it can be expressed as a voltage source in series to the transistor gate with spectrum [69]:

$$d < v^2 >_{flicker} = \frac{k_f}{C_{ox}^2 WL} \frac{df}{f}, \tag{2.18}$$

where k_f is a constant that depends on the specific process, C_{ox} is the gate oxide capacitance per unit of area and W and L are the transistor width and length respectively.

2.4.2 Charge sensitive amplifier

The most commonly used front-end amplifier in pixel sensors is the charge-sensitive amplifier (CSA). The term "charge-sensitive" refers to the fact that the amplifier generates a voltage signal proportional to the input charge. The basic principle of a CSA is shown in Fig. 2.10 and it is realized by means of a high-gain inverting amplifier with a capacitive feedback. The sensor is modelled with a current source I_{IN} in parallel with a capacitance C_S. The nodal equation at the input node reads:

$$I_{IN} + V_{IN} sC_S + (V_{IN} - V_{OUT}) sC_F = 0. \tag{2.19}$$

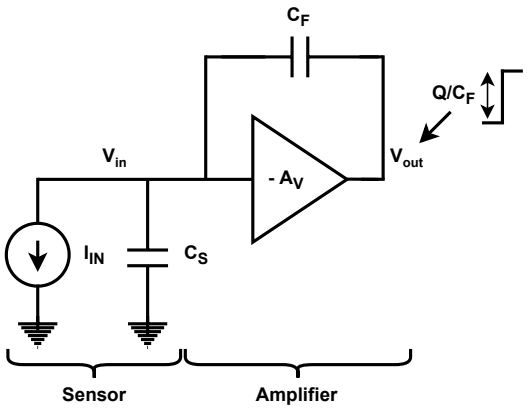


Figure 2.10: Basic principle of a CSA.

Combining this equation with $V_{OUT} = -A_V V_{IN}$, the output voltage is:

$$V_{OUT} = -\frac{A_V I_{IN}}{s[C_S + (1 + A_V)C_F]}. \tag{2.20}$$

If the inverting amplifier has a sufficiently large gain such that $(1 + A_V)C_F >> C_S$, the previous expression can be approximated with:

$$V_{OUT} \approx -\frac{I_{IN}}{sC_F} \rightarrow V_{OUT} = -\frac{Q}{C_F}. \tag{2.21}$$

This relationship shows that the response of a CSA to a current signal is a voltage step proportional to the input charge, as anticipated, and the inverse of the feedback capacitance. As long as the gain of the core amplifier is sufficiently large, the gain of the CSA is independent of the sensor capacitance and any other parameter. In order for the CSA to be able to process subsequent events, the feedback capacitance needs to be discharged. This is typically implemented with a resistor in parallel. In addition, a buffer usually follows the core amplifier to prevent the feedback network from loading it and reduce its gain. A more comprehensive scheme of a CSA is shown in Fig. 2.11. A circuitry, not shown in the figure, is generally used to bias the sensor and compensate its leakage. As the latter varies with temperature and irradiation, leakage compensation schemes are typically based on negative feedbacks which adjust a controllable current source. The bandwidth of this circuit should be very small to react only to slow signals, such as leakage, without affecting the fast currents induced by particles.

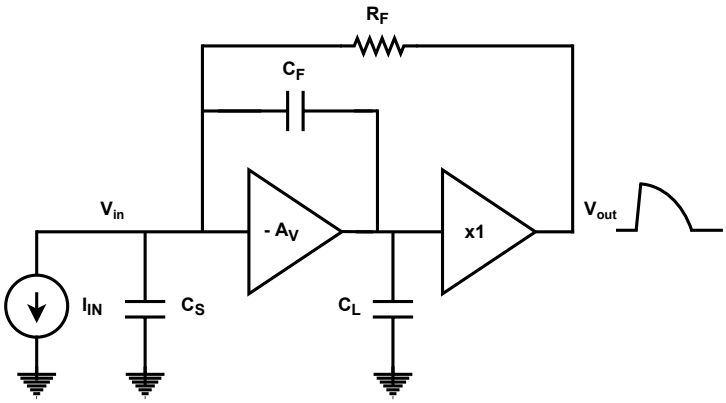


Figure 2.11: CSA with reset resistor and output buffer.

A study of the system in Fig. 2.11 leads to a transimpedance gain equal to [70]:

$$\frac{V_{OUT}}{I_{IN}} = -\frac{R_F}{1 + sC_F R_F + s^2 R_F \frac{\zeta}{g_m}} \quad \text{with} \quad \zeta = C_S C_L + C_S C_F + C_L C_F, \tag{2.22}$$

where g_m is the transconductance of the input transistor of the core amplifier. The denominator of the gain is a second-order polynomial and can be expressed as $(1 + s\tau_{rise})(1 + s\tau_{fall})$ if the poles are real and widely separated, i.e. one time constant is much larger than the other

($\tau_{fall} >> \tau_{rise}$). Under these assumptions, the time constants can be expressed as:

$$\tau_{rise} = \frac{C_S C_L + C_S C_F + C_L C_F}{g_m C_F} \tag{2.23}$$

$$\tau_{fall} = R_F C_F. \tag{2.24}$$

The time response of the CSA to a current pulse is therefore:

$$V_{OUT}(t) = -\frac{Q}{C_F}\frac{\tau_{fall}}{\tau_{fall} - \tau_{rise}}(e^{-\frac{t}{\tau_{rise}}} - e^{-\frac{t}{\tau_{fall}}}), \tag{2.25}$$

whose maximum, occurring at the peaking time $\tau_p = \frac{\tau_{rise}\tau_{fall}}{\tau_{rise}-\tau_{fall}} \ln \frac{\tau_{rise}}{\tau_{fall}}$, is equal to:

$$V_{OUT,peak} = \frac{Q}{C_F}(\frac{\tau_{fall}}{\tau_{rise}})^{\frac{\tau_{rise}}{\tau_{rise}-\tau_{fall}}}. \tag{2.26}$$

The obtained output step voltage is the one obtained with the ideal CSA multiplied by a term which depends on the time constants. In particular, the peak output voltage is close to the ideal value Q/C_F for large ratios between the discharging and rise time constants, e.g. it is 95.4 % of it for $\tau_{fall}/\tau_{rise} = 100$. To obtain high gains, a small feedback capacitance C_F is thus required. This, however, reduces the discharging time constant τ_{fall} and also the phase margin of the loop [70]. To counteract the lower τ_{fall} and increase the phase margin, the transconductance g_m of the input transistor has to be increased which leads to larger power consumption. This is beneficial also for the noise performance. The ENC is indeed given by [70], [71]:

$$ENC^2 = (\frac{a}{\tau} + b)C_{DET}^2 + c\tau. \tag{2.27}$$

In this equation τ is the peaking time of the amplifier or, if present, the time constant of the subsequent shaping filter whereas the terms a, b and c express the noise of the devices in the system. In particular, the first two terms a and b represent the input series thermal and flicker noise of the core amplifier respectively, typically dominated by the input transistor noise which scales down with its transconductance g_m and area. The term c represents instead the input parallel shot noise which, with proper sizing of the feedback resistor R_F, is dominated by the noise generated by the sensor leakage current. A significant point to notice is that the effect of the series noise, both thermal and flicker, is proportional to the sensor capacitance. In order to achieve better performance and lower power consumption, sensors offering a smaller capacitance are thus preferable. This concept is further elaborated in the following section.

2.4.3 Importance of high Q/C

The ratio between the collected charge and the sensor capacitance is an important parameter which has significant consequences on many design aspects, particularly on noise and power consumption. The main aim in the charge signal processing is to minimize the ENC or, which is the same, maximize the SNR. In this context, the SNR can be evaluated comparing the

input signal resulting from the ionization of a high-energy particle to the input-referred RMS noise of the amplifier. The latter is typically dominated by the thermal noise of the input transistor, inversely proportional to the square root of its transconductance g_m. The amplitude of the input signal is instead given by the ratio between the collected charge and the sensor capacitance, i.e. the said Q/C ratio. Assuming that the amplifier power consumption P is dominated by the current in the input branch, as it is usually the case, the SNR can be written as [72]:

$$SNR \approx \frac{Q}{C}\sqrt{\frac{3g_m}{8k_BT}} \propto \frac{Q}{C}\sqrt{g_m} \propto \frac{Q}{C}\sqrt[m]{P}, \tag{2.28}$$

with $m = 2$ if the input transistor works in weak inversion, where the transconductance g_m is proportional to the biasing current, or $m = 4$ if the input transistor works in strong inversion, where the transconductance g_m is proportional to the square root of the biasing current. Rearranging for a fixed SNR and bandwidth:

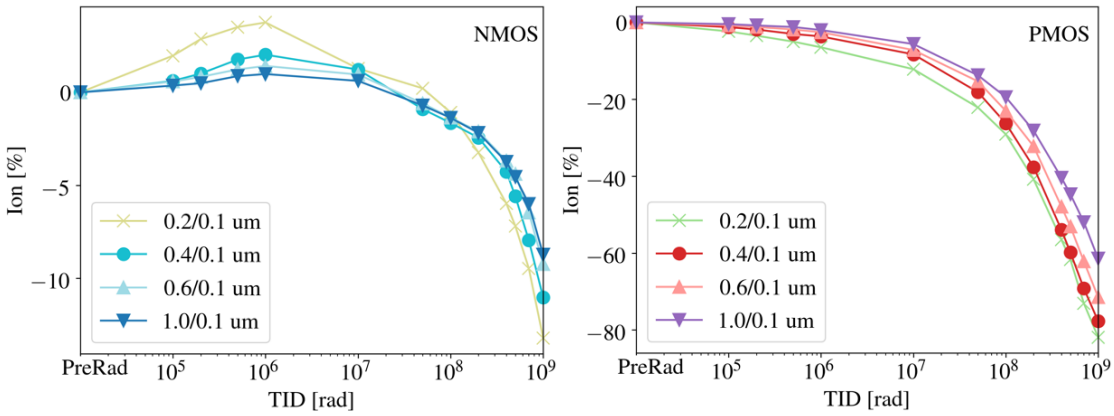
$$P \propto (\frac{Q}{C})^{-m}, \tag{2.29}$$

with $2 \leq m \leq 4$ depending on the operating point of the input transistor. This expression basically states that a high Q/C ratio is key to reduce the power consumption for a given SNR and bandwidth, i.e. for the same analog performance.

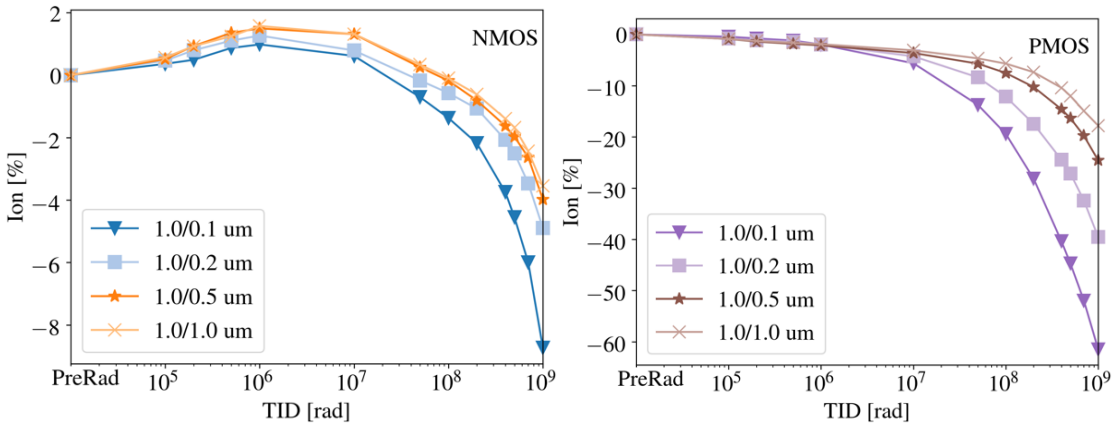
2.4.4 Radiation effects on the electronics - Ionizing radiation

The readout electronics is mainly affected by ionizing radiation. The latter indeed damages the oxides present on the surface of CMOS circuits, such as gate oxides, shallow trench isolations (STI) and gate spacers. The damage depends on the received dose and it is therefore described by a quantity called Total Ionizing Dose (TID), measured in units of Gy, or more commonly of rad. Ionizing radiation generates charge carriers within the oxides and, as in SiO_2 the hole mobility is orders of magnitude lower than the electron mobility, electrons escape more easily leaving behind positive charges which accumulate over time [73]. The build-up of positive charge in the gate oxide causes a shift in the threshold voltage of transistors: for NMOS transistors the threshold voltage decreases, which leads to an increase of the leakage current, whereas for PMOS transistors the threshold voltage increases, which leads to a reduction of their conductivity. Ionizing radiation also causes the formation of dangling bonds at the Si-SiO_2 interface that act as traps [73]. In NMOS transistors, negative charge gets captured in the interface traps partially compensating the effect of the positive charge in the oxide. However, the time scales of the two mechanisms is different. The same effects take place also in the STI oxides. While they cause PMOS transistors to be less conductive, in NMOS transistors, the positive charge in the STI can create a lateral conductive path in the channel even when the main transistor is off, again causing an increase of the leakage current. As for small widths the percentage of channel influenced by the STI is higher, these effects are more pronounced

for narrow transistors and are so called Radiation-Induced Narrow Channel Effects (RINCE) [74]. Likewise, short transistors are also more affected by TID effects, referred to in this case as Radiation-Induced Short Channel Effects (RISCE) [75]. These are mainly caused by the gate spacers: the effects of ionizing radiation on these oxides influence the source and drain diffusions and increase their series resistance. This phenomenon can effectively be modeled as an increase of the channel length and short transistors are therefore proportionally more affected by it. As an example, TID effects on transistors with different sizes in the Tower Partners Semiconductor Co. 65 nm imaging technology are shown in Fig. 2.12. The most affected structures are the ones with smaller dimensions. As for the NIEL-induced damage, annealing can be useful in reversing the effects of TID.



(a)



(b)

Figure 2.12: Variation of I_{ON} ($V_{GS,DS} = 1.2V$) with TID for NMOS and PMOS transistors with (a) a length of 0.1 µm and different widths (b) a width of 1 µm and different lengths [76].

For high tolerance to TID, transistors with minimum sizes have to be avoided. A common technique to improve the TID tolerance of a transistor is to design it with an enclosed gate, as shown in Fig. 2.13. This structure is called Enclosed Layout Transistor (ELT) [77] and its geometry ensures that the source-to-drain current can only flow below the gate and eliminates the possibility of leakage paths along the channel edge. In order to avoid inter-device leakage,

instead, contact rings around the transistors can be used as in the figure. The main disadvantage of the ELT is the larger area it requires for the same channel dimensions. Modern CMOS technologies are inherently more and more tolerant to TID thanks to their thin gate oxides, as the tunneling effect that allows positive charge to escape becomes more prominent [78].

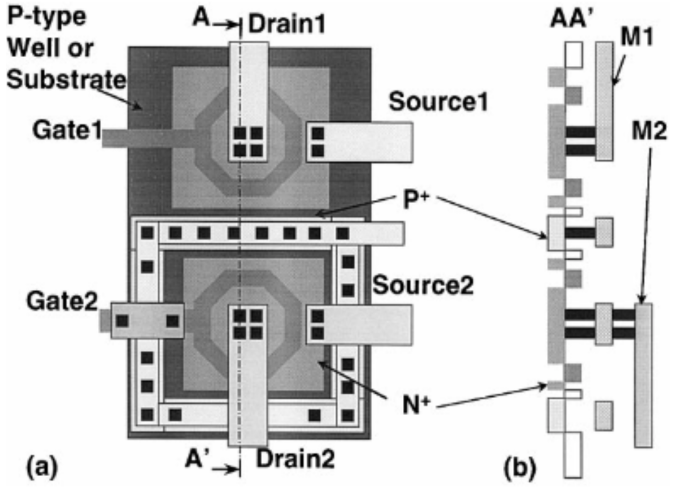


Figure 2.13: (a) Layout and (b) cross-section of two enclosed layout NMOS transistors with a p+ contact guard ring which prevents inter-device leakage [77].

Apart from the cumulative radiation effects described so far, charge deposition by an ionizing particle can cause single-event effects (SEE) [79]. A single-event upset (SEU) occurs if the deposited charge causes a bit value stored in a memory to be flipped, corrupting the stored information. The risk of SEU is typically mitigated by triplicating the critical memory cells and implementing a majority voter such that two uncorrupted bits overrule an upset one. Moreover, if the deposited charge is large enough, it can trigger the intrinsic thyristor structure present in the CMOS well structure to conduct a large current, which can cause damage or even destruction of the circuitry. This process is called single-event latchup (SEL) and can be prevented with a frequent placement of well contacts in order to obtain fast discharging of the wells.

3 Developments in the TowerJazz 180 nm imaging technology

The use of MAPS has been restricted thus far in low radiation environments. Recent advancements in CMOS sensor processing are expected to improve the radiation hardness of MAPS and make it a viable option also for the most extreme radiation environments. Indeed, monolithic sensors were proposed as an alternative to hybrid sensors in the outer pixel layer of the ATLAS ITk where the cost advantage would be significant due to the extensive area that needs to be covered. The interest was raised in particular by the ALPIDE chip [32] in the TowerJazz 180 nm imaging technology, the sensor developed for the current ALICE ITS. Significant effort was invested in optimizing this process and make it suitable also for more demanding applications in terms of radiation hardness. The encouraging results obtained prompted the development of monolithic sensors for harsher environments. One of these developments is MALTA (Monolithic from ALice To Atlas). The MALTA matrix features a sensor with a small collection electrode, an open-loop charge-sensitive front-end amplifier and a fast, low-power, asynchronous digital readout architecture [80]. Measurements on the chip showed a timing response within the specifications [81]. However, the detection efficiency was degraded at the pixel edges already after a NIEL fluence of $10^{14}\,1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$ [82] due to a weak lateral electric field in these regions and a large RTS noise which prevented the chip from being reliable at low thresholds. The process was therefore further optimized to enhance the lateral electric field and the modifications have been successfully verified in a small-scale prototype called Mini-MALTA [83]. The sensor modifications were then implemented in a subsequent large-scale prototype, called MALTA2, along with modifications on the front-end to reduce its noise.

This chapter focuses on the optimization of the MALTA2 front-end and its characterization. In particular, section 3.1 introduces the standard sensor in the TowerJazz 180 nm imaging technology and the implemented modifications to improve its radiation tolerance. Section 3.2 gives a detailed explanation on the design and optimization of the MALTA2 front-end circuit whereas section 3.3 a general overview of the chip. Finally, section 3.4 shows the main characterization results of the front-end. Part of the work presented in this chapter was published form the author on the following paper:

[84] F. Piro et al., "A 1-μW Radiation-Hard Front-End in a 0.18-μm CMOS Process for the MALTA2 Monolithic Sensor", in *IEEE Trans. Nucl. Sci.*, vol. 69, no. 6, pp. 1299-1309, 2022. DOI: 10.1109/TNS.2022.3170729.

3.1 Sensor technology

3.1.1 The standard TowerJazz 180 nm process

The TowerJazz 180 nm CMOS imaging process is a quadruple well technology originally designed for CMOS camera applications. The cross section of the sensor developed in this process is shown in Fig. 3.1. It implements a small collection electrode, defined by an n-well implant, which is located inside the sensing volume. The foundry offers the possibility to use different starting materials which makes it particularly interesting for HEP applications. Indeed, high-resistivity p-type epitaxial layers help to enhance the depletion around the collection electrode. The in-pixel circuitry is placed next to the collection electrode and it is shielded from it by a deep p-well, which avoids collection of the signal charge by parts of the circuit other than the designated collection electrode. The key features of the technology are therefore the deep p-well, which allows full CMOS in-pixel circuitry, and the possibility to use different starting materials compatible with particle detection. For visible light the signal is generated within a depth of a few microns. As mentioned in the previous chapter, charged particles generate instead carriers over the full thickness of the silicon (≈ 60 e/h pairs per µm traversed [51]) which needs to be collected well within the time response required by the experiment for the event reconstruction. A fast charge collection is beneficial also for a high tolerance to NIEL damage. For this reasons, as opposed to the typical usage in CMOS camera applications, the depletion volume within the sensor is increased by increasing the

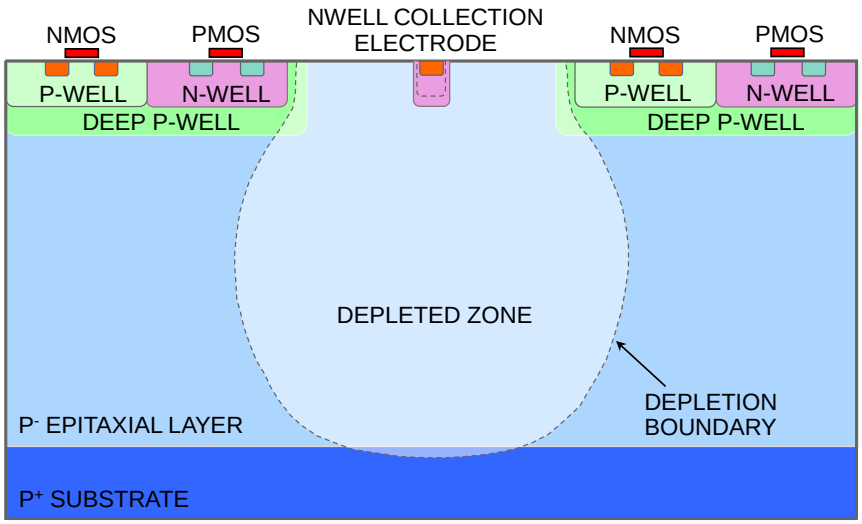


Figure 3.1: Cross-section of the standard sensor in the TowerJazz 180 nm process [85].

reverse bias between the collection electrode and the surrounding p-well and p-type substrate. However, even with a reverse bias, due to the area needed for the readout circuitry, it is difficult to deplete the regions along the pixel edges and achieve full depletion of the sensitive layer. Most of the signal charge is thus generated outside the depleted area and collected mainly by diffusion (with a collection time of ≈ 100 ns). For the ALPIDE development, this sensor provided a NIEL tolerance of up to $10^{13}\,1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$ [32], sufficient for the modest ALICE requirements.

3.1.2 The modified TowerJazz 180 nm process

In the TowerJazz 180 nm modified process, a uniform ion-implanted low-dose n- layer is added under the deep p-well containing the circuitry and covers the entire matrix/pixel area. The cross section of the sensor in the modified process is shown in Fig. 3.2. The modification creates a planar junction deep in the epitaxial layer and the depletion extends immediately over the full pixel area. If sufficiently low dose, the n- layer is fully depleted up to the n-well of the collection electrode already for a low sensor reverse bias and introduces only a small penalty on the sensor capacitance. This also guarantees that the collection electrodes in the matrix are mutually isolated. The n- layer also separates the p-wells containing the electronics from the substrate and allows them to be biased independently. Importantly, the process modification entails the addition of only an extra mask and therefore has a small impact on the manufacturing process and cost overhead. Furthermore, it can be implemented without requiring any changes in the layout of the sensor or of the circuitry.

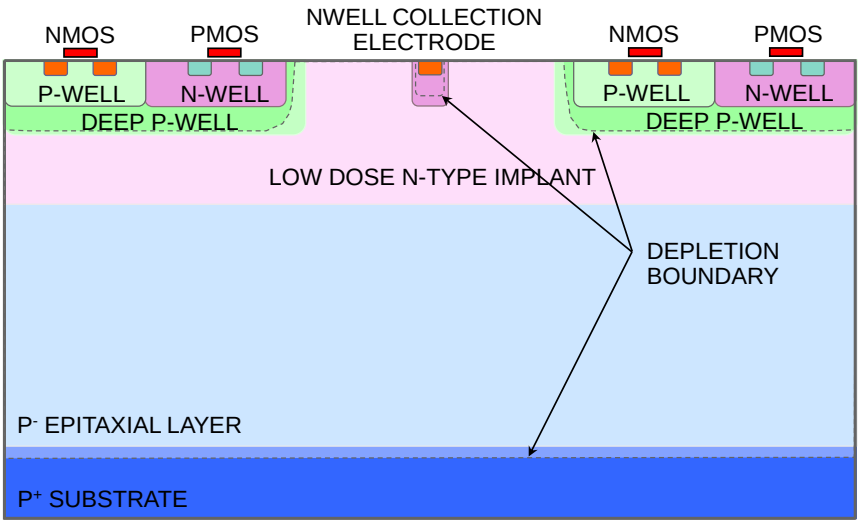
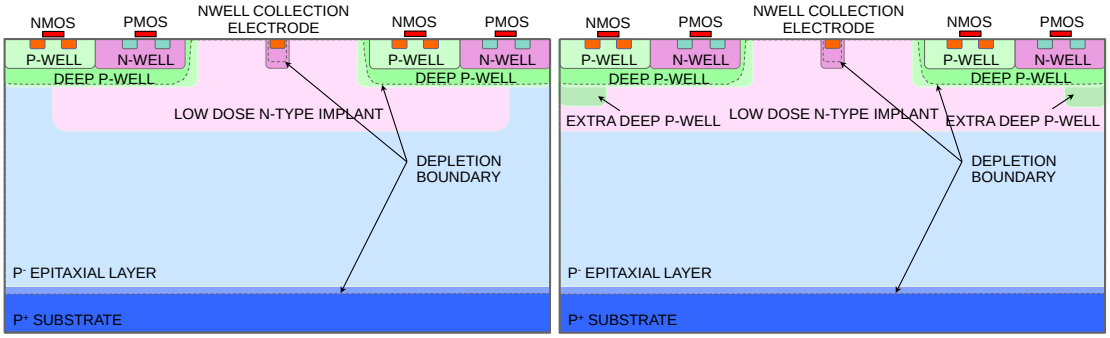


Figure 3.2: Cross-section of the sensor in the TowerJazz 180 nm modified process [85].

While this process modification allows full depletion of the sensitive layer, the charge collection can be further accelerated by enhancing the lateral component of the electric field in the epitaxial layer which would accelerate the charge towards the collection electrode. This is beneficial especially in the pixel corners and along the pixel edges where the path of the

carriers and thus their collection time is longer. The lateral electric field in these regions can be increased and the charge collection strongly accelerated by introducing a junction along the sensor depth. This can be achieved either by patterning the n- layer, i.e. removing it along the pixel edges, or with the introduction of an extra deep p-well implant [86]. The cross section of the sensor for both solutions is shown in Fig. 3.3. In the case of the gap in the n- layer, only a change in the mask of its implant is required. For the additional extra deep p-well, instead, an additional mask is needed, but this implant is already available in the foundry. Therefore, also these additional modifications have a minimal impact on the sensor fabrication.



(a) (b)

Figure 3.3: Cross-section of the sensor in the TowerJazz 180 nm modified process with: (a) gap in the low-dose n- implant (b) extra deep p-well [83].

Regarding tolerance to ionizing radiation, transistors in this technology showed a good tolerance to TID thanks to their thin gate oxide (3 nm) [87]. In the case of the chip developed and tested in this work, the pixels implement the sensor from Fig. 3.3b. The epitaxial layer is 30 µm thick and has a resistivity $\gtrsim 1\,\mathrm{k}\Omega\mathrm{cm}$. The collection electrode is an octagonal-shaped n-well with a diameter of 2 µm, distanced 4 µm from the surrounding p-well containing the circuitry. The extra deep p-well is 1.2 µm wide and centered along the pixel edges. This geometry is the result of a trade-off between a small sensor capacitance (<5 fF) and a large lateral electric field [88], with collection times in the nanosecond range [86].

3.2 Analog front-end

The analog front-end circuit designed for the developments in the TowerJazz 180 nm imaging technology is a continuously active circuit which performs the reset of the collection electrode, the amplification of the collected charge, the shaping of the analog signal and the digitization of this signal through a discrimination stage. The basic principle of the amplification stage is illustrated in Fig. 3.4. The continuous reset mechanism uses the diode D1 to hold the collection electrode voltage. The latter is set at ≈ 1 V for proper operation of the front-end by tuning the V_{RESET} bias according to the sensor leakage current. As the circuit is DC coupled to the sensor, to increase the sensor reverse bias, the voltage of the p-type substrate and p-well containing the circuitry is decreased. In particular, these voltages can be biased down to −6 V,

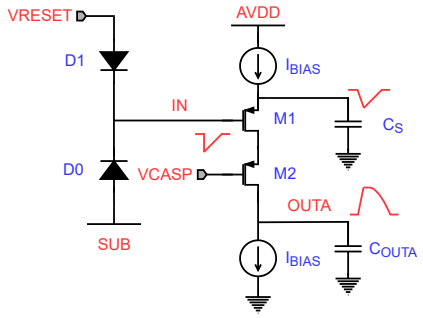


Figure 3.4: Principle of the front-end amplifier.

allowing to achieve a reverse bias of the sensor of ≈ -7V. When no charge is collected by the collection electrode, D1 is biased by the leakage current of the sensor diode D0. Upon a particle crossing, the generated charge is collected by the collection electrode and integrated onto its capacitance. A negative voltage step with an amplitude of $\Delta V = Q/C$ is thus generated on it. This causes the reset diode to conduct more current and to slowly charge the input node back up to its original value, which can take several hundreds of µs. The reset diode is implemented with a small p+ implant in the n-well of the collection electrode, adding only a small contribution to the sensor capacitance whose low value needs to be preserved for good performance. In contrast to a conventional CSA typically used in these applications, the proposed architecture implements a voltage amplifier. In this case, indeed, reducing the feedback capacitance of a CSA to a much smaller value than the sensor capacitance to avoid a noise penalty can be challenging. Furthermore, this would make it also typically lower than the output capacitance, degrading the speed of the circuit (see equation 2.23). To profit from the low sensor capacitance and overcome the aforementioned limitations, the proposed architecture integrates the generated charge on the sensor itself and processes the obtained signal with an open-loop voltage amplifier, resulting in a simpler and more power-efficient solution. The input node (gate of the transistor M1) is connected directly to the collection electrode. The input transistor M1 acts as a source follower and, when the input voltage drops because of the collected charge, forces its source to follow transferring charge from the capacitance C_S to the output node capacitance C_{OUTA}. Ideally, for the voltage on OUTA, one can write:

$$\Delta V_{OUTA} = \frac{Q_S}{C_{OUTA}} = \frac{C_S \cdot \Delta V_{IN}}{C_{OUTA}} = \frac{C_S}{C_{OUTA}} \frac{Q_{IN}}{C_{IN}}. \tag{3.1}$$

A large gain is therefore obtained for $C_S \gg C_{OUTA}$. The overall effective sensor capacitance is the sum of the sensor junction capacitance, the reset diode parasitic capacitance, the input line and the input transistor gate capacitance. After settling, the follower action on the source of the input transistor reduces the contribution of its gate-source capacitance to the total capacitive load on the electrode. Furthermore, the cascode transistor M2 reduces the gain (impedance) on the drain of the input transistor and mitigates the Miller effect on its gate-drain capacitance. A more practical implementation of the circuit is shown in Fig. 3.5a. Since

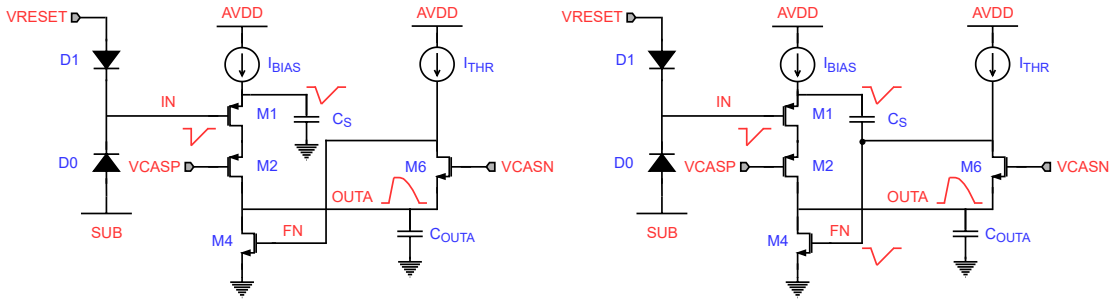


Figure 3.5: Principle of the front-end amplifier: (a) with feedback to adjust the operating point (b) with additional gain mechanism.

the two I_{BIAS} current sources are difficult to match, a low-frequency feedback which sets the operating point of the transistor M4 is introduced: its gate voltage is now adjusted for it to sink $I_{BIAS} + I_{THR}$, where I_{THR} is a small fraction of the main biasing current I_{BIAS}. The introduced branch also defines the DC voltage of the amplifier output node and its return to baseline. Upon a particle hit, when the voltage on OUTA rises, the gate-source voltage of the transistor M6 reduces, forcing I_{THR} to charge up the gate of the transistor M4, discharging OUTA and bringing it back to its baseline value. An additional gain mechanism is introduced by connecting the capacitance C_S to the gate of the transistor M4, as done in Fig. 3.5b. A part of the signal on the input transistor source is now transferred to the gate of the transistor M4, which behaves as a common-source device. In this scheme, the capacitance C_S plays an important role not only for the gain of the amplifier but also in determining its return to baseline since it is connected to the feedback node (FN), i.e. the gate of the transistor M4. A larger I_{THR} increases the speed of the feedback loop, resulting in a faster return to baseline, but could also provide an excessive filtering at low frequencies on the gate of the transistor M4, reducing the amplifier gain.

As will be shown later, the circuit behaviour is non-linear. However, a small-signal analysis helps to gain insights into its operation. The amplifier small signal model is shown in Fig. 3.6. In the circuit, only the relevant components have been considered. Furthermore, an ideal unity gain buffer is interposed between the source of the input follower, loaded with the equivalent impedance seen on its source, and the capacitance coupling to the gate of the

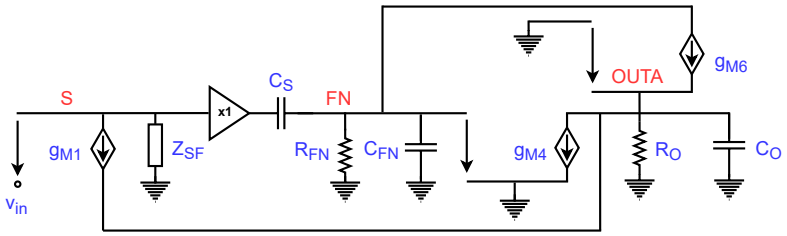


Figure 3.6: Small signal model of the front-end amplifier.

transistor M4. This technique allows to decouple the contributions to the gain from the input transistor and the common-source transistor, simplifying the equations while still providing an accurate analysis. The nodal equations of the small-signal model are:

$$\begin{cases} (V_S - V_{FN})sC_S = V_{FN} / \frac{R_{FN}}{1+sC_{FN}R_{FN}} - g_{M6}V_{OUTA} \\ g_{M4}V_{FN} + V_{OUTA} / \frac{R_O}{1+sC_OR_O} = -g_{M6}V_{OUTA} + g_{M1}(V_S - V_{IN}) \\ g_{M1}(V_S - V_{IN}) + V_S / Z_{SF} = 0 \end{cases}. \tag{3.2}$$

By solving the system, the following gain expression can be found:

$$\frac{V_{OUT}}{V_{IN}} = -\frac{g_{M1}R_O(1 + sR_{FN}(C_{FN} + C_S) + sC_Sg_{M4}Z_{SF}R_{FN})}{(1 + g_{M1}Z_{SF})(g_{M6}g_{M4}R_{FN}R_O + (1 + g_{M6}R_O + sR_OC_O)(1 + sR_{FN}(C_{FN} + C_S)))}. \tag{3.3}$$

Isolating the expression of the output impedance Z_{OUTA} which can be found from the small-signal model by injecting a current into the output node or more easily with the systems theory as $Z_{OUTA,OPEN\_LOOP}/1 + LG$ where LG is the loop gain and is

$$Z_{OUTA} = \frac{R_O(1 + sR_{FN}(C_{FN} + C_S))}{g_{M6}g_{M4}R_{FN}R_O + (1 + g_{M6}R_O + sR_OC_O)(1 + sR_{FN}(C_{FN} + C_S))}, \tag{3.4}$$

the gain expression can be re-written as:

$$\frac{V_{OUTA}}{V_{IN}} = -\frac{g_{M1}}{g_{M1} + 1/Z_{SF}}(\frac{Z_{OUTA}}{Z_{SF}} + g_{M4}Z_{OUTA}\frac{sR_{FN}C_S}{1 + sR_{FN}(C_{FN} + C_S)}). \tag{3.5}$$

In this expression, the term outside the bracket represents the transfer function of the input follower whereas the two terms inside the bracket represent the contributions to the gain given by the input follower and the common-source device M4, as they both inject current

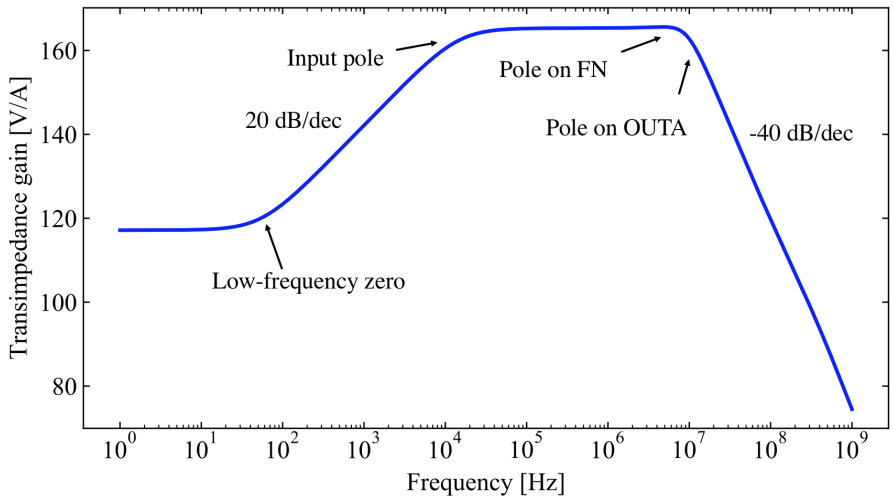


Figure 3.7: AC simulation of the transimpedance gain of the amplifier.

into the output node. The impedance Z_{SF} is stable up to very large frequencies outside the bandwidth of the amplifier due to the low resistance offered by the source of the input transistor and therefore effectively introduces no pole or zero in the frequency range of interest. The gain therefore follows the trend of the output impedance which features a zero at low frequencies, given by the low-frequency loop which suppresses slow signals, and two poles at high frequencies, given by high-impedance nodes, namely the feedback node (FN) and the output one (OUTA). Due to the AC coupling between the input follower and the gate of the transistor M4, the gain contribution of the latter features the zero already in DC. An AC simulation of the transimpedance gain of the amplifier is shown in Fig. 3.7. As the simulation is performed with an input current, an additional pole given by the sensor, modelled with a capacitance, in parallel with the small-signal resistance of the diode biased by the sensor leakage current is present in the transfer function. Overall, the circuit is characterized by a bandpass response. The bandwidth of the amplifier can be optimized for the signal bandwidth to improve the SNR or, in other words, reduce the ENC. No additional shaping is implemented after the amplification stage.

The complete front-end circuit which includes the amplifier and discriminator is shown in Fig. 3.8. The capacitance C_S is implemented with a PMOS device whose source, bulk and drain are connected together to exploit the capacitance of the MOS structure in inversion. The capacitor C_{OUTA} includes only the parasitic contributions of transistors connecting to it, since it needs to be as low as possible. The input transistor M1 is placed together with the capacitor C_S in a separate n-well connected to its source to eliminate the body effect and achieve a gain close to unity for the input source follower. An improvement to the circuit from Fig. 3.5b is provided by cascoding the transistor M4. For good timing performance, a large transconductance is required for this transistor. However, a too large aspect ratio would increase the output parasitic capacitance, detrimental both for gain and speed itself. The cascode decouples the transistor M4 from the output node, giving more freedom in its sizing, and is optimized for a reduced output capacitance. Additionally, it increases the output impedance of the amplifier, which is thus dominated by the transconductance of the transistor M6 which works in weak inversion, leading to a higher gain. The discriminator consists of a common-source

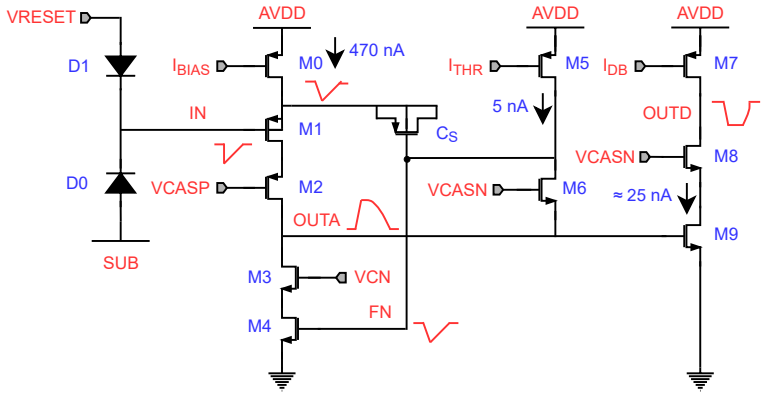


Figure 3.8: Complete schematic of the front-end with discriminator.

amplification stage, the transistors M7-M9, which can be better seen as a current comparator. In steady state, the output baseline of the amplifier sets the standby current of the transistor M9, while the transistor M7 is biased to provide a current I_{DB} higher than the DC current forced by the transistor M9, charging the node OUTD to the supply voltage. As the signal on OUTA rises upon a particle hit, the current drawn by the transistor M9 increases, eventually exceeding I_{DB} and discharging the output node to ground. The threshold of the discriminator is therefore controlled by the I_{DB} current setting and the amplifier output baseline (through V_{CASN}). The cascode transistor M8 is again used to reduce the large capacitance penalty on OUTA due to the Miller effect on the transistor M9 and the coupling between this node and the rail-to-rail OUTD signal. In the actual front-end implementation, three parallel NMOS switches are placed between the source of the transistor M9 and the ground. These switches are controlled by three different digital signals which are connected to all the pixels in a row, column or diagonal. If all the switches are open, the discriminator is disabled and cannot generate an output pulse. This logic gives the possibility to address a pixel and mask it in case it generates an excessive noise hit rate.

The circuit is designed to have peaking times in the order of tens of ns with a low power consumption. For the input follower action, the input transistor more quickly discharges its load capacitance with a larger transconductance g_m. The speed of the gain contribution provided by the transistor M4 is defined by its transconductance g_m and the output capacitance C_{OUTA}. Therefore, the peaking time decreases with a higher transconductance g_m of the amplifying devices and a lower output capacitance C_{OUTA}. The transistors' dimensions and the layout are therefore optimized to reduce C_{OUTA} to less than $\approx 5\,\mathrm{fF}$. To reach the target timing response, the main biasing current I_{BIAS} needs to be $\approx 470\,\mathrm{nA}$ for a sufficiently large g_m of the amplifying devices. The I_{THR} current, typically a few nA, and the discriminator off current, typically a few tens of nA, need to be added to the I_{BIAS} current to obtain the total current consumption which is $\approx 500\,\mathrm{nA}$. With a supply voltage of 1.8 V, the total power consumption is less than $1\,\mu\mathrm{W}$ per pixel, which is used efficiently thanks to the current reuse between the input follower M1 and the common-source device M4. A parasitic-extracted simulation of the transient waveforms at the input IN, analog output OUTA and discriminator output OUTD of the front-end with the charge threshold set to $100\,\mathrm{e}^-$ is shown in Fig. 3.9. The solid lines show the response for a collected charge of $250\,\mathrm{e}^-$, the dashed lines for a charge of $1000\,\mathrm{e}^-$. The simulation was performed using a current pulse at the input, i.e. by injecting the input charge uniformly in a collection time of 1 ns. The sensor is modelled as a capacitance of 2.5 fF, which is a value previously measured on prototype chips [89], in parallel with a leakage current source of 10 pA. The red curves represent the input signals and show that the voltage step on the electrode is proportional to the collected charge. The blue curves represent the amplified signals on OUTA. The front-end gain is non-linear since the transistor M6 dynamically turns off as the output voltage rises, offering a larger impedance on the output node. At threshold, the gain is $\approx 1.9\,\mathrm{mV/e}^-$, whereas for a charge of $250\,\mathrm{e}^-$, as seen in Fig. 3.9, it is $\approx 2.5\,\mathrm{mV/e}^-$. For larger charges, the analog output signal on OUTA is sufficiently large to push the cascode transistor M2 out of saturation and the front-end gain drops. This makes

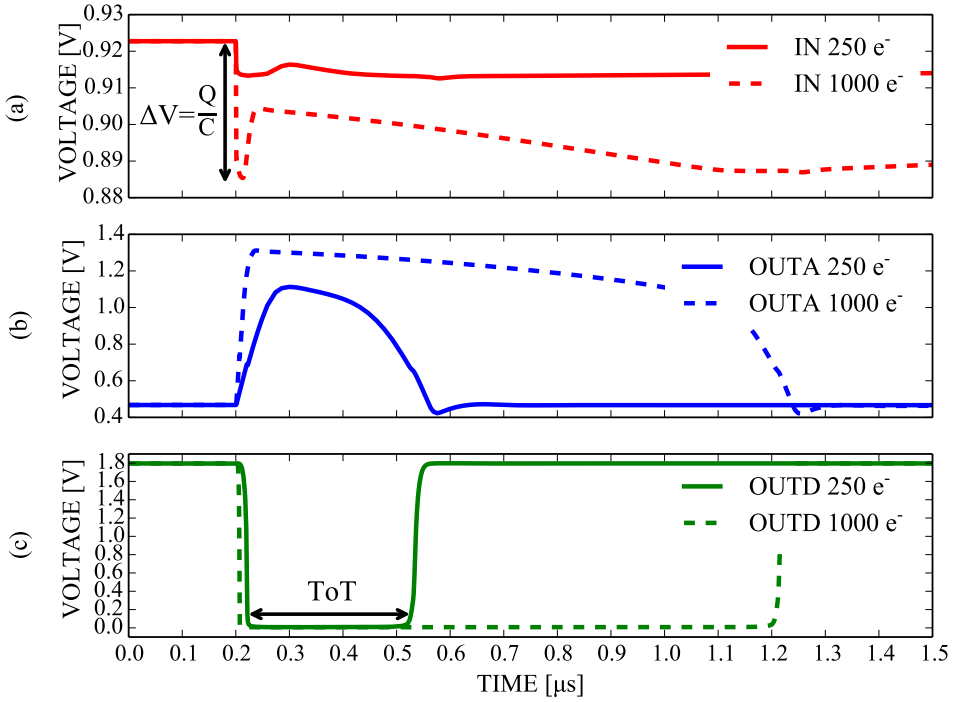


Figure 3.9: Front-end simulated transient response with a 100 e^- threshold: signals at the (a) sensing node (b) output of the amplifier (c) output of the discriminator.

the cascode ineffective, so the equivalent input capacitance increases due to the Miller effect. The input signal is therefore lower during this transition time, as shown from the dashed red curve of Fig. 3.9, and a saturation of the analog output signal is reached. However, its Time over Threshold (ToT), i.e. the duration of the discriminator output pulse shown in green in Fig. 3.9, has a linear dependence on the input charge. Indeed, the ToT depends on the time required for the feedback circuit to charge up the capacitor C_S through the current I_{THR}.

For a reliable operation of the sensor at low thresholds, the front-end ENC has to be minimized. Noise simulations show that the main noise contributors are the amplifying devices M1 and M4 due to their large transfer function to the output node. A small percentage of noise comes also from the I_{THR} current source and the device M6 which define the baseline of the amplifier output. An important noise source not accounted for in the simulation is the Random Telegraph Signal (RTS) noise. Indeed, the simulation models do not include RTS noise which is therefore difficult to estimate during design. The most critical devices for RTS noise are again the transistors M1 and M4 whose sizing required an iterative process [83]. Increasing the gate area of the transistor M1 to combat RTS noise results in a larger effective sensor capacitance. A gate area of 0.18 μm^2 has been chosen since it is a good compromise between capacitance penalty and noise. A gate area of 2.4 μm^2 has instead been chosen for the transistor M4 since it exhibits a larger noise transfer function to the output node and the RTS noise is typically larger in NMOS transistors. The enlargement of this transistor and the introduction of its cascode to prevent it from excessively loading the output node are

among the main modifications of the front-end from the previous prototypes. One of the main parameters for good noise performance is the size of the capacitance C_S: a larger capacitance provides dynamically more charge to the output node for the same input signal and improves the coupling between the source of the input transistor and the gate of the transistor M4. From a frequency standpoint, it widens the amplifier passband towards lower frequencies where the input signal has a large frequency content. The output signal therefore increases more than the noise level and a larger SNR is obtained. Also the area of the C_S capacitance has been increased with respect to previous front-end implementations. The PMOS transistor implementing this capacitor has a gate area of $\approx 14.24\,\mu m^2$, providing a capacitance of $\approx 114\,fF$, and it is one of the largest components of the circuit. The front-end ENC can be evaluated with the s-curve, i.e. the front-end probability to generate a hit as a function of the input charge due to noise. Fig. 3.10 shows this curve obtained with transient noise simulations. It follows the behaviour of a Gaussian error function, therefore, by fitting it to such function, the average of the fit provides an estimate of the nominal threshold whereas its standard deviation an estimate of the front-end ENC. In this case, the charge threshold and ENC are respectively $\approx 100\,e^-$ and $\approx 6.4\,e^-$.

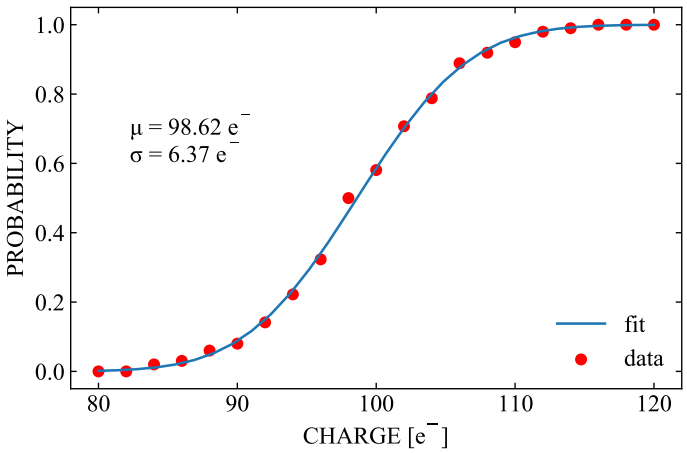


Figure 3.10: Hit probability as a function of the input charge with simulated transient noise.

Apart from the noise, another limit to the minimum operating threshold is the pixel-to-pixel variation of the transistor parameters which causes the threshold to vary over the matrix. It is well known that the transistors' mismatch scales down with the square root of their area [90]. The pixel size is however limited and often dictated by the target sensor spatial resolution. To optimize the space, it is therefore necessary to identify the devices with the largest impact on the threshold dispersion and increase their area. In the amplification stage, the most critical devices are the transistors M5 and M6. As previously said, the current I_{THR} of the transistor M5 defines the speed of the feedback loop and significantly influences the amplifier gain. Regarding the transistor M6, its gate-source voltage directly defines the amplifier output baseline, setting the discriminator DC current and hence its switching threshold. The transistor M5 is biased with a current of only a few nA and therefore operates in weak inversion, which makes the impact of its mismatch even more prominent. For these reasons, it is designed with

a low aspect ratio and a large area ($20\,\mu m^2$). The size of the transistor M6, however, cannot be increased to the same extent because of the capacitance penalty on the output node. In the discriminator stage, the input transistor M9 is the main critical device: a variation of its threshold voltage directly shifts the switching point of the discriminator, appearing effectively as an offset. As for the transistor M6, it has to be kept small to prevent increasing the amplifier output capacitance and it represents the largest contribution (nearly 50 %) to the overall threshold dispersion. Fig. 3.11 shows the s-curve obtained with Monte Carlo simulations for transistors' mismatch. The fit to a Gaussian error function indicates an average threshold of $\approx 100\,e^-$ with a pixel-to-pixel threshold variation of $\approx 2.5\,e^-$.

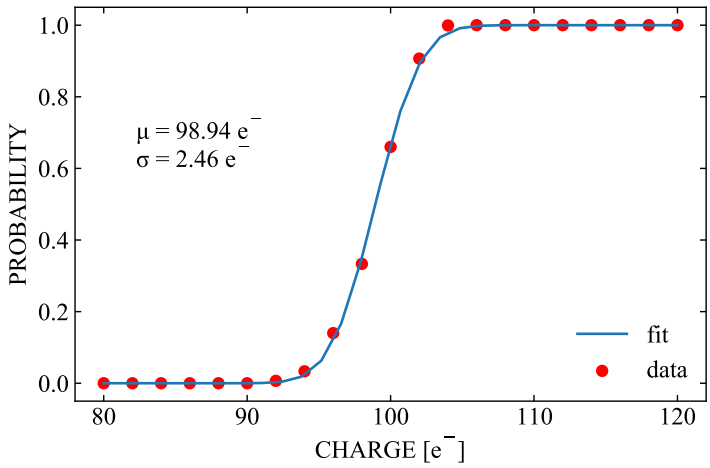


Figure 3.11: Hit probability as a function of the input charge with simulated transistors' mismatch.

The radiation effects that influence the front-end operation and performance include the increase of the sensor leakage current due to NIEL and the TID effects that affect the transistor characteristics. For a good tolerance to NIEL, the front-end has been designed to cope with a wide range of sensor leakage currents. As for the tolerance to TID, minimum dimensions have been avoided for the critical devices to mitigate RINC- and RISC-effects [75]. Leakage currents in the order of a hundred pA have been measured for NMOS transistors in this technology after 20 Mrad of TID [91]. Since the I_{THR} current can be below 1 nA, the transistor M6 has been designed as an ELT [77] and is surrounded by a p+ guard ring to prevent any leakage to neighboring devices. These precautions double its area but they are necessary to ensure radiation hardness. The layout of the pixel is shown in Fig. 3.12. The 2 µm octagonal collection electrode, distanced 4 µm from the surrounding p-well of the circuitry, is placed in the centre of the pixel. The front-end circuit occupies an area of $\approx 160\,\mu m^2$ and is placed to the left of the collection electrode with other analog circuitry. The latter includes decoupling capacitors and a testing circuit which can capacitively inject a tuneable amount of charge into the collection electrode. The rest of the pixel is occupied by the digital readout circuitry for a total pixel area of $36.4\,\mu m \times 36.4\,\mu m$, leading to an analog power density of $\approx 75\,mW\,cm^{-2}$ over the matrix, well within the requirements for the ATLAS ITk.

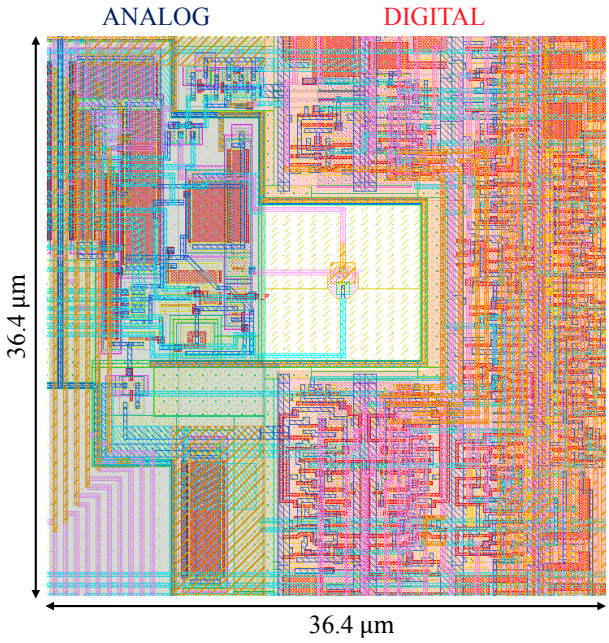


Figure 3.12: Layout of the MALTA2 pixel.

3.3 The MALTA2 chip

The MALTA2 chip has a size of 10.12 mm × 20.2 mm and integrates a matrix of 224 × 512 pixels. The power pads are distributed along the left and right sides of the chip. The biases of the in-pixel circuits are adjusted through DACs placed in the periphery at the bottom of the matrix and distributed along its entire width. A serial logic, referred to as "slow control" as it operates with a 40 MHz clock, allows to communicate with the chip and write in the registers used to adjust the different settings of the chip. The matrix features an asynchronous readout logic, inherited from the previous prototype. Upon a particle hit, the in-pixel digital circuitry instantaneously sends a pattern of short pulses corresponding to the pixel address to the periphery at the bottom of the matrix on a digital data bus. For the readout, the pixels are organized in double columns and each double column has a dedicated bus for the transmission of the data down to the periphery. In this approach, the digital power consumption in the matrix to transmit the hits is dependent on the hit rate. For the target hit rate of 100 MHz cm^{-2} of the ATLAS ITk outer layer, this asynchronous logic provides a significant power reduction compared to a synchronous readout which distributes a clock at 40 MHz, the bunch crossing frequency, over the matrix [92]. In the periphery, a binary tree-like structure that merges the hits of the whole matrix onto a single bus is implemented. In the case of simultaneous events, this logic delays one of the two hits in time while keeping track of the delay for later reconstruction. The final word containing the pixel address and introduced time shift is 40 bits wide and is transmitted off-chip via LVDS drivers, which are designed to operate at a maximum speed of 5 Gbps [93], sufficiently high for the pixel sensor to cope with the ATLAS ITk outer layer hit rate. A picture of the MALTA2 chip is shown in Fig. 3.13.

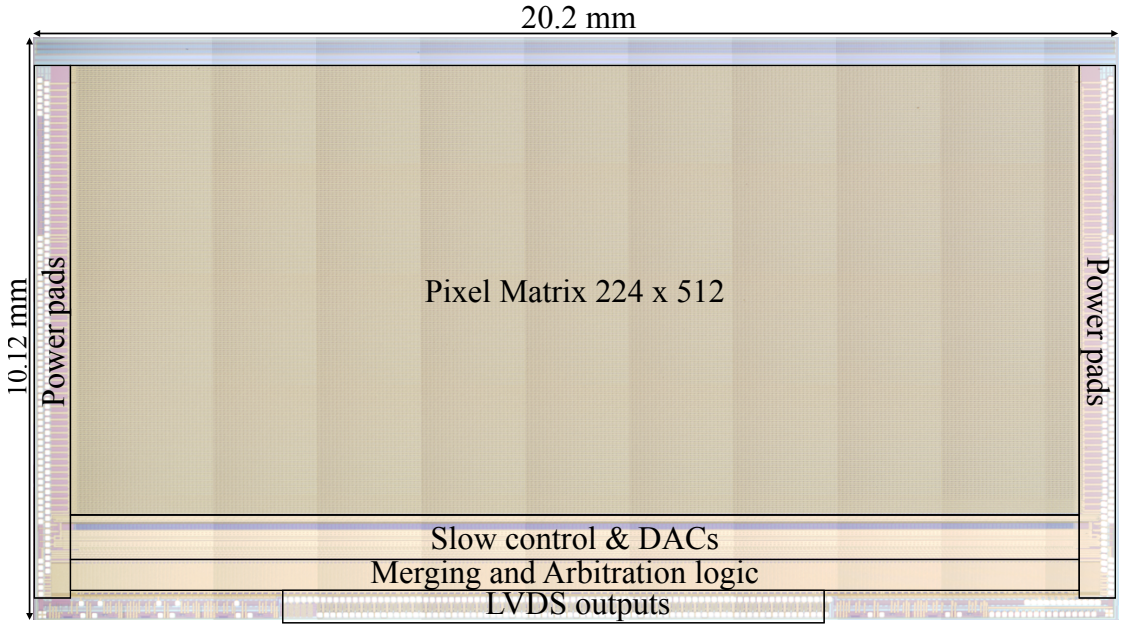


Figure 3.13: Picture of the MALTA2 chip.

3.4 Front-end characterization

The readout system used to characterize the MALTA2 sensor is based on a custom PCB where the chips are wirebonded and the LVDS output connected to an FPGA used as a data acquisition system. To test the performance of the front-end, a special set of pixels which allows the monitoring of their analog output has been included on the left and right side of the matrix. In these pixels, the front-end analog output is buffered to an output pad with a two-stage source follower with a gain close to 1. The first stage is optimized to match closely the discriminator input capacitance to have the same amplifier output load as in the other pixels of the matrix. An oscilloscope is used to monitor the output pad through a low-capacitance active probe and the full buffering system is designed not to degrade the signal timing. The front-end speed can be evaluated with the plot of Fig. 3.14 which shows the time walk curve, i.e. the time for the amplifier output to reach the discriminator threshold as a function of the input charge. The conversion between charge and amplitude is derived through the charge injection circuitry. The injection capacitance was calibrated with ToT measurements of signals from test pulses and an ^{55}Fe source. For this measurement, the front-end operates with the nominal bias settings ($\approx 1\,\mu W$ power consumption, as confirmed by measuring the total analog current consumption of the matrix) and the oscilloscope is set to trigger with a signal of $\approx 100\,e^-$. The waveforms are collected while exposing the chip to a ^{90}Sr radioactive source which undergoes β^- decay emitting electrons that generate an ionization signal close to a MIP. The most probable value of charge deposition for a MIP in the 30 µm thick epitaxial layer is $\approx 1800\,e^-$ [51]. The signal is collected by a cluster up to 4 pixels and the seed pixel, the one with the largest signal, has a charge $\geq 1/4$ of a MIP charge. Events with high charges ($\gtrsim 1200\,e^-$) have a threshold crossing time close to the minimum value of ≈ 10 ns. With respect to the ATLAS

application, an event is considered in time when it falls within 25 ns from this value. As can be noticed in Fig. 3.14, the in-time threshold corresponds to an input charge of $\approx 200\,e^-$. Less than 10 % of the hits are below the in-time threshold. Statistically, these are mostly caused by non-seed pixels, with a neighboring seed pixel which is likely to collect a charge above the in-time threshold.

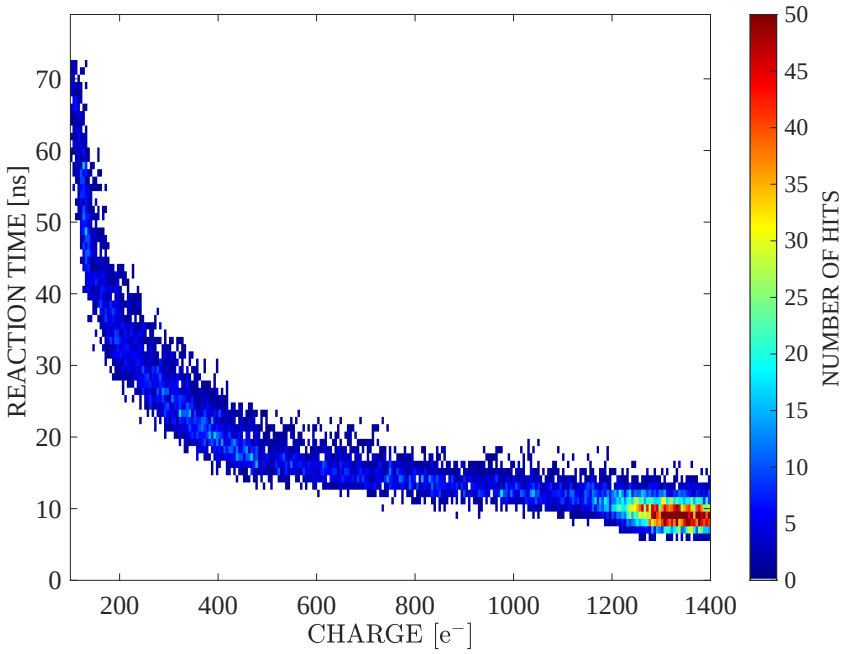


Figure 3.14: Time walk curve obtained with a ^{90}Sr source.

The front-end timing can also be studied through the matrix digital readout. An increasing amount of charge can be injected in a specific pixel with the aforementioned charge injection circuitry. The time of arrival of the generated hits can then be compared to a time reference. This procedure has been performed using as time reference the charge injection trigger pulse sent to the chip. In order to do so, this signal is also sent to an external 3 ps binning TDC [94] together with a fast-OR signal from the chip. The mean difference between these two signals' time of arrival provides a time walk curve compatible with the one in Fig. 3.14. This methodology, however, allows to better study the front-end jitter by evaluating instead the RMS difference of the two signals' time of arrival, which is plotted as a function of the injected charge in Fig. 3.15. For each charge, ten thousand events are acquired. The time jitter of the reference pulse has been estimated to be below 100 ps, therefore, the values in Fig. 3.15 are dominated by the front-end jitter which reduces from 4.7 ns at threshold, down to 0.16 ns for very high input charges ($\gtrsim 1200\,e^-$).

The charge injection circuitry also allows to extract information such as threshold and noise: varying the charge injected into a pixel, an s-curve as the one in Fig. 3.10 can be obtained and the front-end threshold and noise extracted through the Gaussian error function fit as done before. Fig. 3.16 shows the threshold and noise distributions for an entire matrix with

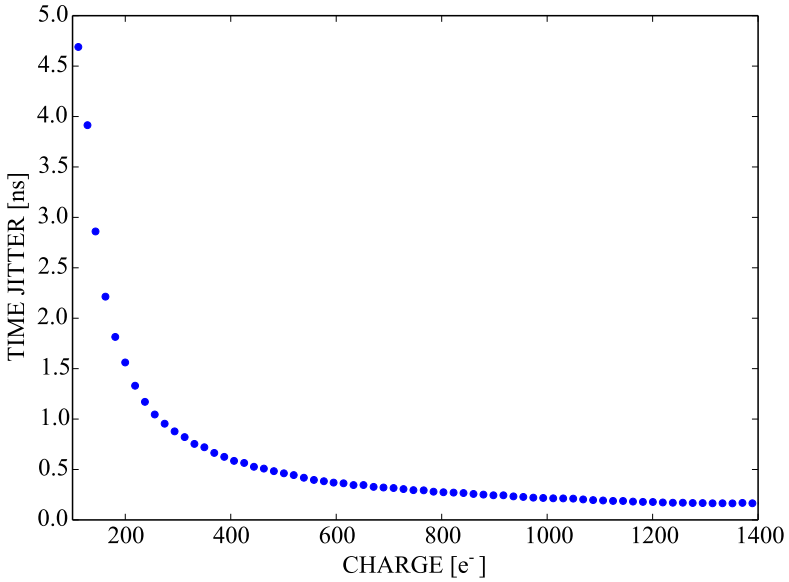


Figure 3.15: Dependence of front-end time jitter on the input charge.

nominal front-end settings. The average threshold is $\approx 100\,e^-$ with a variation of $\approx 6\,e^-$, more than a factor of 2 higher than the simulated value shown in Fig. 3.11. The noise distribution has an average of $\approx 6.5\,e^-$ with a low spread, matching fairly well the simulations. 2D maps of the pixels' threshold and noise are shown in Fig. 3.17. No systematic effects are observed for the noise. As for the threshold, it is possible to notice a variation of its average over different vertical sections of the matrix. This effect strongly correlates with the scheme of the front-end biasing which is adjusted through DACs in the bottom periphery. As mentioned, the power pads are distributed only along the left and right side of the matrix. For this reason, a horizontal power voltage drop is inevitably present and is estimated to reach $\approx 12\,mV$ in the middle of the matrix. To compensate this effect and avoid a systematic threshold gradient, the biasing DACs have a dedicated mirroring stage for every 32 columns of the matrix which shares their local power supply. To better study the threshold behaviour, the distribution of the threshold along the columns with one RMS error bar is shown in Fig. 3.18a and here the threshold average variation at every biasing group is more clearly visible. A straightforward solution to increase the biasing transistors' area and mitigate this effect is to connect more mirroring stages together, trading-off with the power voltage drop compensation accuracy. This is envisaged for a future prototype. Fig. 3.18b illustrates the distribution of the threshold along the rows showing a slight vertical gradient which is caused by the mirroring stages at the matrix bottom which load the matrix power grid and introduce a vertical power voltage drop. Considering only pixels within the same biasing group and correcting the systematic vertical gradient, the threshold variation is $\approx 5.1\,e^-$. The variation of the NMOS transistors' output conductance with a high reverse bias to the bulk (beyond the normal supply voltage) is not fully covered by the simulation models and this is thought to be the cause of the discrepancy between the simulated and measured threshold variation. Even with a larger pixel-to-pixel mismatch, the chip can be operated reliably with thresholds of $\approx 100\,e^-$.

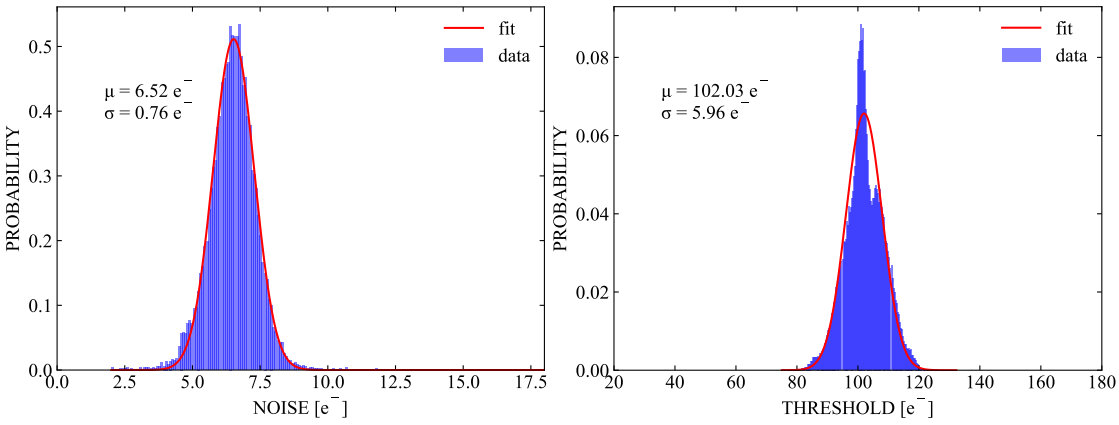


Figure 3.16: Distribution of (a) ENC (b) threshold.

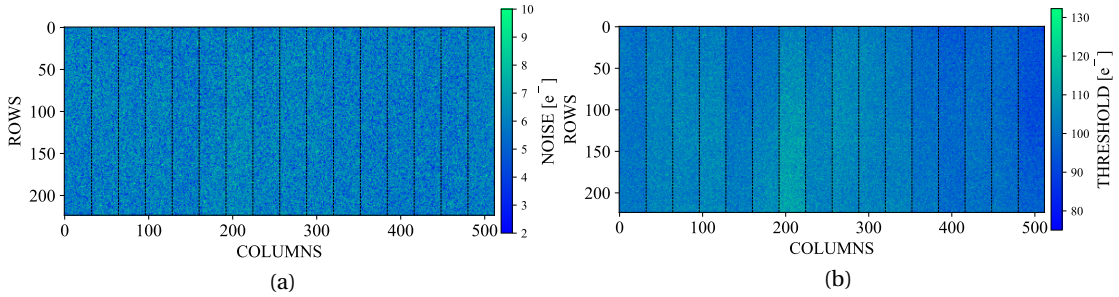


Figure 3.17: 2D map for (a) ENC (b) threshold.

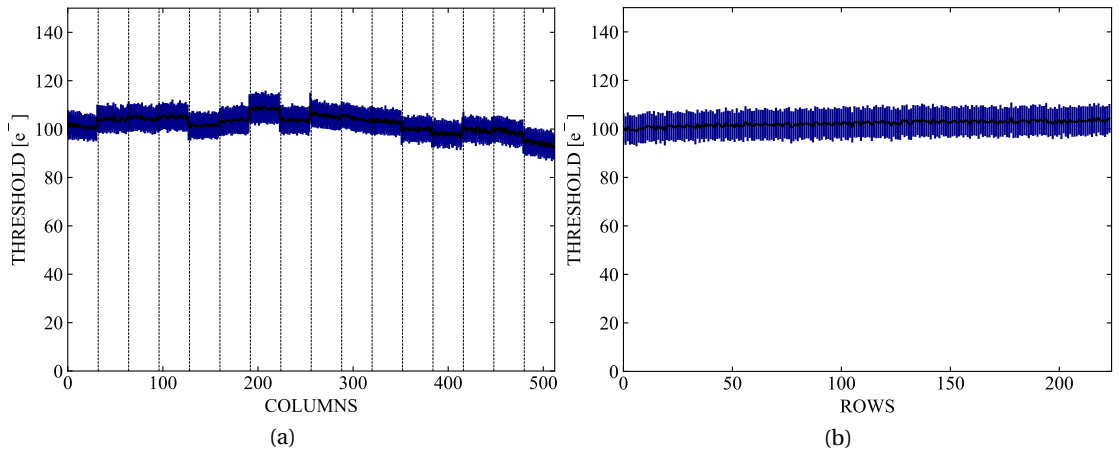
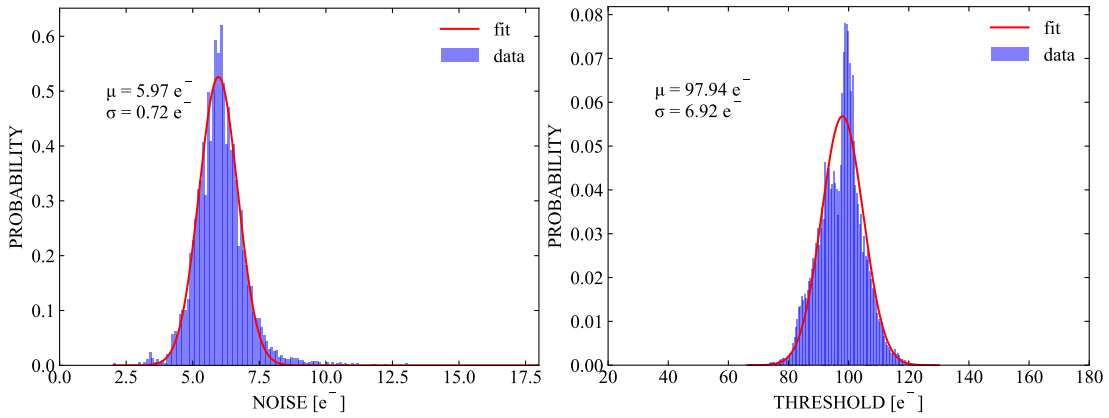
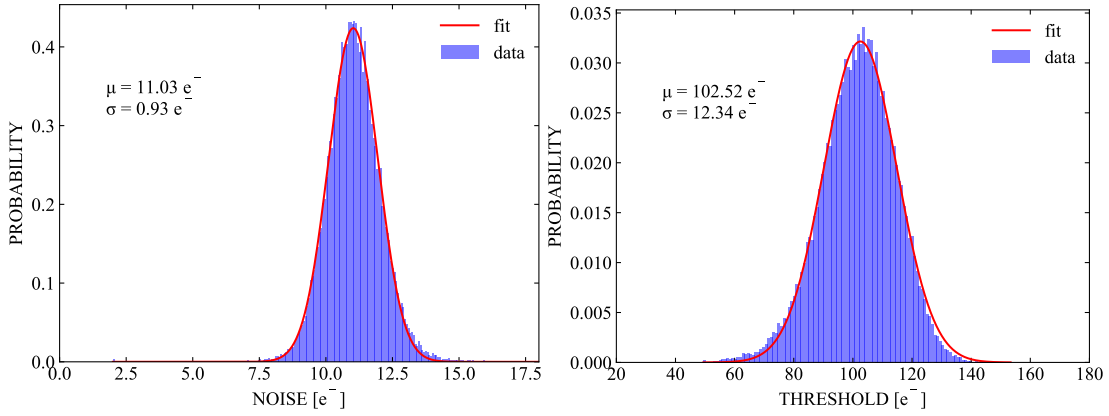


Figure 3.18: Distribution of the threshold vs the (a) columns and (b) rows.

A number of chips have been irradiated with neutrons at the TRIGA reactor in Ljubljana [95] up to $3 \cdot 10^{15}\,1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$ of NIEL fluence. The chips also received a background TID of 1 Mrad for every $10^{15}\,1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$. During irradiation, the chips were not powered. After irradiation, the chips are stored at low temperature, below −20 °C, to avoid annealing of



(a)



(b)

Figure 3.19: Distribution of ENC and threshold with a threshold of $\approx 100\,e^-$ for (a) an unirradiated sample and (b) a sample irradiated at $3 \cdot 10^{15}\,1\,MeV\,n_{eq}\,cm^{-2}$ and 3 Mrad.

the radiation damage. For the same reason, all the measurements of irradiated samples are performed at $-20\,^\circ C$ which also helps to contain the increase of the sensor leakage current. The chip still shows full functionality after these levels of irradiation. Charge injection tests have been performed on these samples with a step of $10^{15}\,1\,MeV\,n_{eq}\,cm^{-2}$ and 1 Mrad. For a fair comparison, the measurement of the unirradiated sample has been repeated at $-20\,^\circ C$ and the I_{THR} current setting of the front-end has been adjusted to obtain similar thresholds in all the cases. An increasing level of ENC and threshold dispersion as a function of the irradiation level has been noticed. The distributions of ENC and threshold dispersion for an unirradiated sample and a sample irradiated to $3 \cdot 10^{15}\,1\,MeV\,n_{eq}\,cm^{-2}$ and 3 Mrad are shown in Fig. 3.19 for reference. The noise average for the unirradiated sample is $\approx 6\,e^-$, slightly lower than in the previous case due to the lower temperature, and increases to $\approx 11\,e^-$ for the sample irradiated at $3 \cdot 10^{15}\,1\,MeV\,n_{eq}\,cm^{-2}$ and 3 Mrad. The lack of substantial non-Gaussian tails suggests a negligible contribution of RTS noise. The pixel-to-pixel threshold variation increases from $\approx 7\,e^-$ in the unirradiated case to $\approx 12.5\,e^-$ for the sample irradiated at $3 \cdot 10^{15}\,1\,MeV\,n_{eq}\,cm^{-2}$ and 3 Mrad.

To evaluate the front-end performance for higher TID levels, the chip has been irradiated with X-rays at a dose rate of 25 krad/min up to 100 Mrad. The chip still shows full functionality at this TID level. The irradiation was stopped at different doses to perform basic functionality tests and evaluate the front-end performance. In order to reproduce the typical operating conditions, the chip was powered and biased during irradiation. Additionally, to minimize annealing effects, the chip was kept at a low temperature (−10 °C) during the whole process. The measured ENC and pixel-to-pixel threshold variation as a function of TID are shown in Fig. 3.20. The threshold was adjusted to ≈ 100 e^- at each step of the measurement. The ENC grows monotonically from ≈ 5.9 e^- before irradiation (first data point) to ≈ 22.5 e^- at 100 Mrad. High levels of RTS noise are present in the ENC distributions for TID levels higher than 1 Mrad. However, already after 24 hours of annealing at room temperature, the RTS noise disappears and the mean ENC drops from ≈ 22.5 e^- to ≈ 19 e^- (data point at 250 Mrad). After other 24 hours of annealing at 80 °C, the mean ENC reduces to ≈ 14 e^- (last data point). As for the pixel-to-pixel threshold variation, it increases from ≈ 6.8 e^- before irradiation to ≈ 23 e^- at 100 Mrad. The threshold dispersion more rapidly increases with TID compared to the noise, but it settles around ≈ 23 e^- already at 1 Mrad. After 24 hours of annealing at room temperature, it drops to ≈ 14.5 e^- and it further drops to ≈ 9 e^- after other 24 hours of annealing at 80 °C.

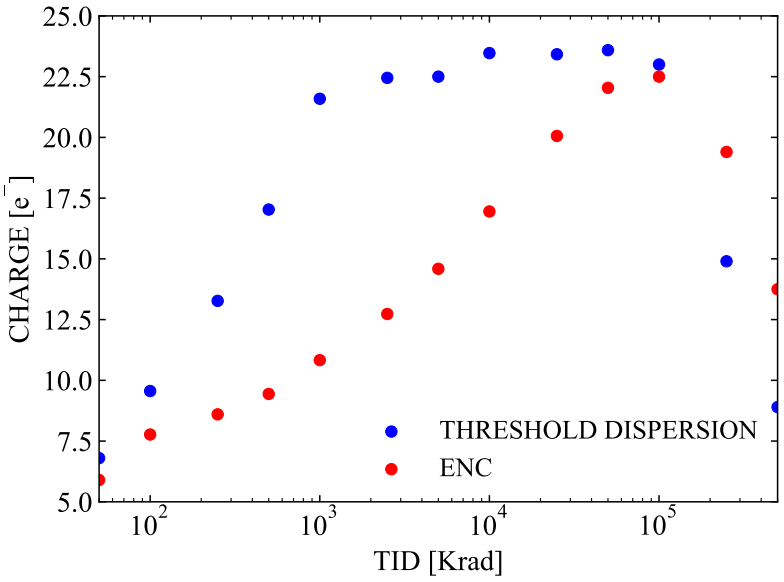
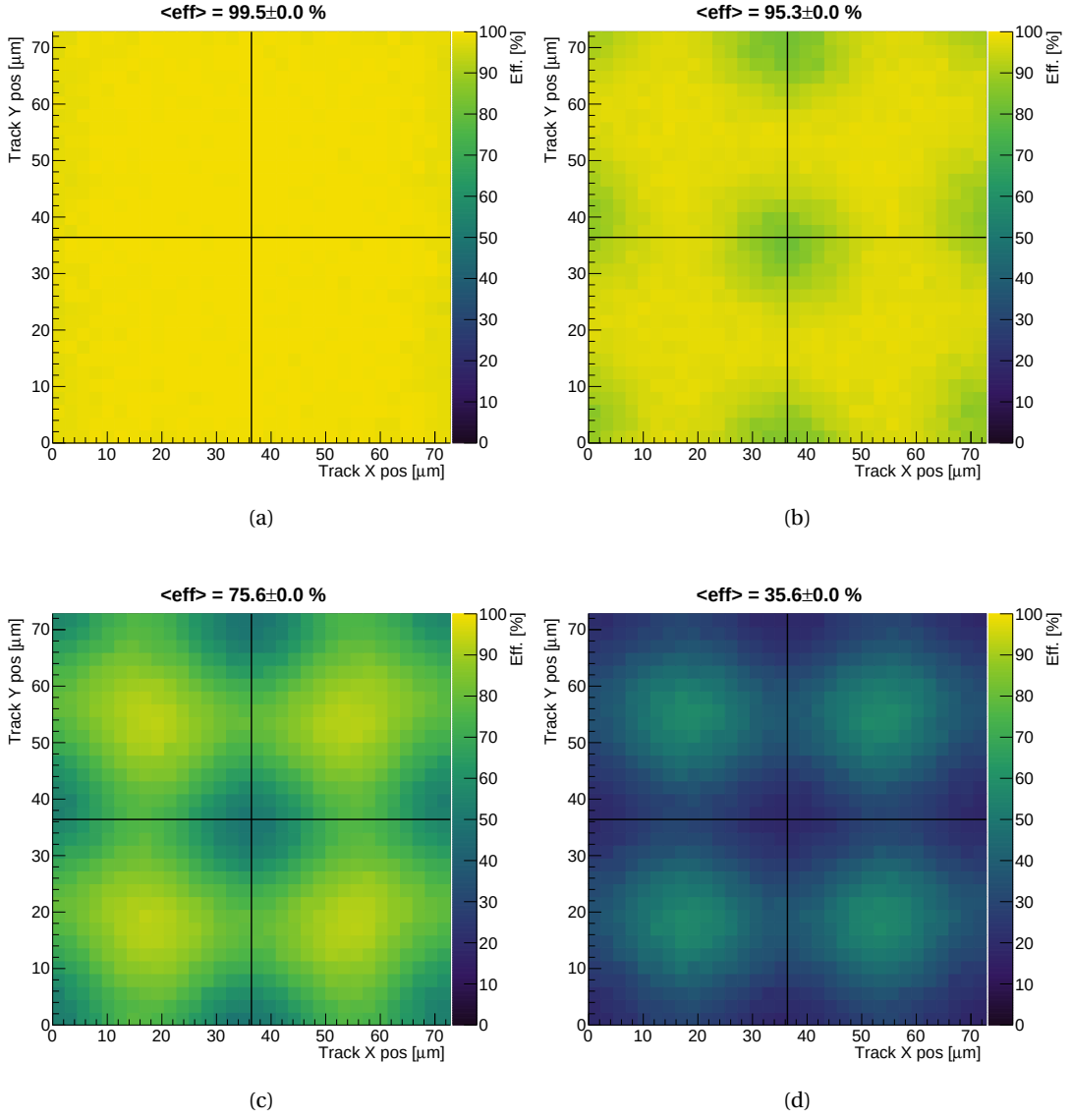


Figure 3.20: Dependence of ENC and threshold dispersion on TID with a threshold of ≈ 100 e^-. The points at 50 krad represent measurements before irradiation. The points at 250 Mrad and 500 Mrad correspond to measurements after 24 hours annealing at room temperature and additional 24 hours of annealing at 80 °C, respectively.

Measurements with test beams have been performed to evaluate the detection efficiency of the MALTA2 sensor. For these tests, MALTA2 samples were exposed to the 120 GeV hadron beam of the CERN SPS and placed between six reference planes, three upstream and other three downstream, which form the so-called beam telescope. The tracks of the particles in the

beam reconstructed with the telescope have a spatial accuracy of ≈ 5 µm on the DUT. Thanks to this, the detection efficiency can be determined with sub-pixel precision. The efficiency over an area of 2 × 2 pixels with a charge threshold of ≈ 100 e^- is shown in Fig. 3.21 for an unirradiated sample and samples irradiated up to $3 \cdot 10^{15}$ 1 MeV n_{eq} cm^{-2} and 3 Mrad with a step of 10^{15} 1 MeV n_{eq} cm^{-2} and 1 Mrad.



(a) (b)

(c) (d)

Figure 3.21: In-pixel efficiency over an area of 2 × 2 pixels with a charge threshold of ≈ 100 e^- for (a) an unirradiated sample and samples irradiated up to $3 \cdot 10^{15}$ 1 MeV n_{eq} cm^{-2} and 3 Mrad with a step of 10^{15} 1 MeV n_{eq} cm^{-2} and 1 Mrad in (b), (c) and (d) respectively.

During the measurement, the irradiated samples are placed in a cooling box at −20 °C to avoid annealing of the radiation damage and reduce the sensor leakage current. For the unirradiated sample, the efficiency has an average of ≈ 99.5 % and is uniform over the pixels. For the device

irradiated to 10^{15} 1 MeV $n_{eq}\,cm^{-2}$ and 1 Mrad, the efficiency has an average of ≈ 95.3 % and is highest in the center, where it still reaches values close to ≈ 100 %, and drops to ≈ 70 % in the pixel corners. In these regions, indeed, the charge is shared among more pixels and, due to the longer path to the collection electrode, is also more likely to get trapped by the radiation-induced defects. For events occurring in the pixel corners, therefore, the charge induced on the electrode might not be large enough to trigger the discriminator, hence the lower efficiency. This effect is even more prominent in the sample irradiated to $2 \cdot 10^{15}$ 1 MeV $n_{eq}\,cm^{-2}$ and 2 Mrad as in the pixel corners it features an efficiency of ≈ 50 %. The sample irradiated to the highest level of irradiation has instead an average efficiency of ≈ 35.6 %. In this case, indeed, the density of radiation-induced traps is so high that the detection efficiency is low even in the pixel centers.

3.5 Conclusions

Monolithic sensors were considered as cost-effective solutions for the outer pixel layer of the ATLAS ITk. The MALTA monolithic sensors target this application, very demanding in terms of radiation-hardness. This chapter described the optimization of the front-end integrated in the MALTA2 sensor. The circuit is implemented in the TowerJazz 180 nm CMOS imaging technology. The sensor features a small octagonal collection electrode with a diameter of 2 µm to obtain a low sensor capacitance (<5 fF), which is key to achieve high analog performance. Process modifications have been introduced to fully deplete the sensor and enhance the lateral electric field along the pixel edges for good tolerance to NIEL. The front-end is a continuously active open-loop amplifier followed by a high-gain common-source discriminator stage. It is designed for peaking times in the order of tens of ns, requiring a power <1 µW and an area of ≈ 160 μm^2. Furthermore, it has a gain of ≈ 2 mV/e^-. MALTA2 samples were extensively characterized to evaluate the front-end performance before and after irradiation. The main front-end metrics with a threshold of ≈ 100 e^- are summarized in Tab. 3.1. For the same threshold, the efficiency of the sensor is ≈ 99.5 % and drops to ≈ 35.6 % after $3 \cdot 10^{15}$ 1 MeV $n_{eq}\,cm^{-2}$ and 3 Mrad mainly due to a loss of collected charge when the particle hits occur along the pixel edges.

Table 3.1: Front-end specifications with a threshold of $100\,e^-$.

Parameter		**Value**
Area		$160\,\mu m^2$
Power consumption		$1\,\mu W$
In-time threshold (for a 25 ns time window)		$200\,e^-$
Time jitter	at threshold	4.7 ns
	for high charges ($\gtrsim 1200\,e^-$)	0.16 ns
ENC	unirradiated	$6.5\,e^-$
	$3\cdot10^{15}$ 1 MeV $n_{eq}\,cm^{-2}$, 3 Mrad	$11\,e^-$
	100 Mrad	$22.5\,e^-$
Threshold dispersion	unirradiated	$6\,e^-$
	$3\cdot10^{15}$ 1 MeV $n_{eq}\,cm^{-2}$, 3 Mrad	$12.5\,e^-$
	100 Mrad	$23\,e^-$

4 Developments in the TPSCo 65 nm imaging technology

Small linewidth technologies are highly desirable in the context of pixel sensors for HEP experiments. Indeed, more advanced nodes offer greater integration densities and lower power consumption for the same performance, beneficial to meet the requirements of high-resolution and low-mass vertex detectors for future experiments. Significant expertise has been gained in the TowerJazz 180 nm imaging technology with the ALPIDE and MALTA developments. The possibility of moving future monolithic developments in a more advanced technology has been explored in the framework of the EP-R&D program at CERN [96]. The 65 nm imaging technology from the same foundry, more precisely the Tower Partners Semiconductor Company (TPSCo), has been considered as a possible candidate for these developments. One of the main applications targeted by the R&D effort is the ALICE ITS upgrade. Several test structures have been submitted to verify the suitability of this technology for HEP applications. In particular, transistors test structures (TTS) have been developed to evaluate the tolerance of the transistors to TID (some measurements on this structure were actually provided as an example of TID effects on transistors in section 2.4.4). Analog pixel test structures (APTS) have been developed instead to characterize the analog behaviour of the sensor and its tolerance to NIEL. A digital pixel test structure (DPTS), so called as its readout chain features a front-end amplifier and discriminator followed by digital readout logic, has also been developed. This chip allows to validate not only the performance of the sensor but also novel readout circuit solutions.

This chapter focuses on the design and characterization of the APTS and DPTS chips. In particular, section 4.1 illustrates the sensor in the TPSCo 65 nm imaging process, highlighting the differences with the previous technology. Section 4.2 explains the APTS and its readout circuitry which allows to monitor the analog behaviour of the sensor. Section 4.3 explains instead the DPTS with particular focus on its analog front-end circuit. For both structures, the main characterization results are shown in their corresponding section.

4.1 Sensor technology

The cross-section of the standard sensor developed in the TPSCo 65 nm imaging process is very similar to that in the 180 nm process, as the main differences are only in the doping levels and dimensions involved [97], [98]. It is shown again in Fig. 4.1a for reference. In this process, the epitaxial layer is 10 µm thick. Despite the smaller feature size of the transistors, achieving pixel pitches smaller than 15 µm is challenging. Due to the larger ratio between the pitch and thickness of the pixel, the process modifications to fully deplete the epitaxial layer and enhance the lateral electric field to speed up the charge collection are even more needed in this case. These entail the introduction of a low-dose n- implant implant, as shown in Fig. 4.1b, and the cutting of this implant along the pixel edges, as in Fig. 4.1c. The modification with the extra deep p-well has not been pursued as it provides similar performance to the one with the cut in the n- layer but requires an extra implant. In the pixels implemented in this work, the collection electrode is an octagonal-shaped n-well with a diameter of 1.14 µm and is distanced 1.93 µm from the surrounding p-well containing the circuitry. The cut in the n- implant is 2.5 µm wide and centered along the pixel edges. This geometry is the result of a trade-off between a small sensor capacitance (<5 fF) and a large lateral electric field, with collection times in the sub-nanosecond range, as shown later in this chapter.

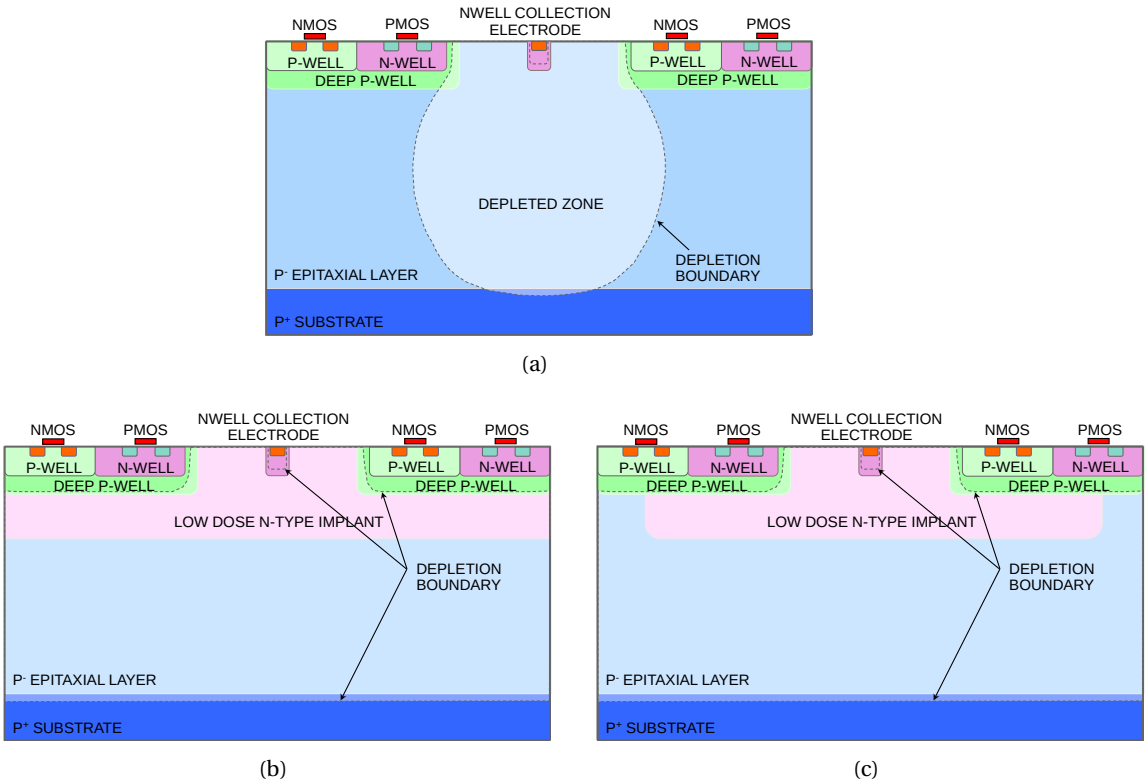


Figure 4.1: Cross section of the sensor in the TPSCo 65 nm imaging technology (a) standard process (b) modified process with low-dose n- implant (c) with gap in the low-dose n- implant.

4.2 Analog Pixel Test Structure

The goal of the APTS is to study the analog properties of the sensor. In order to do so, it integrates a small matrix of 4 × 4 pixels with the collection electrode of each pixel connected to an output pad via a buffer chain, enabling the simultaneous off-chip visualization of signals from all electrodes. For a proper characterization of the sensor, the buffer chain must alter as little as possible the electrode signals. To achieve this, it is required to have a small input capacitance to avoid an excessive increase of the effective sensor capacitance and enough bandwidth to be able to resolve the sensor timing performance in the sub-nanosecond range.

The buffer chain is partially implemented in the pixel and partially in the matrix periphery. The schematic of the in-pixel circuit is shown in Fig. 4.2. The reset of the collection electrode and compensation of the sensor leakage current is performed by the PMOS transistor M0. This transistor is biased to provide a current I_{RESET} larger than the sensor leakage current. In DC conditions, as it is forced to conduct the sensor leakage current, it operates in the linear region and the voltage on the electrode is close to the potential on its source, i.e. V_{RESET}. The circuit is DC coupled to the sensor and, as done also for the developments in the TowerJazz 180 nm process, to increase the sensor reverse bias, the voltage of the p-type substrate and p-well containing the circuitry are decreased. These voltages can be biased down to −6 V, allowing to achieve a reverse bias of the sensor slightly larger than this value. Upon a particle crossing, as a negative voltage step develops on the collection electrode, the drain-source voltage of the transistor M0 increases and so does its current. For a sufficiently large input signal, the transistor enters in the saturation region and its current saturates to I_{RESET}. The current in excess of the sensor leakage charges the electrode back up to V_{RESET}, which can take several µs. The electrode signal is buffered by a PMOS source-follower stage consisting of the input transistor M2 and the transistor M1 which provides the biasing current I_{BIAS}. The buffered signal on the source of the input transistor is then given to a second NMOS source-follower stage consisting of the transistor M3 and the transistor M4 which conducts the current I_{BIASN}. The output of the second follower stage is fed back to the drain of the input transistor. Source followers are widely used in MAPS to read out the electrode voltage as they offer a low input capacitance: due to the following action on the source of the input

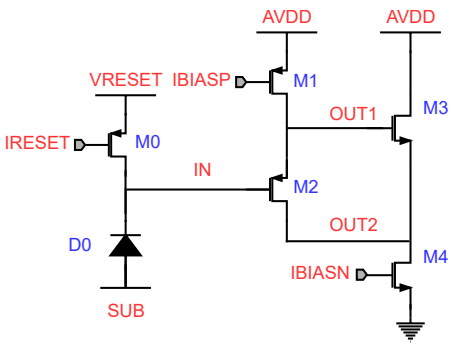


Figure 4.2: Buffer circuit in the APTS pixel.

transistor, its gate-source capacitance is not charged and therefore does not load the input node. In the implemented circuit, both the source and drain of the input transistor follow the voltage on the input node, therefore, also the gate-drain capacitance of the input transistor is compensated. In practice, the gain of the follower stages is not unitary and these capacitances are not perfectly compensated but strongly reduced. To obtain a gain close to unity, the n-well of the input transistor can in principle be connected to its source to avoid the body effect. However, this connection has not been made as it introduces a speed penalty due to the additional capacitive load on the node OUT1. The bulk of the NMOS amplifying device M3, instead, cannot be connected to its source as in the pixel a deep n-well structure to isolate it cannot be implemented. The layout of the pixel is shown in Fig. 4.3a. The pixel has a pitch of 10µm. The 1.14µm octagonal collection electrode, distanced 1.93µm from the surrounding p-well of the circuitry, is placed in the center of the pixel. The two-stage buffer is placed to the right of the collection electrode and, apart from it, the pixel integrates other analog circuitry, i.e. a testing circuit which can capacitively inject a tuneable amount of charge into the collection electrode and decoupling capacitors. The layout of the matrix is shown in Fig. 4.3b. The active matrix of 4 × 4 pixels is surrounded by a ring of dummy pixels to suppress edge effects on the core pixels which are actually read out. The dummy pixels only contain the circuitry to bias the sensor.

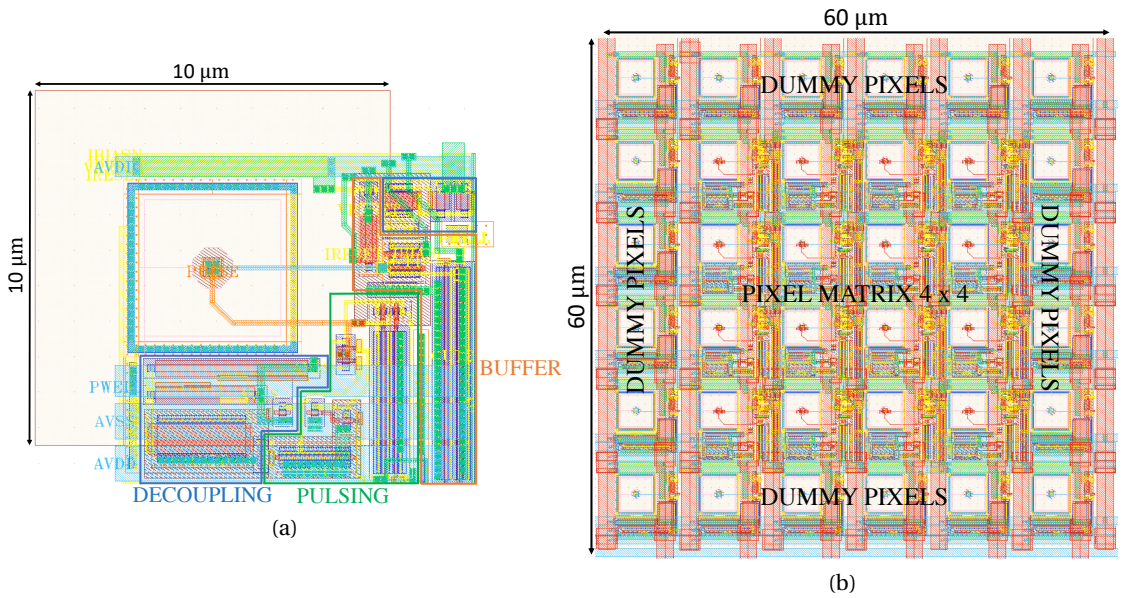


Figure 4.3: Layout of the APTS (a) pixel and (b) matrix.

The buffer circuit which connects the pixels to the output pads is shown in Fig. 4.4. This is realized by means of an OPAMP connected in unity-gain buffer configuration with a common-source stage loaded with a 500 Ω resistance. The OPAMP implements a folded cascode topology with a PMOS input differential pair. The output of the OPAMP is connected also to a second common-source circuit in 1:10 ratio with the one in the feedback loop. The second common-source stage is loaded externally with a 50 Ω resistance and therefore transmits

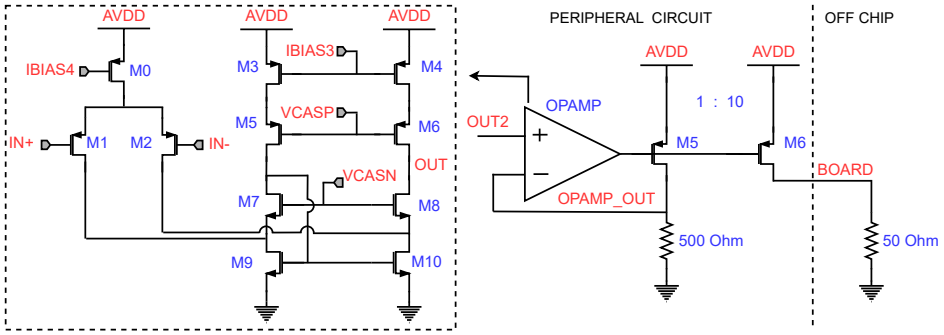


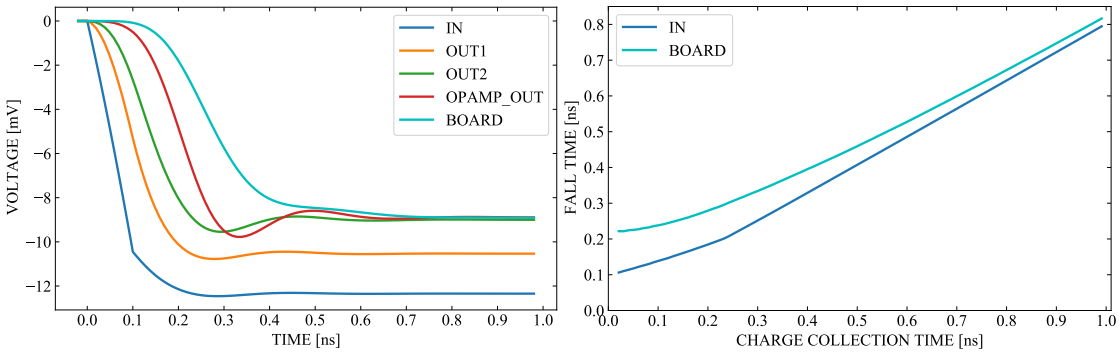
Figure 4.4: Buffer circuit in the APTS periphery.

off-chip a copy of the buffered signal generated by the internal loop. The capacitance on the output of the internal common-source stage is smaller and more controllable than the external capacitive load of the output stage. Therefore, closing the loop with an internal replica of the output driver allows to achieve a faster buffer and better stability.

A parasitic-extracted simulation of the transient waveforms (without their DC component) from the collection electrode to the output pad is shown in Fig. 4.5a. The settings used for the circuitry are reported in Tab. 4.1. The simulation was performed using a current pulse at the input, i.e. by injecting the input charge uniformly in a collection time of 100 ps. The sensor is modelled with a capacitance of 1 fF in parallel with a leakage current of 10 fA. Furthermore, the off-chip capacitive load of the output driver, given e.g. by the pad on the carrier board, is taken

Table 4.1: Settings of the APTS buffer circuits.

I_{RESET}	V_{RESET}	I_{BIASP}	I_{BIASN}	I_{BIAS3}	I_{BIAS4}	V_{CASP}	V_{CASN}	I_{BOARD}
100 pA	250 mV	10 µA	75 µA	200 µA	2.5 mA	300 mV	750 mV	5 mA



(a) (b)

Figure 4.5: Simulated response of the buffer chain in the APTS with an input charge of 150 e^-: (a) transient waveforms for a charge collection time of 100 ps (b) fall time of the signal on the electrode and on the off-chip 50 Ω resistance as a function of the charge collection time.

into account with a lumped capacitance of 2 pF. The compensation mechanism of the input transistor capacitances is limited by the reaction time of the in-pixel circuit. Therefore, the signal amplitude on the input node continues to decrease even after the 100 ps of the charge injection as the compensation evolves. The amplitude of the buffered signal monitored on the external 50 Ω resistance is 30 % smaller than the one of the input signal. The amplitude loss is entirely caused by the buffer circuit in the pixel. The fall time of the propagated signal degrades from $\approx$ 100 ps to $\approx$ 250 ps. Up to the output of the unity-gain buffer, the speed penalty is negligible. The main speed degradation is indeed given by the output stage which drives a large capacitive load. Fig. 4.5b shows the fall time of the signal at the input and on the 50 Ω termination as a function of the charge collection time. This curve shows that for slower charge collections, the speed degradation introduced by the buffer chain reduces and already for charge collection times $\gtrsim$ 700 ps it becomes negligible (<5 %). Even though the buffer chain may not be able to exactly reproduce the speeds of the sensor signals for collection times well below 1 ns, it still resolves them with a different fall time and provides insights into the timing performance of the sensor.

A version of the chip that uses two source-follower stages instead of a unity-gain OPAMP buffer in the matrix periphery to connect the pixels to the output pads has also been implemented. The schematic of the buffer chain for this version of the chip is shown in Fig. 4.6. This design is more robust and inherently stable. In this case, the output signal is monitored on a high-impedance node rather than on a resistance of 50 Ω. As a result, the trace on the carrier board that connects the chip to the signal probing point represents a large capacitive load to the last buffering stage of the readout chain. In the OPAMP-based solution, instead, the capacitance of the line on the carrier board is part of its characteristic impedance, matched to a resistance of 50 Ω. A large amount of power ($\approx$ 9.3 mW per readout channel) is dissipated in the OPAMP-based readout to extract the sensor timing information out of the chip. Due to the large capacitive load on the carrier board, even with comparable levels of power consumption, the same speeds cannot be attained with the readout based instead on source followers. This readout solution is used therefore to extract only the amplitude information of the sensor signals and, as the speed is no longer a requirement, it is operated with a much lower power

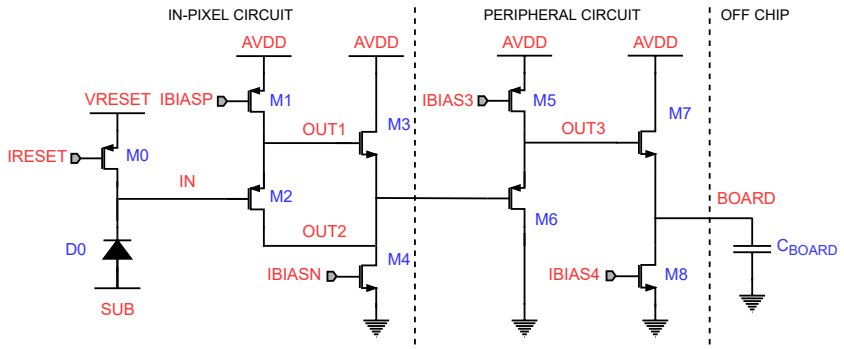


Figure 4.6: Buffer chain in the APTS version with readout entirely based on source followers.

consumption. Apart from being implemented with different readout schemes, to study the process and the sensor characteristics, the chip has been produced also with the three different sensor flavors (the standard one, the one with the additional n- implant and with the additional n- implant cut along the pixel edges, see Fig. 4.1), and with various pixel pitches (from 10 µm to 25 µm with a step of 5 µm).

A picture of the APTS is shown in Fig. 4.7. The chip measures 1.5 mm × 1.5 mm. The matrix of 4 × 4 pixels surrounded by a dummy pixel all around is placed in the center of the die. The connections between the unity-gain buffers (or third source-follower stage), placed close to the matrix, and the output stages, placed right next to the pads, have been equalized and matched to the longest one. This has been done also for the connections between the in-pixel circuits and the peripheral ones to guarantee a uniform response from all the pixels. In addition to the analog output pads and power pads, the rest of the pads in the pad ring are dedicated either to the pulsing logic, to select one or more pixels and trigger a charge injection, or to provide and externally adjust the biases of the buffer circuitry.

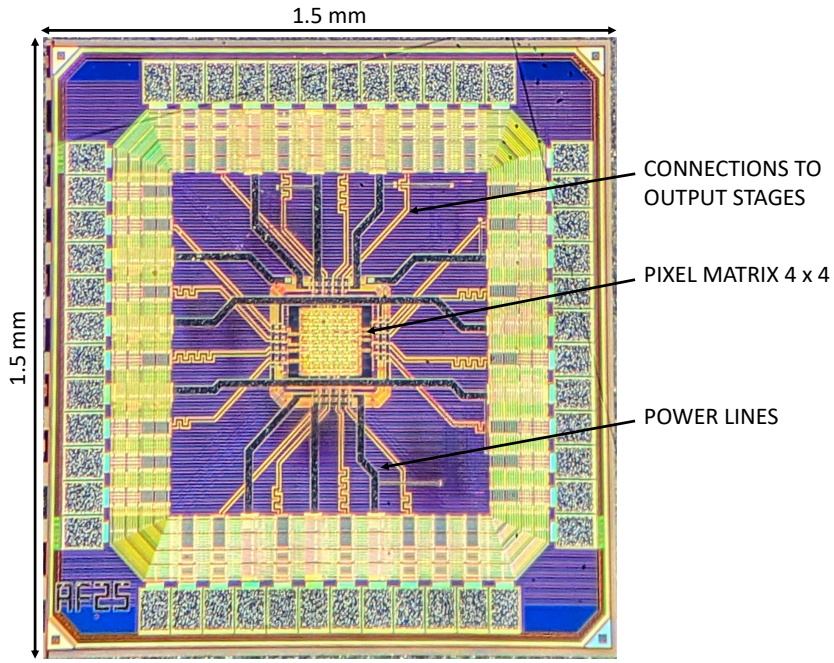


Figure 4.7: Picture of the APTS chip.

4.2.1 Characterization

A custom system which supplies biases and control signals has been used to characterize APTS chips. In the case of the APTS with readout based on source followers, the analog outputs are read out via ADCs with a sampling rate of 4 MS/s and an FPGA is used as data acquisition system. In the case of the APTS with readout based on OPAMPs in unity-gain configuration, the analog outputs are monitored via an oscilloscope with a sampling rate of 40 GS/s. The

readout chain has been characterized using the in-pixel charge injection circuitry. To fairly characterize the sensor in different biasing conditions, the operating point of the readout was optimized for a uniform behaviour with the voltage applied to the p-type substrate and p-well of the in-pixel circuitry. The latter, indeed, affects the in-pixel NMOS transistors through the body effect which appeared to be underestimated in the simulation models for significant sensor reverse biases (<1.2 V) [99]. The injection capacitance was calibrated comparing the amplitude of output signals obtained through charge injections and irradiation with an ^{55}Fe source. The main performance parameters of the OPAMP-based readout are reported with their RMS variation over the 16 channels in Tab. 4.2. The pulsing circuit injects the charge into the collection electrode in $\lesssim 100$ ps, i.e. the transition time of a digital signal buffered in the pixel. The fall time of the output signals recorded on the oscilloscope is thus comparable with the simulated one. The jitter of the readout circuitry has also been measured and it is reported as a function of the injected charge in Fig. 4.8. For this measurement, two pixels have been pulsed a thousand times with the same charge-injection trigger signal. The jitter is then evaluated as the RMS of the difference between the times of arrival of the two output signals on the oscilloscope, considered as the times when the signals cross 50 % of their excursion. For an input charge of 300 e^-, which is ≈ 10 times larger than the readout ENC, the jitter is ≈ 28 ps whereas it reduces to ≈ 5 ps for large input charges ($\gtrsim 1500\,e^-$).

Table 4.2: Measured performance of the APTS readout chain with unity-gain OPAMP buffer.

Parameter	Value
Power	9.3 mW/channel
Gain	68 % ± 2 %
Fall time	245 ps ± 8 ps
ENC	27.2 e^- ± 2.6 e^-
Jitter (with ≈ 300 e^-)	28.3 ps ± 0.8 ps

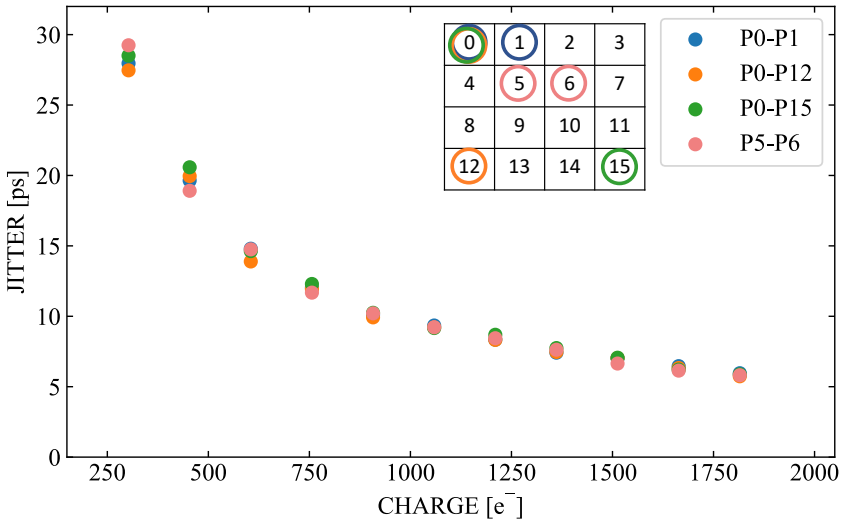


Figure 4.8: Jitter of the buffer chain in the APTS as a function of the input charge.

The properties of the sensor have been evaluated mainly by exposing chips to an ^{55}Fe source. Fig. 4.9 shows the spectrum of the charge collected for each event by the seed pixel, i.e. the one with the largest signal, for the different sensor flavors. For the standard sensor, the characteristic K_α peak at 1640 e^- and K_β peak at 1800 e^- of an iron source are rather low. A large peak is present instead at low charge values which indicates that the generated charge is most of the times shared among more pixels. For the modified sensor with the low-dose n-implant, the charge sharing peak is strongly suppressed and the characteristic iron peaks are much larger. In this case, therefore, the generated charge is mainly collected only by the seed pixel. In the case of the sensor with cut in the n- implant, the characteristic iron peaks are even more pronounced over the charge sharing one, indicating that this additional modification actually enhances the lateral component of the electric field and additionally reduces charge sharing.

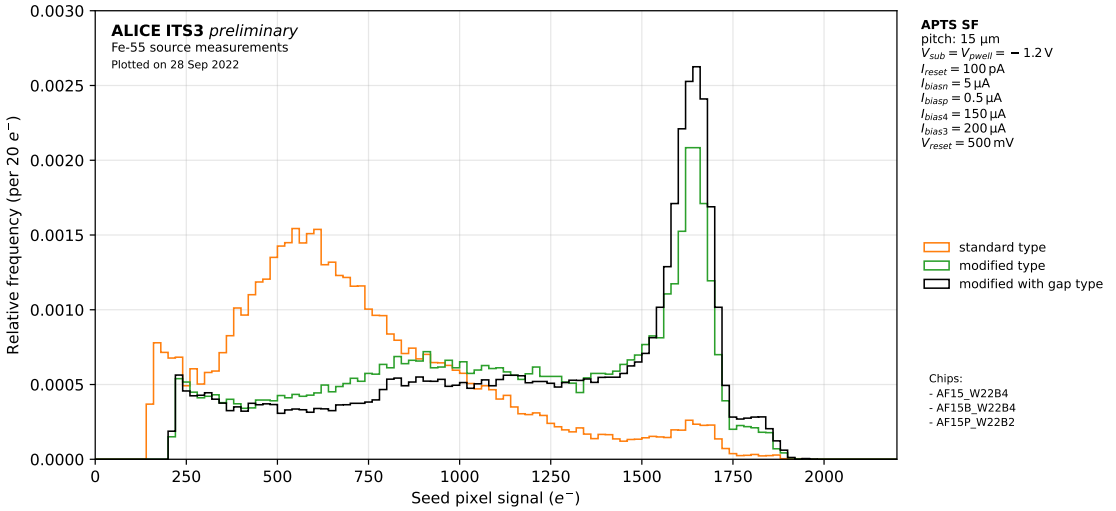


Figure 4.9: Spectrum of the charge collected by the seed pixel under exposure to a ^{55}Fe source for the different sensor flavours [100].

The impact of the sensor optimizations is further highlighted by comparing for the different sensor flavors the spectrum of the charge collected by the seed pixel under exposure to a ^{55}Fe source with different pixel pitches. Fig. 4.10 shows this for the standard sensor. In this case, larger pixel pitches increase the probability of charge sharing as indicated by the increase of the charge sharing peak and reduction of the characteristic iron peaks. The same comparison is made in Fig. 4.11 for the modified sensor with cut in the the low-dose n- implant. As also pointed out from Fig. 4.9, the sensor modifications strongly suppress the charge sharing. Furthermore, for this sensor flavor, in stark contrast with the standard one, the pixel pitch has a minimal impact on the charge sharing as the spectra in all the cases are nearly identical.

In order to study the sensor capacitance, rather than the spectrum of the collected charge, the spectrum of the signal amplitude on the seed pixel should be analyzed instead. Fig. 4.12 shows this spectrum for the modified sensor with the cut in the low-dose n- implant and

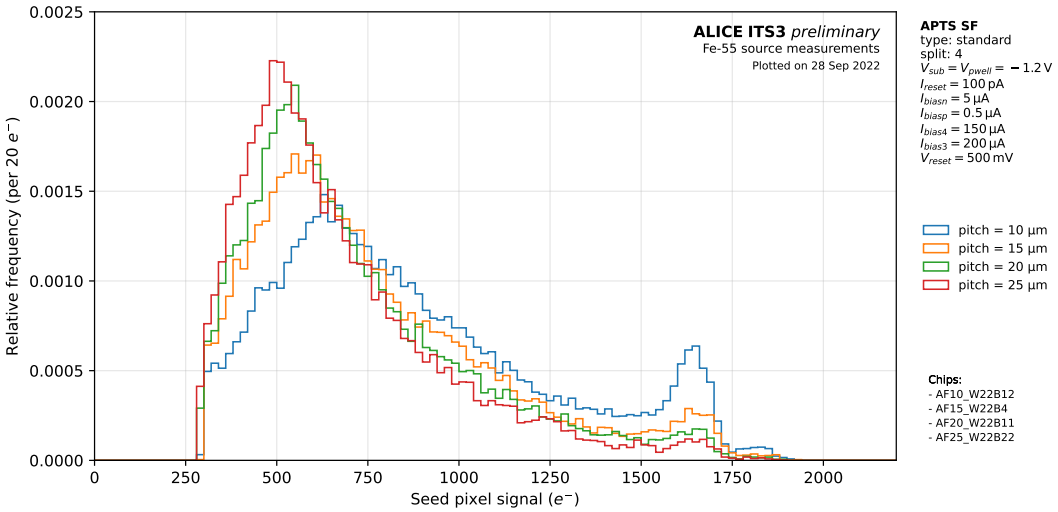


Figure 4.10: Spectrum of the charge collected by the seed pixel under exposure to a ^{55}Fe source for the standard sensor and different pixel pitches [100].

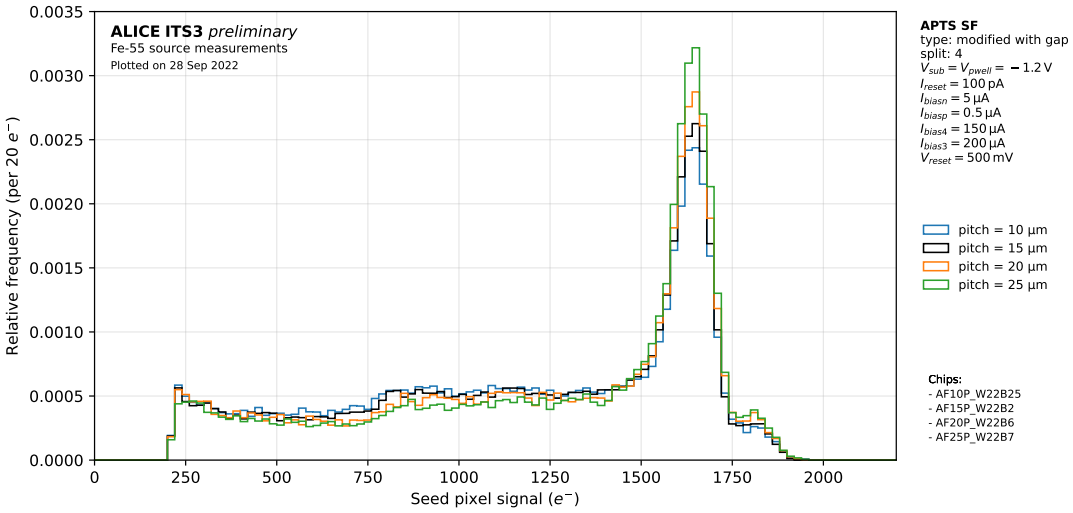


Figure 4.11: Spectrum of the charge collected by the seed pixel under exposure to a ^{55}Fe source for the modified sensor with cut in the low-dose n- implant and different pixel pitches [100].

different bias voltages applied to the p-type substrate of the sensor. The plot shows that for lower substrate voltages the amplitude of the characteristic iron peaks are shifted to larger amplitudes revealing a reduction of the sensor capacitance. A lower substrate voltage, indeed, increases the sensor reverse bias, enhances the depletion of the epitaxial layer and, more importantly, of the low-dose n- implant around the collection electrode leading to this effect. The value of the sensor capacitance as a function of the substrate voltage for the different sensor flavors is shown in Fig. 4.13. The modified flavors have essentially the same sensor capacitance, meaning that it does not vary regardless of a cut in the low-dose n- implant along the pixel edges. Furthermore, for low sensor reverse biases, the modified flavors exhibit a

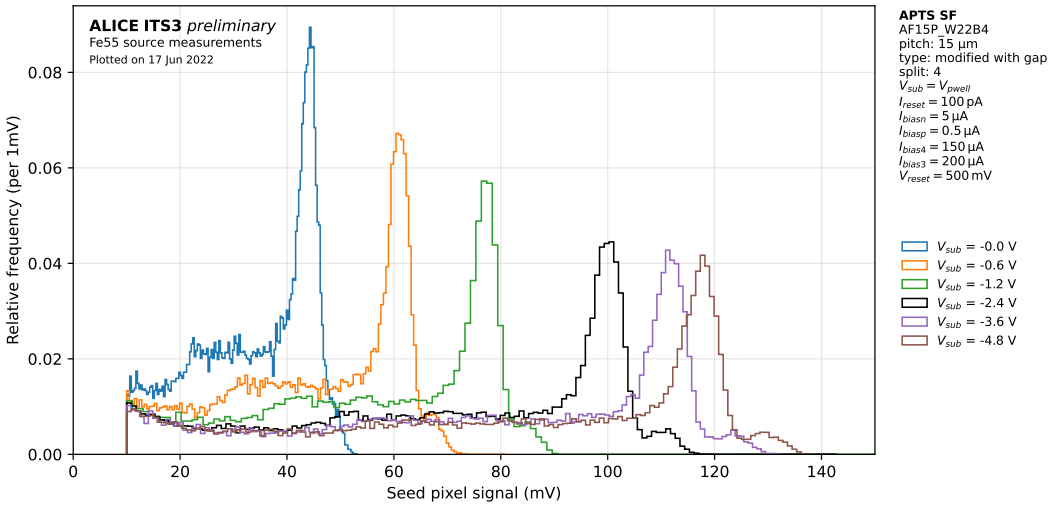


Figure 4.12: Spectrum of the signal on the seed pixel under exposure to a ^{55}Fe source as a function of the substrate voltage for the modified sensor with cut in the low-dose n- implant.

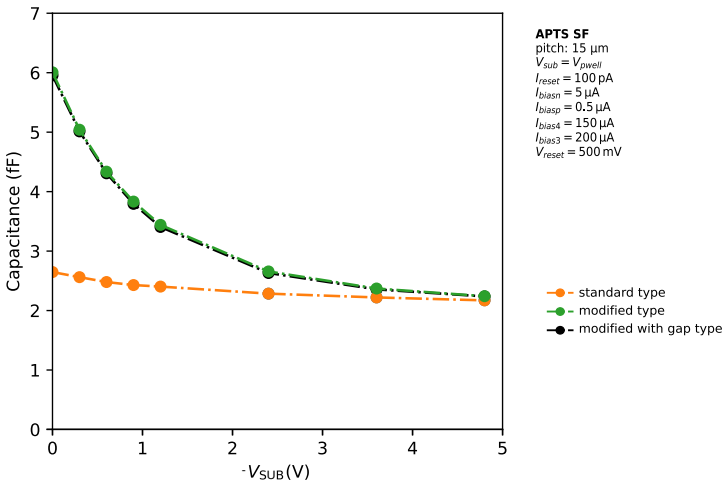


Figure 4.13: Sensor capacitance as a function of the substrate voltage for the different sensor flavors.

larger capacitance than the standard one. If the low-dose n- implant is not fully depleted up to the n-well of the collection electrode, indeed, the process modifications result in a penalty on the sensor capacitance. As a larger depletion of the this implant is achieved for larger sensor reverse biases, this penalty reduces. As shown in the Fig. 4.13, at a substrate voltage of -4.8 V, all the sensor flavors feature a capacitance of ≈ 2.2 fF and the penalty is thus negligible. It is worth noting that this value includes not only the junction capacitance of the collection electrode but also contributions from the in-pixel circuit such as the input routing line, the gate capacitance of the input transistor and the drain capacitance of the reset transistor.

The timing properties of the sensor can be analyzed with the measurements reported in Fig. 4.14. These plots show 2D distributions of the output signals from the seed pixel in an

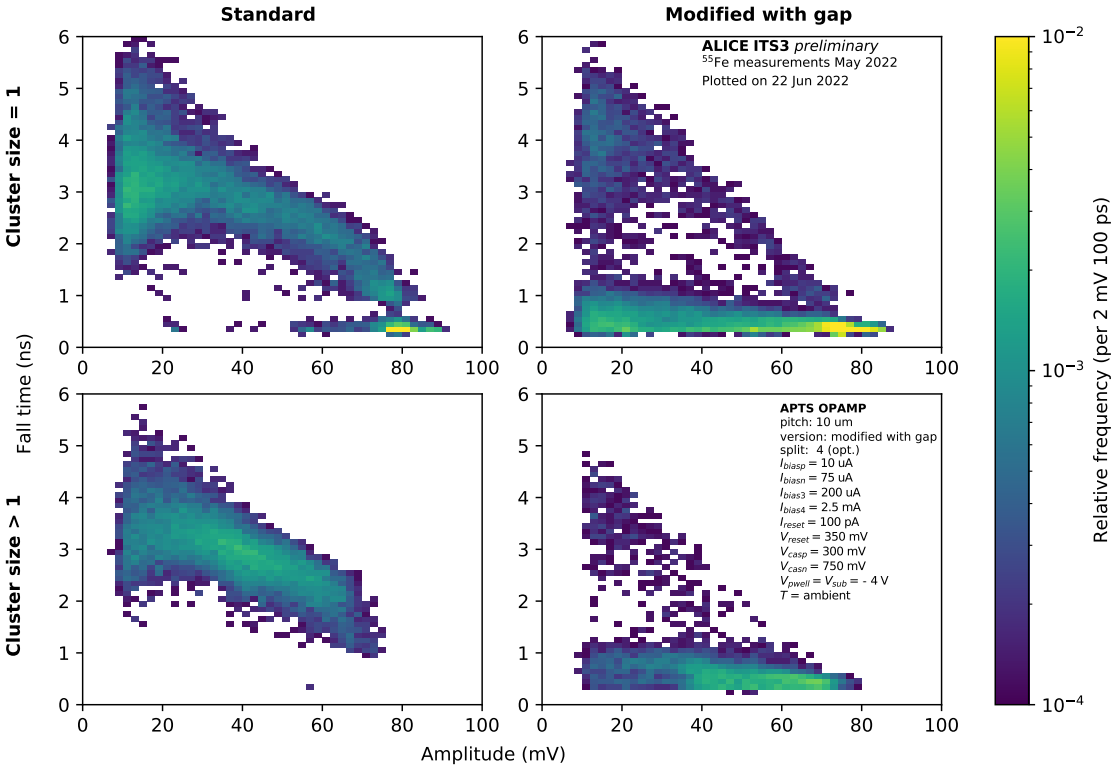


Figure 4.14: 2D distributions of the output signals from the seed pixel in an amplitude-fall time plane obtained by exposing the chip to a ^{55}Fe source comparing different sensor flavors and cluster sizes [101].

amplitude-fall time plane obtained by exposing the chip to a ^{55}Fe source. The distributions in the left column are obtained with the standard sensor whereas the ones on the right with the modified sensor with cut in the low-dose n- implant. The distributions on the top row report only the events with a cluster size equal to 1 whereas the ones in the bottom row the events with cluster size > 1. The cluster size is defined as the number of pixels firing with the same particle hit. For the standard sensor, the distribution of events with a unitary cluster size shows a peak around fall times $\lesssim 400$ ps and amplitudes of ≈ 78 mV. This peak is the result of charge generated in the depleted volume of the epitaxial layer and therefore collected by drift by a single pixel. In the distribution of events with a cluster size > 1, instead, all the events are broadly spread in the region with fall times $\gtrsim 1$ ns and amplitudes $\lesssim 70$ mV. These events are the result of charge generated in the non-depleted volume of the epitaxial layer which is collected mainly by diffusion and by more than one pixel. In the case of the modified sensor, the distribution of events with unitary cluster size still peaks around fall times $\lesssim 400$ ps and amplitudes of ≈ 72 mV (slightly lower due to a larger capacitance). A similar but broader distribution is obtained considering events with cluster size > 1. Indeed, most of the events are concentrated in the region with fall time $\lesssim 1$ ns regardless of their cluster size. This indicates that the charge is collected by drift even in the pixel corners where it can be shared among more pixels. The sensor modifications therefore allow depletion of the epitaxial layer over

the entire pixel area and accelerate the charge collection which occurs typically in less than a nanosecond.

In order to study the NIEL effects on the sensor, a number of APTS samples have been irradiated with neutrons at the TRIGA reactor in Ljubljana [95] up to a NIEL fluence of $10^{16}\,1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$. The chips also received for every $10^{15}\,1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$ a background TID of 1 Mrad. During irradiation, the chips were not powered. Afterward, the chips are stored at low temperature (below −20 °C) to prevent annealing of the radiation effects. The measurements on these samples are however performed at 14 °C. Fig. 4.15 shows the spectrum of the charge collected by the seed pixel by exposing chips irradiated at different levels of NIEL fluences to a ^{55}Fe source. For a fair comparison, all the chips in the figure implement the modified sensor with cut in the low-dose n- implant and are operated with the same bias settings and temperature. Up to a NIEL fluence of $10^{14}\,1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$, the spectra are very similar to the unirradiated case. Beyond this value, trapping of the generated charge by the radiation-induced defects is more and more relevant. The characteristic iron peaks therefore start to reduce while events at low charges become more frequent. For the chip irradiated to $10^{16}\,1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$, most of the events feature a small collected charge.

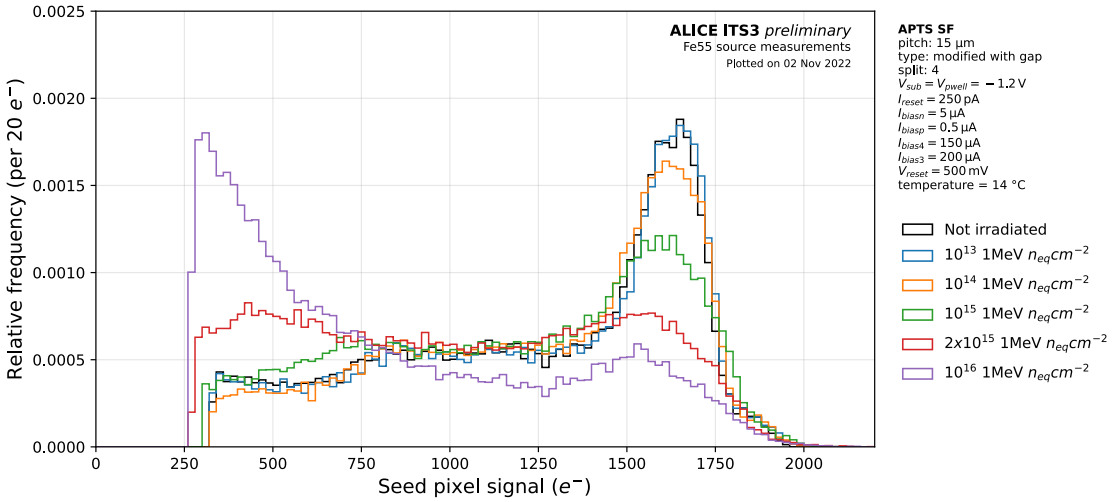


Figure 4.15: Spectrum of the charge collected by the seed pixel under exposure to a ^{55}Fe source for the modified sensor with cut in the low-dose n- implant at different levels of irradiation [100].

4.3 Digital pixel test structure

Monitoring the analog behavior of the sensor is a viable solution only for very small matrices. Already for a matrix of 4 × 4 pixels, due to the amount of pads required, the sensitive area of the chip is very small compared to its total area. For a more area-efficient solution, outputs from more pixels have to be combined and this can more easily be done in the digital domain. To characterize the sensor over larger sensitive areas, a structure with a readout chain composed

of an analog front-end amplifier and discriminator followed by a digital readout logic has thus been developed. This prototype is called DPTS. It integrates a matrix of 32 × 32 pixels with a pitch of 15 µm. In this structure, the hits from all the pixels are merged into a single line, requiring only one pad to read out the entire matrix. The structure features a novel analog front-end topology and an unconventional digital readout logic. Therefore, it has been designed not only to validate the sensor but also its operation in combination with the analog front-end circuit, as well as the concept of the digital readout.

4.3.1 Front-end

The analog front-end implemented in the DPTS is a continuously active circuit which performs the reset of the collection electrode, the amplification of the generated charge and the digitization of the amplified signal through a discrimination stage. As for the developments in the TowerJazz 180 nm technology, a conventional CSA has not been chosen as solution for the amplification stage. To avoid a noise penalty introduced by the feedback capacitor, which cannot easily be made much smaller than the sensor capacitance, and fully benefit from the low sensor capacitance offered by the technology, a solution which integrates the charge onto the sensor capacitance itself has again been pursed. A pixel pitch comparable to the thickness of the epitaxial layer, in this case 10 µm, is typically targeted. A pixel with this pitch would have an area ≈ 10 times smaller compared to those developed in the TowerJazz 180 nm technology, such as the MALTA2 pixel whose pitch is 36.4 µm. Even with a channel length ≈ 3 times smaller, more compact front-end solutions have to be explored. The front-end topology developed for the MALTA2 chip has therefore been modified to obtain a more compact design. The basic principle of this front-end is based on the transfer of charge from a capacitance to the output node. Furthermore, the same capacitance is used to couple a buffered version of the input signal to the gate of an amplifying device. Large values are required for this component to achieve high gains. The circuit has therefore been modified to remove it while still obtaining sufficiently high gains. The amplification principle of the front-end in the DPTS is shown in Fig. 4.16. The circuit is directly coupled to the sensor, represented by the diode D0. The input

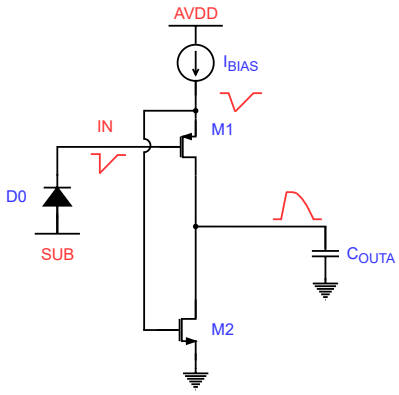


Figure 4.16: Principle of the front-end amplifier.

transistor is a PMOS device, the transistor M1, connected in source-follower configuration to reduce the contribution of its gate-source capacitance to the sensor capacitance. It is biased by the I_{BIAS} current source which is the main current of the front-end. The input transistor is loaded with the NMOS device M2. The gate of the transistor M2 is connected the source of the input device. Upon a particle hit, the negative voltage step on the collection electrode is buffered by the input transistor and provided to the gate of the transistor M2. The latter behaves as a common-source device and a voltage signal is thus obtained across the drains of the two transistors. As opposed to the previous front-end topology, in this scheme, the amplifying NMOS device below the input transistor is DC coupled to its source, removing the need of the large coupling capacitance. To avoid coupling large signals to the collection electrode, a cascode is used to move the high-impedance output node from the drain of the input transistor over another branch, as shown in Fig. 4.17a. In this circuit, as the drain of the input transistor exhibits a lower impedance and hence a lower gain, the Miller effect on its gate-drain capacitance is reduced and so is its contribution to the sensor capacitance. The I_{BIASN} current source introduced to bias the cascode draws nominally 1/10 of the I_{BIAS} current from the main branch. The output node therefore features a higher impedance compared to the previous scheme and larger gains are also achieved. In these circuits, both the current in the input transistor and the potential on its source are defined. The DC voltage on the input node must therefore be set in a narrow range of values for the input transistor to be in saturation. A more practical implementation of the circuit is shown in Fig. 4.17b. In this scheme, an input-output feedback which adjusts the input voltage has been introduced. This feedback also resets the front-end after a particle hit and compensates the sensor leakage. A small fraction of the I_{RESET} current is indeed taken by the sensor leakage current. The remaining current flows in the transistor M7 which is the feedback element connected across the input and output node. Upon a hit, as the voltage on the collection electrode drops and the output voltage rises, the gate-source voltage of the transistor M7 reduces, forcing more current from the I_{RESET} current source into the collection electrode which charges it back to its original

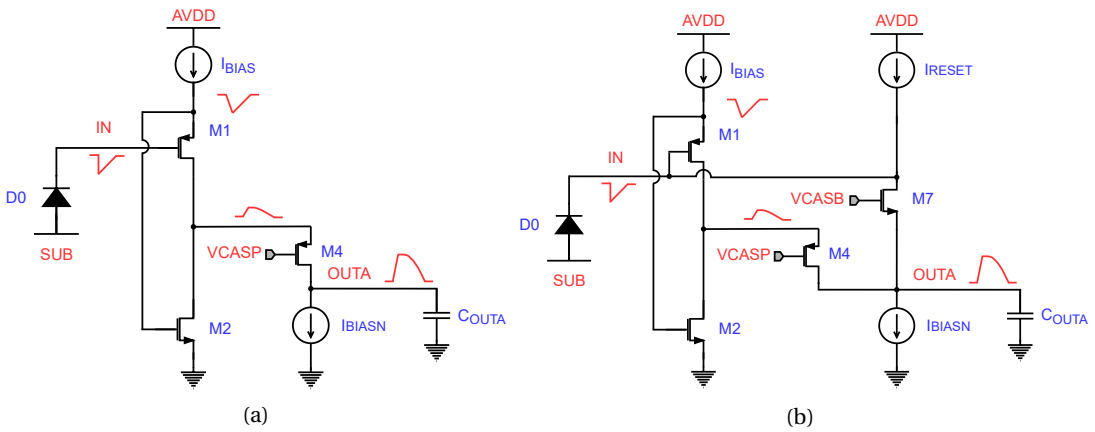


Figure 4.17: Principle of the front-end amplifier: (a) with cascode and (b) with feedback and reset mechanism.

value. The I_{RESET} current is typically orders of magnitude lower than the main I_{BIAS} current. A low value for this current is necessary to implement a sufficiently low-frequency feedback and avoid clipping the high-frequency input signals.

As for the previous topology, the circuit behaviour is non-linear. However, a small-signal analysis helps to gain insights into its operation. For the sake of simplicity, the analysis is performed on the circuit without the cascode, as shown in Fig. 4.18a. The small signal model is shown in Fig. 4.18b. In this circuit, only the relevant components are included and the source-follower input transistor is replaced with an ideal unity-gain buffer. As proved later, these simplifications still provide an accurate model. The nodal equations of the small-signal model with an input current source are:

$$\begin{cases} V_{IN}/R_{RST} + V_{IN}sC_S + I_{IN} + g_{M7}(-V_{OUTA}) + (V_{IN} - V_{OUTA})sC_F = 0 \\ g_{M2}V_{IN} + V_{OUTA}/R_O + V_{OUTA}sC_O = g_{M7}(-V_{OUTA}) + (V_{IN} - V_{OUTA})sC_F \end{cases}. \tag{4.1}$$

From this system, the transimpedance gain of the amplifier can be obtained and it is

$$\frac{V_{OUTA}}{I_{IN}} = \frac{(g_{M2} - sC_F)R_O R_{RST}}{(g_{M7} + sC_F)(g_{M2} - sC_F)R_{RST}R_O + (1 + sR_{RST}(C_F + C_S))(1 + g_{M7}R_O + sR_O(C_F + C_O))}. \tag{4.2}$$

In this equation, the expression of the output impedance Z_{OUTA} can be isolated. The latter can be found from the small-signal model by injecting a current into the output node or more easily with the systems theory as $Z_{OUTA,OPEN\_LOOP}/1 + LG$ where LG is the loop gain and is

$$Z_{OUTA} = \frac{R_O(1 + sR_{RST}(C_F + C_S))}{(g_{M7} + sC_F)(g_{M2} - sC_F)R_{RST}R_O + (1 + sR_{RST}(C_F + C_S))(1 + g_{M7}R_O + sR_O(C_F + C_O))}. \tag{4.3}$$

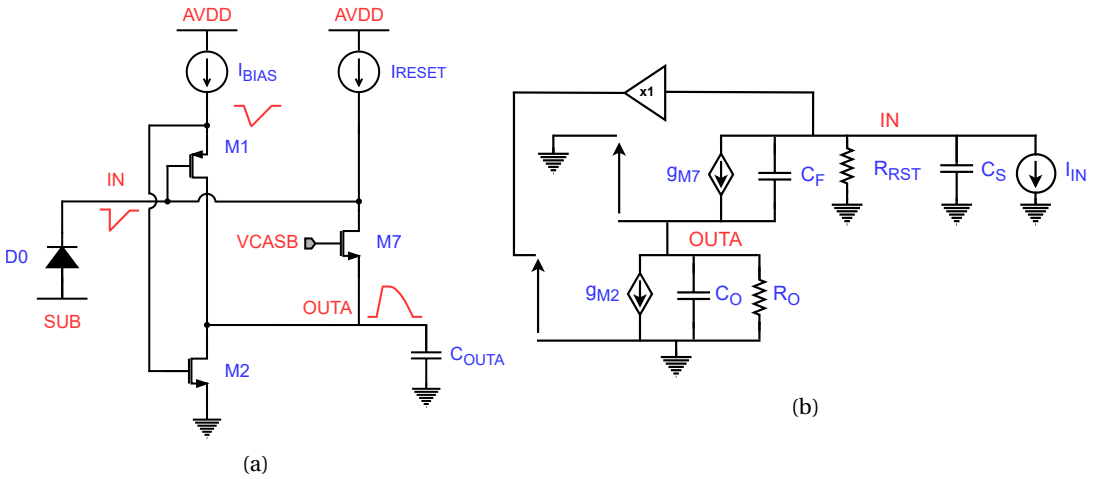


Figure 4.18: Schematic of the (a) front-end amplifier without cascode and (b) small-signal model.

The gain can therefore be expressed also as:

$$\frac{V_{OUTA}}{I_{IN}} = (g_{M2} - sC_F) Z_{OUTA} \frac{R_{RST}}{(1 + sR_{RST}(C_F + C_S))}, \tag{4.4}$$

and it is basically given by the product of the transconductance g_{M2} of the amplifying device with the output impedance Z_{OUTA}. The latter features a zero at low frequencies, given by the feedback mechanism previously explained which suppresses slow signals, and two poles given by the high-impedance nodes, namely the input (IN) and output (OUTA) nodes. As the circuit is stimulated with an input current, an additional integration pole is also present. This is represented by the last term of the equation and compensates the zero of the output impedance. The sC_F term subtracting the transconductance g_{M2} introduces an actual zero in the transfer function. The capacitance C_F indeed shunts the feedback device and represents an additional parallel path between the input and output node. However, this zero appears at frequencies well beyond the amplifier bandwidth. An AC simulation of the transimpedance gain of the amplifier is shown in Fig. 4.19. The poles on the IN and OUTA nodes occur at similar frequencies resulting in a slight peaking of the transfer function. Overall, the transimpedance gain is characterized by a low-pass response. No additional shaping is implemented after the amplification stage.

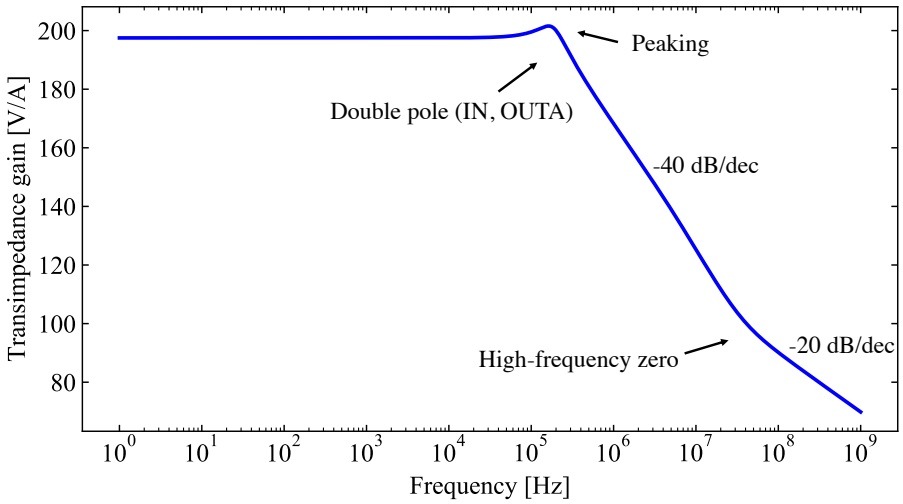


Figure 4.19: AC simulation of the transimpedance gain of the amplifier.

The complete front-end circuit which includes the amplifier and discriminator is shown in Fig. 4.20. In the amplification stage, a diode-connected NMOS transistor has been inserted between the source of the transistor M2 and the ground to shift up its source voltage. With this modification, part of the buffered signal on the source of the input transistor drops across the diode-connected device reducing the signal available to the amplifying device M2 and therefore the front-end gain. This modification is however necessary to obtain sufficient margins for the input transistor in all the operating conditions. The input transistor is placed, together with the transistor M4, in a separated n-well connected to its source to eliminate the body effect and achieve a gain closer to unity for the input follower. The I_{BIASN} current source,

implemented by the transistor M9, is cascoded to increase the output impedance and therefore the gain. The I_{RESET} current source, implemented by the transistor M5, is also cascoded to reduce the systematic variations on this current. The cascode transistor M6, as well as the feedback transistor M7, are designed with minimum width to minimize as much as possible the capacitive load on the collection electrode. The discrimination stage follows the same principle as in the previous front-end topology and implements a high-gain common-source stage, the transistors M10 and M11, better seen in this case as a current comparator. The output baseline of the amplifier is defined by the I_{RESET} current and the V_{CASB} voltage. These settings therefore also define the stand-by current in the transistor M11 and, in combination with the amplifier gain and the I_{DB} current setting of the discriminator, the overall charge threshold of the front-end. The front-end also adopts the same masking logic as in the previous topology: three switches, controlled by three different signals connected to all pixels in a row, column or diagonal, are placed between the source of the transistor M11 and ground. These allow to select a pixel and disable the output of the discriminator in case it generates an excessive noise hit rate.

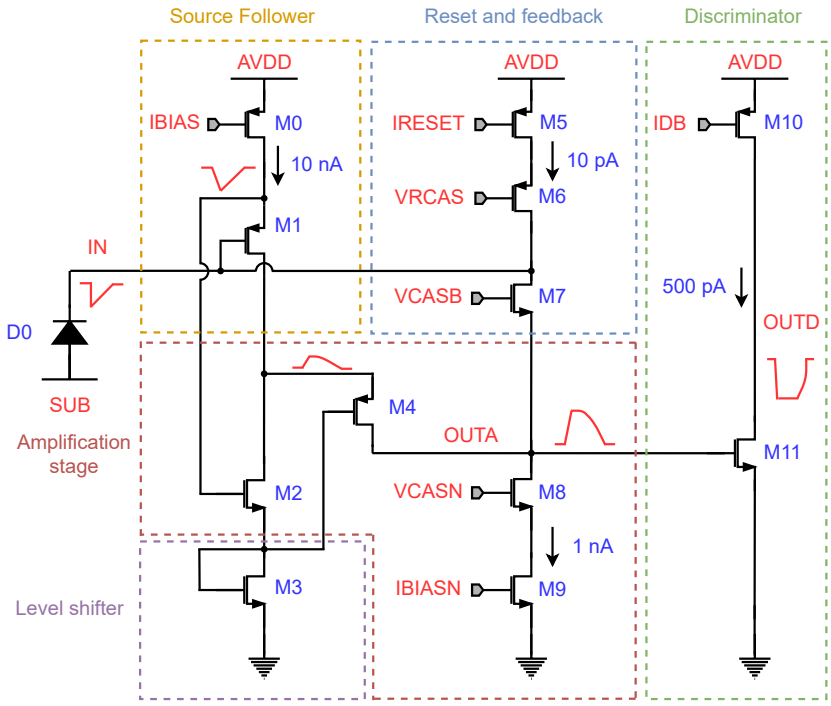


Figure 4.20: Complete schematic of the front-end with discriminator.

The front-end was designed to be within the specifications of the ALICE ITS upgrade and dissipate, for the target pitch of 15 µm, a power density of $\approx 5\,\mathrm{mW\,cm^{-2}}$ while featuring sub-µs reaction times. The circuit is therefore designed for optimized timing performance given this power budget. The bandwidth of the input follower is mainly related to the transconductance g_m of the input transistor and its load capacitance, dominated by the gate capacitance of the device M2. The gain-bandwidth product of the amplification provided by the transistor M2 is defined by its transconductance g_m and the output capacitance C_{OUTA}. Essentially, the

peaking time of the output waveform decreases with a higher transconductance g_m of the amplifying devices and a lower output capacitance C_{OUTA}. The transistors' dimensions and the layout are therefore optimized for a large transconductance g_m of the amplifying devices and a low output capacitance C_{OUTA}, which is $\lesssim 5$ fF. To satisfy the power requirement, the main biasing current I_{BIAS} needs to be within 10 nA. The I_{BIASN} current is set typically 10 times lower than I_{BIAS} and so to 1 nA. The I_{RESET} current is instead set to 10 pA, small enough to avoid filtering low-frequency components of the input signals within the bandwidth of the amplifier with these currents. The quiescent current in the discriminator can be set as low as hundreds of pA thanks to the large gain provided by the amplification stage. With a supply voltage of 1.2 V, the total power consumption of the front-end is ≈ 12 nW. Although the circuit is optimized for low power consumption, all its parameters can be varied across a wide range of values. In particular, to enhance the front-end speed, its power consumption can be increased by raising the I_{BIAS} and I_{BIASN} currents maintaining a 10:1 ratio. A parasitic-extracted simulation of the front-end with a charge threshold set to ≈ 140 e^- is shown in Fig. 4.21. The solid lines show the response for an input charge of 150 e^- whereas the dashed lines for a charge of 500 e^-. In the simulation, the sensor is modelled as a capacitance of 1 fF in parallel with a leakage current source of 10 fA. The charge is injected with a current pulse on the sensor, i.e. uniformly in a collection time of 100 ps. The red curves represent the input signals and show that the voltage step on the collection electrode is proportional to the injected charge. The blue curves represent instead the amplified signals on OUTA. The front-end gain

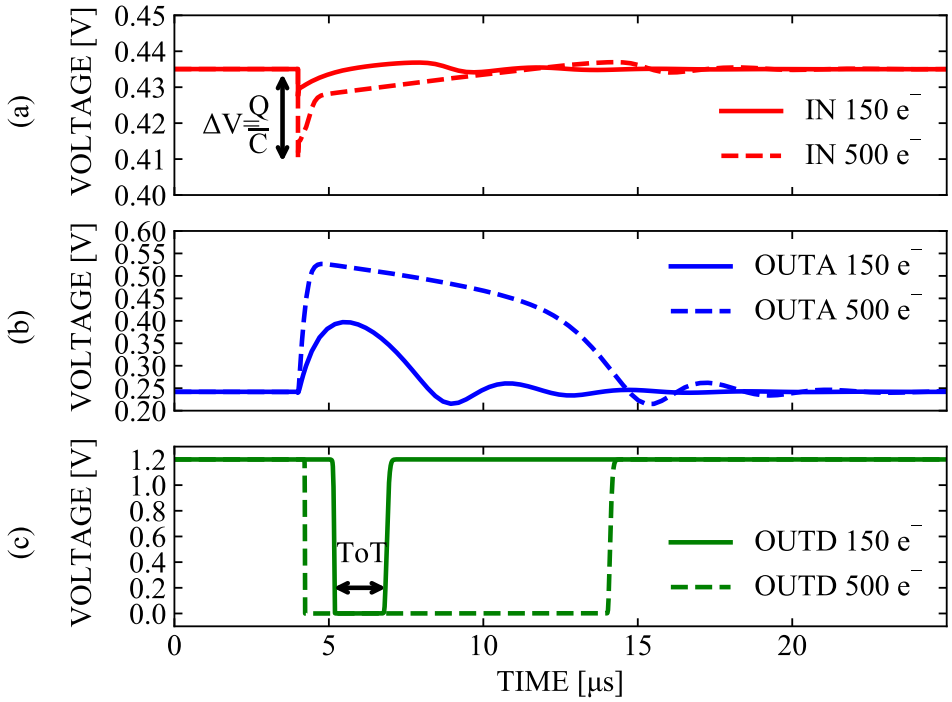


Figure 4.21: Front-end simulated transient response with a threshold of ≈ 140 e^-: (a) signals on the collection electrode, (b) signals at the output of the amplifier, (c) signals at the output of the discriminator.

is inherently non-linear: as the voltage on OUTA rises, the transistor M7 dynamically turns off, offering a larger impedance on the output node and the gain increases. Indeed, for a charge of a few electrons, the gain is $\approx 0.7\,\mathrm{mV/e^-}$ whereas it is $\approx 1\,\mathrm{mV/e^-}$ with an injected charge of $\approx 150\,\mathrm{e^-}$, i.e. around threshold, as shown in Fig. 4.21. For larger charges, the analog output signal on OUTA dynamically pushes the cascode transistor M4 out of the saturation region which makes it ineffective and the front-end gain therefore starts to drop. For an injected charge of $500\,\mathrm{e^-}$ as in the simulation, the front-end gain is $\approx 0.57\,\mathrm{mV/e^-}$. Conversely, the Time over Threshold (ToT) of the analog output signal, i.e. the duration of the discriminator output pulses shown in green in Fig. 4.21, has a linear dependence on the input charge in a wide range of values. The ToT is indeed related to the time needed for the collection electrode to be reset to its steady-state value after a particle hit. As the input charge is large enough for the analog output signal to completely shut off the feedback device M7, the I_{RESET} current entirely flows into the collection electrode which is therefore charged back up linearly with a constant current.

To enable operation of the sensor at low thresholds, the front-end has been optimized also for low ENC and pixel-to-pixel threshold variation. Apart from the main amplifying devices which have a large transfer function to the output node, a relevant noise contributor is the transistor M5 which provides the I_{RESET} current. This current is directly connected to the collection electrode and contributes to the input parallel noise. For this reason, a sufficiently low value has to be ensured to prevent it from excessively increasing the input noise. On the other hand, this current has to be higher than the sensor leakage in order for the feedback network to be able to perform the leakage current compensation. Fig. 4.22 shows the front-end s-curve obtained with transient noise simulations. The simulation has been performed with an I_{RESET} current of 10 pA, high enough to operate the chip even after some level of irradiation. The Gaussian fit to the simulation data suggests a nominal charge threshold of $\approx 140\,\mathrm{e^-}$ and a simulated ENC of $\approx 14.7\,\mathrm{e^-}$.

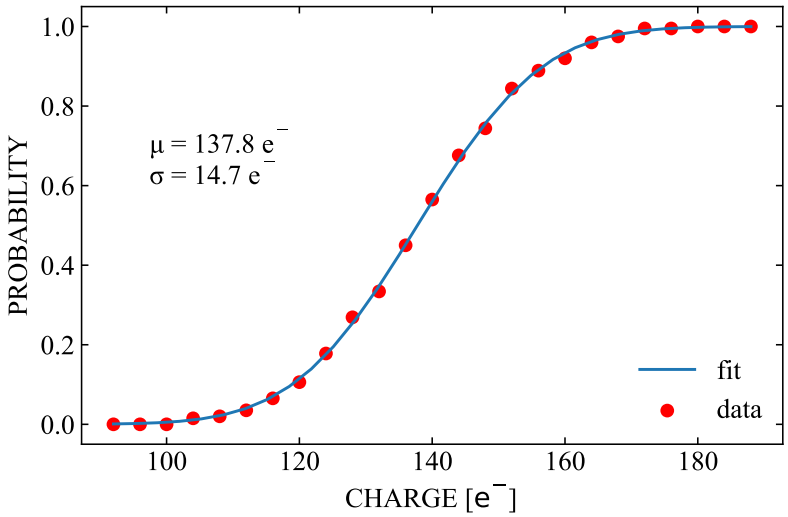


Figure 4.22: Hit probability as a function of the input charge with simulated transient noise.

Regarding the pixel-to-pixel threshold variation, the transistor M5 represents again one of the main critical devices. In fact, the I_{RESET} current defines the transconductance g_m of the feedback device M7 and has a large impact on the feedback speed and amplifier gain. For this reason, it is designed with a low aspect ratio and a large area, representing one of the largest components of the circuit. Another relevant contribution to the pixel-to-pixel threshold variation in the amplification stage is the transistor M7. The gate-source voltage of this transistor, in combination with the I_{RESET} current, defines the amplifier output baseline and thus the stand-by current in the discriminator and its switching threshold. In the discriminator stage, instead, the main critical device is the input transistor M11. A variation of its threshold voltage directly shifts the switching point of the discriminator, resulting basically in an input offset. As both the transistors M7 and M11 load the output node, they have to be kept small to prevent increasing the output capacitance and their size results from a compromise between gain, speed and threshold dispersion. Fig. 4.23 shows the s-curve obtained with Monte Carlo simulations for transistors' mismatch. The nominal threshold is again $\approx 140\,e^-$ whereas its pixel-to-pixel variation is $\approx 12.6\,e^-$.

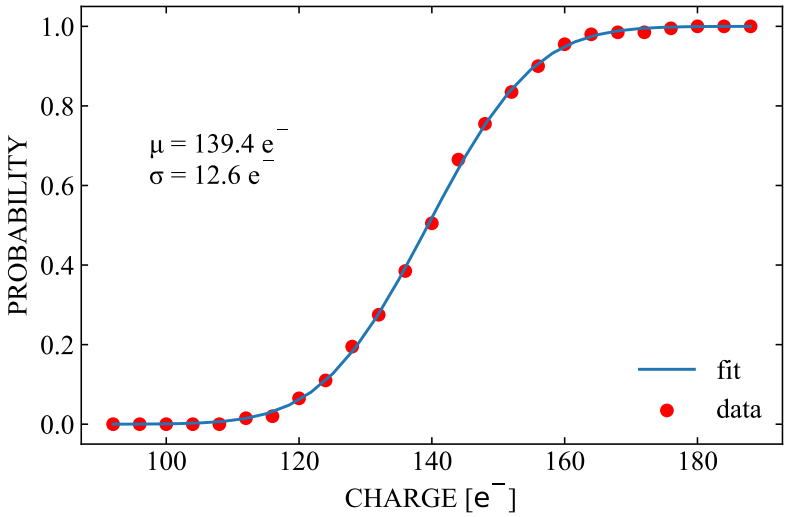


Figure 4.23: Hit probability as a function of the input charge with simulated transistors' mismatch.

The layout of the pixel is shown in Fig. 4.24. The 1.14 µm octagonal-shaped collection electrode, distanced 1.93 µm from the p-well containing the circuitry, is placed in the center. The potential on the collection electrode is adjusted by the negative input-output feedback of the front-end, typically to a few hundreds mV. As for the other prototypes, the sensor reverse bias is increased by lowering the voltage of the p-type substrate and p-well of the in-pixel circuitry, down to −6 V. The front-end is placed below the collection electrode in the layout and occupies, together with a decoupling capacitor of 20 fF, an area of $\approx 42\,\mu m^2$. A testing circuit which capacitively injects a tuneable charge into the collection electrode is integrated also in the DPTS pixel. It is placed above the sensor p-well opening in the layout and requires an area of $\approx 17\,\mu m^2$. The rest of the pixel is occupied by the gates of the digital readout. A

Schmitt trigger is placed after the discriminator to sharpen the edges of its output signal before providing it to the digital readout. The total pixel area is 15 µm × 15 µm. With this pitch, the analog power density over the matrix is ≈ 5.3 mW cm^{-2}, well within the requirements of the ALICE ITS upgrade.

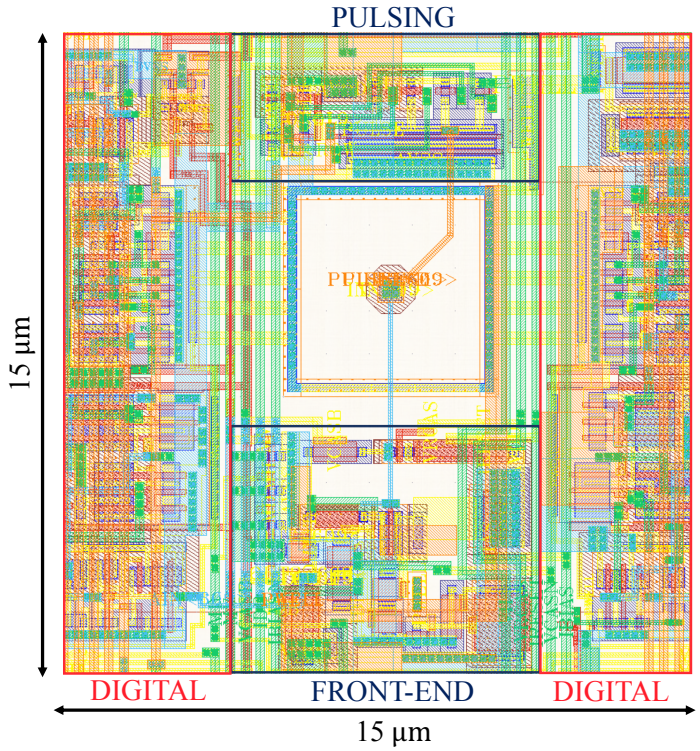


Figure 4.24: Layout of the DPTS pixel.

4.3.2 Digital readout

The digital readout implements an asynchronous event-driven logic [102]. It encodes in time the events' information, i.e. it generates a stream of pulses whose time distance and duration are in one-to-one correspondence with the coordinates of the hit pixels and also their ToT. These pulses are sent immediately upon a hit to the periphery on a single-bit bus which reads out the entire matrix.

The hit encoding circuit is independent for each column. Its components are shown in Fig. 4.25. A delay chain based on XNOR gates is distributed over a column as shown in Fig. 4.25a. The output of the front-end discriminator in each pixel is connected to XNOR gates in two mirrored positions of the folded chain. When a pixel is hit, it therefore injects transitions into the chain in these positions. The transition injected closer to the end of the chain (marked in red in the figure) is cleared by the opposite one (marked in blue in the figure) after a T_{ROW} interval. A pulse with a duration equal to T_{ROW} is thus obtained at the output of the chain. Nominally, $T_{ROW} = T_H + ROW \cdot \delta R$ where δR is the delay of two loaded XNOR gates in the chain. T_H is

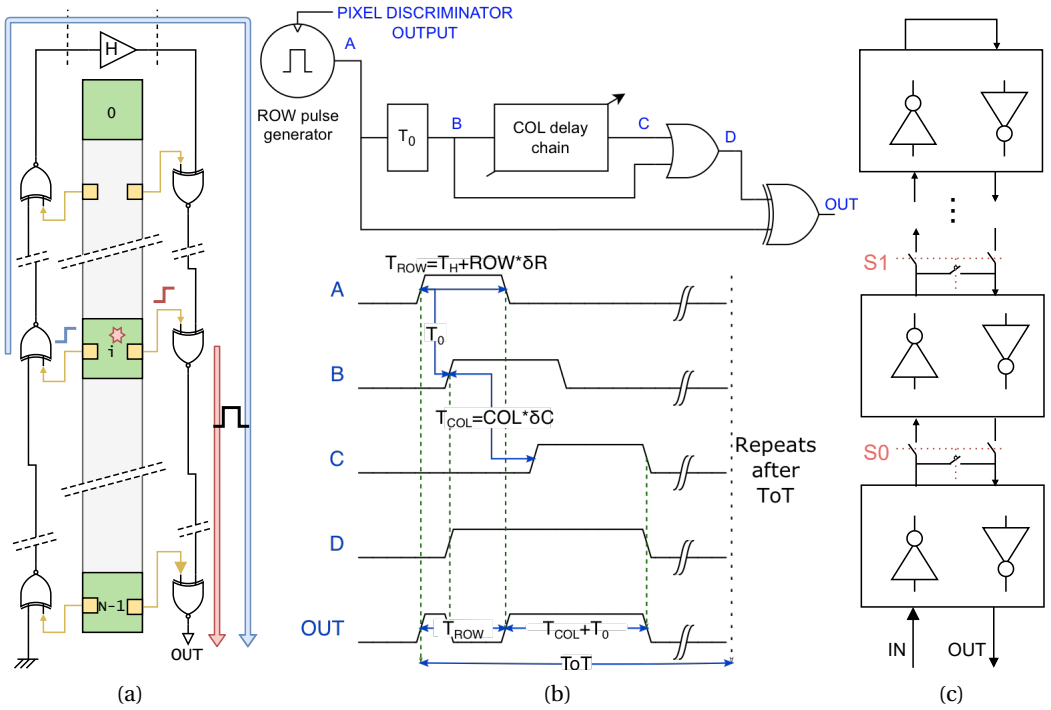


Figure 4.25: Readout components for the time encoding of the pixel coordinates and ToT: (a) row pulse generator (b) processing logic (c) column delay chain.

instead a fixed delay added at the turning point of the chain to guarantee a minimal pulse length when the top pixel is hit. The obtained pulse is further processed with the logic shown in Fig. 4.25b. Specifically, the pulse is first sent to a block which shifts it by a fixed delay T_0 and then to a second delay line which shifts the obtained signal by a duration T_{COL}, variable for each column. This delay is obtained with a chain of inverters arranged over the column as in Fig. 4.25c, i.e. similarly to the XNOR gates in the row pulse generator. At each stage, there are three switches which allow to determine the turning point of the chain and thus the length of the introduced delay. This structure is repeated in each column of the matrix sequentially increasing the number of used stages from the first to the last column thanks to the control bits of the switches. This results in a delay $T_{COL} = COL \cdot \delta C$ where δC is the delay of two loaded inverters in the chain. As shown in Fig. 4.25b, the output of this chain is OR-ed with its input to stretch this pulse by a duration T_{COL}. Finally, the output sequence is obtained by feeding to a XOR gate the original pulse with duration T_{ROW} and its delayed and stretched version at the output of the OR. The final waveform is a train of two pulses separated by a distance equal to T_{ROW}. The duration of the second pulse is instead equal to $T_{COL}(+\ T_0)$. For a better understanding, Fig. 4.25b also shows a time diagram of the waveforms. As the XNOR gates in the row pulse generator are sensitive to transitions of both polarities of their input signals, the two pulses encoding the pixel coordinates are generated upon both edges of the discriminator output allowing to reconstruct also the ToT information. The nominal values of the timing parameters obtained with parasitic-extracted simulations are reported in Tab. 4.3.

Table 4.3: Simulated timing parameters of the DPTS digital readout.

T_0	T_H	δR	δC
≈ 1 ns	≈ 4.4 ns	≈ 0.15 ns	≈ 0.15 ns

To merge the hit sequences from all the columns into a single-bit bus, the outputs of their hit encoding circuit are connected to a common XNOR chain. Encoding the hits with only four pulses, this digital readout results in a low-power logic. As the pulses are generated and transmitted asynchronously, its power consumption is strictly dependent on the hit rate. For the hit rate of the inner layer of the ALICE ITS upgrade, which is $2.2\,\mathrm{MHz\,cm^{-2}}$, the digital power density over the matrix is estimated to be $\approx 0.6\,\mathrm{mW\,cm^{-2}}$, well within the power budget allowed by the experiment. Furthermore, this digital readout requires only 8 custom gates per pixel with the largest measuring $\approx 12\,\mu\mathrm{m}^2$. Therefore, it also represents a compact solution well-suited for small pixel pitches.

4.3.3 Chip overview

A picture of the DPTS is shown in Fig. 4.26. The chip has a size of $1.5\,\mathrm{mm} \times 1.5\,\mathrm{mm}$. The matrix of 32×32 pixels with a pitch of $15\,\mu\mathrm{m}$ is placed in the center of the die. This prototype has only been implemented with the most optimized sensor flavor, i.e. the one with the low-dose

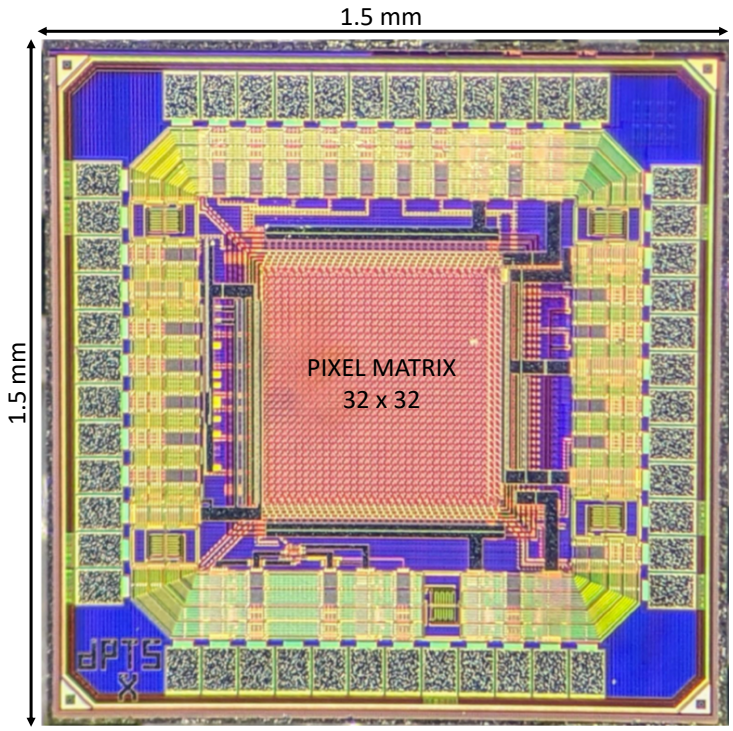


Figure 4.26: Picture of the DPTS chip.

n- implant cut along the pixel edges. A serial interface allows to communicate with the chip and write in a shift register. The latter is responsible for configuring the chip and allows to mask the pixels or to select them for charge injection, which is then triggered externally by sending a pulse on a specific interface pad. The bits of the shift register are triplicated to reduce the probability of SEU. The single-bit bus at the output of the digital readout logic is streamed off-chip via a differential CML driver. The biases of the analog circuits are provided by biasing structures in the periphery and can be adjusted externally through dedicated interface pads. The prototype features three distinguished power domains: one for the analog blocks, such as the front-end and bias circuitry, a second one for the digital blocks, such as digital readout and shift register, and a third one for the output CML buffer.

4.3.4 Characterization

The DPTS has been tested with a custom setup which allows to supply the biases, write in the control registers and sample the output of the CML driver using an oscilloscope with a sampling rate of 5 GS/s and a bandwidth of 500 MHz. The readout electronics has been tested mainly with charge injections through the in-pixel pulsing circuitry. The validity of the digital readout has been demonstrated by pulsing each pixel a thousand times while recording the output of the CML buffer. The measured time intervals encoding the pixel coordinates are shown in Fig. 4.27. The data points are gathered into clusters, one for each of the 32×32 pixels. The hits can be decoded by associating their measured time intervals to the cluster with the closest center of gravity. The clusters are non-overlapping and a proper decoding of the hits

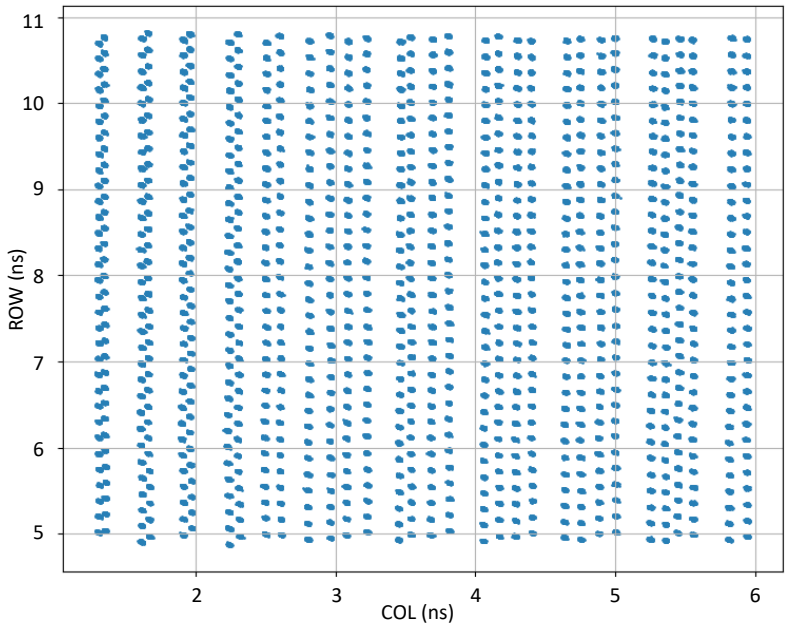


Figure 4.27: Map of time intervals encoding the row and column information of a hit measured by pulsing each pixel a thousand times.

is possible. Process, voltage and temperature (PVT) variations affect the delay of the digital gates changing the calibration map. The decoding of the hits proved to be feasible in all the measured conditions. However, the possibility of tuning the delay of the digital cells might be introduced to make the readout less sensitive to PVT variations.

To characterize the front-end with the in-pixel pulsing circuitry, the injection capacitance has been calibrated by comparing the ToT of signals obtained with charge injections and exposure to an ^{55}Fe source. The front-end speed has been measured with the time walk curve for different settings of the circuit where a faster reaction is obtained by increasing the power consumption from 12 nW up to 600 nW. The obtained results are shown in Fig. 4.28. The

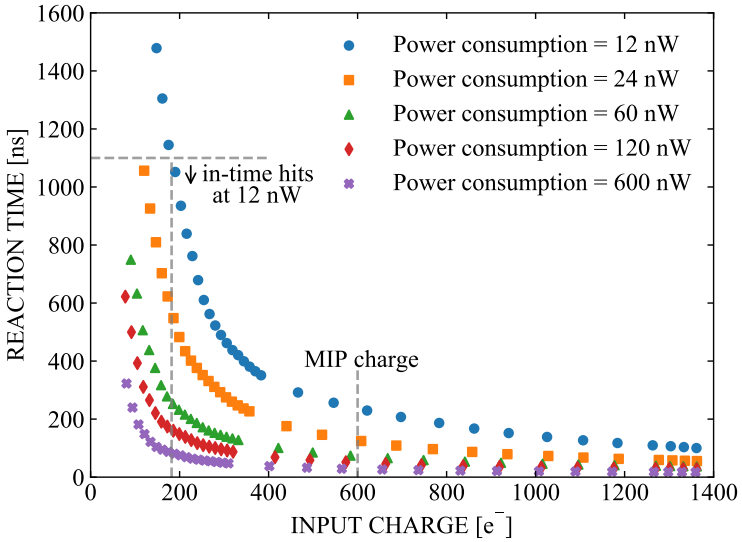


Figure 4.28: Front-end time walk curve.

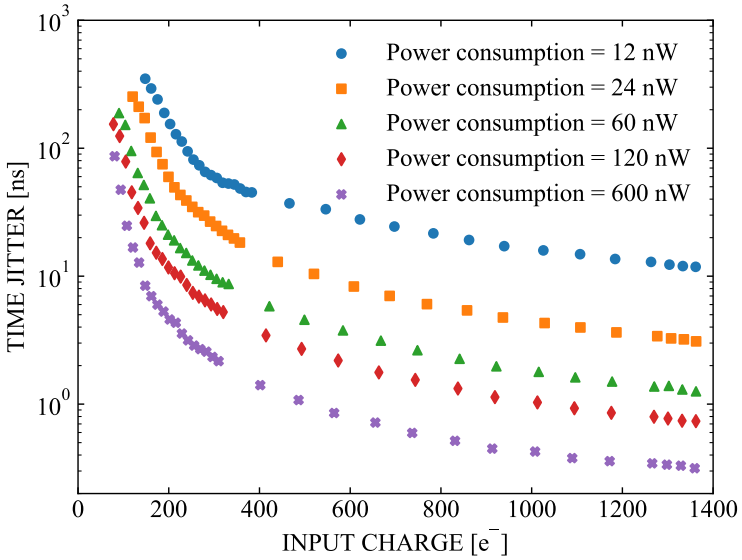


Figure 4.29: Front-end time jitter.

measurement has been performed using as time reference the charge injection trigger pulse sent to the chip. In order to do so, this signal is sent to the readout oscilloscope together with the chip CML output. The difference between the time of arrival of these two signals provides the delay of the entire readout chain which is however largely dominated by the contribution of the front-end amplifier. The plotted values are the average of the results obtained by pulsing each pixel 25 times. The RMS of these values are instead plotted in logarithmic scale in Fig. 4.29 and provide the corresponding front-end time jitter. As can be noticed from Fig. 4.28, in the lowest power mode, hits with charges $\gtrsim 1200\,e^-$ have a delay close to the minimum value of $\approx 100\,ns$. For the ALICE experiment, an event is in time if it arrives within $1\,\mu s$ from the lowest possible delay. In-time events are obtained for input charges $\gtrsim 200\,e^-$, which is $\approx 35\,\%$ of the charge released by a MIP in the epitaxial layer of the sensor. For this input charge, the front-end jitter is $\approx 150\,ns$ and reduces down to $\approx 10\,ns$ for high input charges ($\gtrsim 1200\,e^-$). If a time response within $25\,ns$ is required, as in other experiments at the LHC, this can be obtained for charges $\gtrsim 350\,e^-$ by increasing the power consumption to $600\,nW$. The larger power consumption also reduces the front-end time jitter which spans from a few ns at the in-time threshold charge, down to $0.3\,ns$ for high input charges ($\gtrsim 1200\,e^-$).

Charge test injections have been performed also to extract figures as threshold and noise for each pixel. The distributions of noise and threshold of an entire matrix operating the front-end with a power consumption of $12\,nW$ are shown in Fig. 4.30. The average threshold is $\approx 140\,e^-$ with a standard deviation of $\approx 15.8\,e^-$. The noise distribution has an average of $\approx 15.4\,e^-$. These values match fairly well the simulated ones shown in Fig. 4.22 and 4.23. Bidimensional maps of noise and threshold of each pixel are reported in Fig. 4.31. These plots show random patterns indicating the absence of systematic effects over the matrix. As for the timing measurements, the same procedure has been repeated with higher levels of power consumption and the results are summarized in Fig. 4.32. In particular, the plot in Fig. 4.32a reports the nominal thresholds as a function of the power consumption which shows a decreasing trend. As the discriminator configuration is unvaried in the different settings,

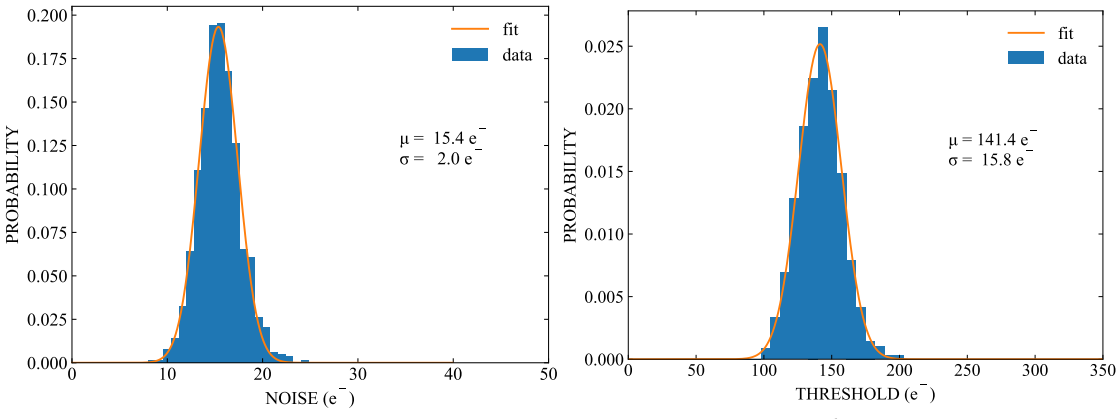


Figure 4.30: Distribution of (a) ENC (b) threshold.

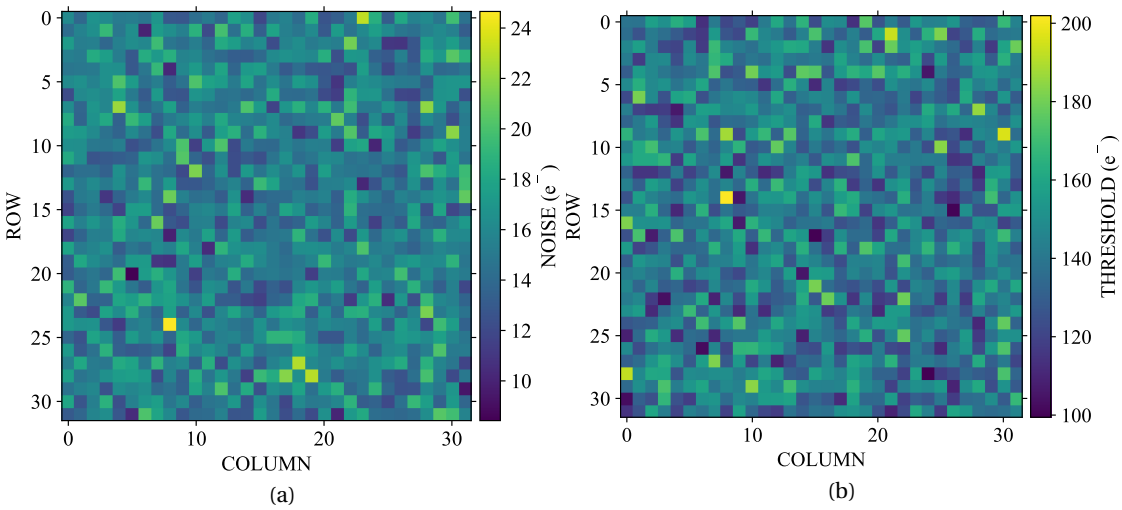


Figure 4.31: 2D map for (a) ENC (b) threshold.

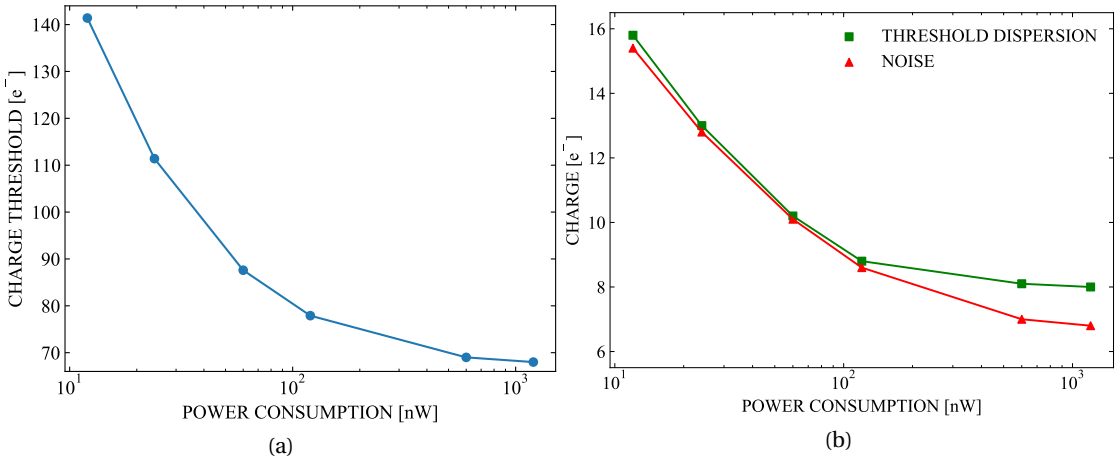
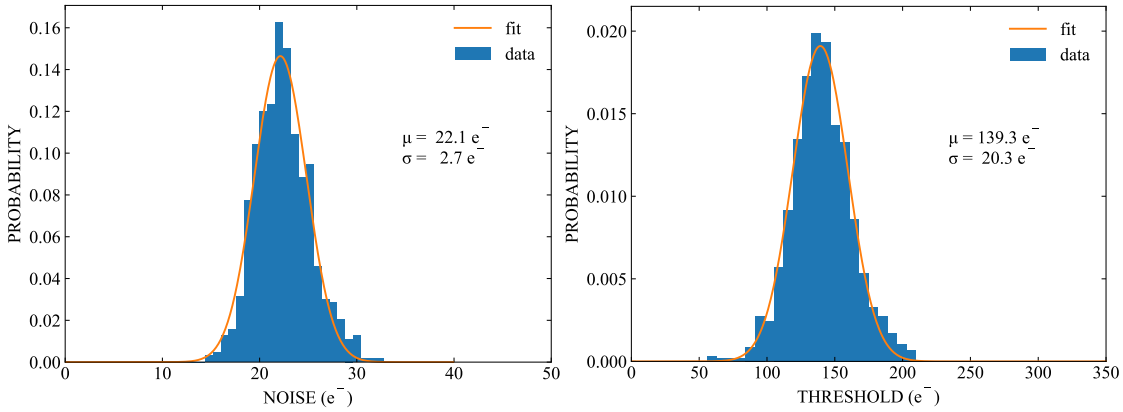


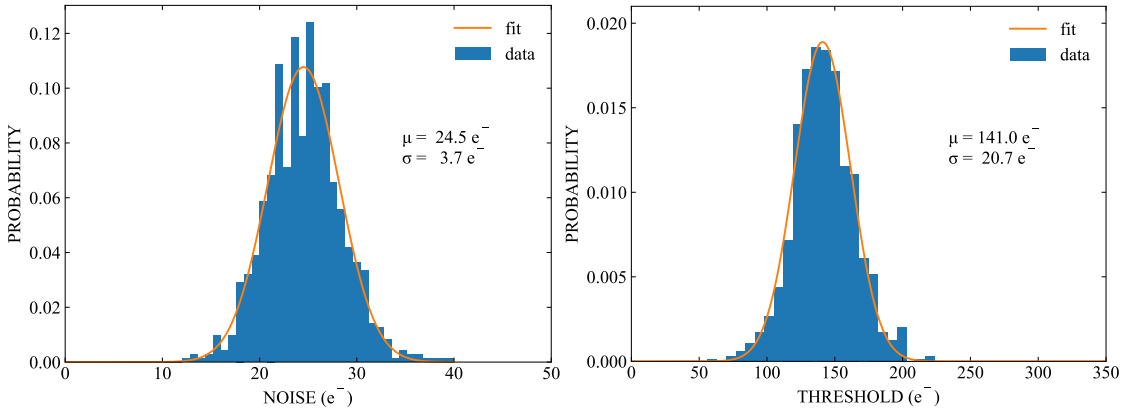
Figure 4.32: Trend of (a) nominal threshold, (b) threshold dispersion and noise with the power consumption.

this indicates a larger amplifier gain which, in combination with the higher currents, leads to a lower threshold dispersion and noise, reported in Fig. 4.32b. In all the configurations, the nominal threshold is at least 8 times larger than the ENC. During these measurements also the fake-hit rate (FHR) has been monitored. This is defined as the number of hits per pixel and second in the absence of external stimuli and is calculated as the number of hits in randomly triggered oscilloscope acquisitions divided by the their duration and total number of pixels. No pixel-by-pixel tuning of the threshold is possible with this prototype. In all the configurations, the FHR stays below a value of 10^{-2} pixel^{-1}s^{-1} with a minimal amount of masked pixels (< 5).

Also a number of DPTS samples have been irradiated with neutrons at the TRIGA reactor in Ljubljana [95]. During irradiation, the chips were not powered. Afterward, the chips are stored



(a)



(b)

Figure 4.33: Distributions of ENC and threshold with a nominal threshold of ≈ 140 e⁻ for (a) an unirradiated sample and (b) a sample irradiated at 10^{15} 1 MeV n_{eq} cm^{-2} and 1 Mrad.

at low temperature (below −20 °C) to avoid annealing of the radiation effects. The measurements on these samples are however performed at room temperature. Charge injection tests have been performed on samples irradiated up to a NIEL fluence of 10^{15} 1 MeV n_{eq} cm^{-2} and a TID of 1 Mrad due to background radiation. The chip still shows full functionality after these levels of irradiation. However, a larger I_{RESET} current has to be set for the reset network to be able to perform the compensation of the sensor leakage current, which increased due to the irradiation. For a fair comparison, tests on unirradiated samples have been repeated with the same I_{RESET} current. The distributions of ENC and threshold for an unirradiated sample and a sample irradiated at 10^{15} 1 MeV n_{eq} cm^{-2} and 1 Mrad with a front-end power consumption of 12 nW are shown in Fig. 4.33. The discriminator settings have been adjusted to obtain a charge threshold of ≈ 140 e⁻ in both cases. The ENC of the unirradiated sample is ≈ 22.1 e⁻, larger than the value shown in Fig. 4.30 due to the larger I_{RESET} current, and increases to ≈ 24.5 e⁻ for the sample irradiated at 10^{15} 1 MeV n_{eq} cm^{-2} and 1 Mrad. The threshold dispersion is ≈ 20.3 e⁻ for the unirradiated sample and marginally higher for the irradiated one at ≈ 20.7 e⁻.

To assess the front-end behaviour at higher TID levels, the chip has been irradiated with X-rays up to 50 Mrad with a dose rate of 100 krad/min. During irradiation, the chip was kept at room temperature. Furthermore, it was powered and biased to replicate the typical operating conditions. The chip remained fully functional after the irradiation. The distributions of ENC and threshold of the irradiated sample after an annealing of ≈ 30 days at room temperature are shown in Fig. 4.34. For a fair comparison with the chips irradiated with neutrons, these measurements have been performed with the same configuration used for those chips with the larger I_{RESET} current, tuning only the discriminator settings to obtain a nominal charge threshold of ≈ 140 e^-. The threshold dispersion is ≈ 20.4 e^-, comparable to the one of the unirradiated sample with the same currents and of the sample irradiated to 10^{15} 1 MeV n_{eq} cm^{-2} and 1 Mrad. The ENC is on average ≈ 27.2 e^-. This value and its dispersion are larger than the ones obtained with the unirradiated sample and also with the neutron-irradiated one.

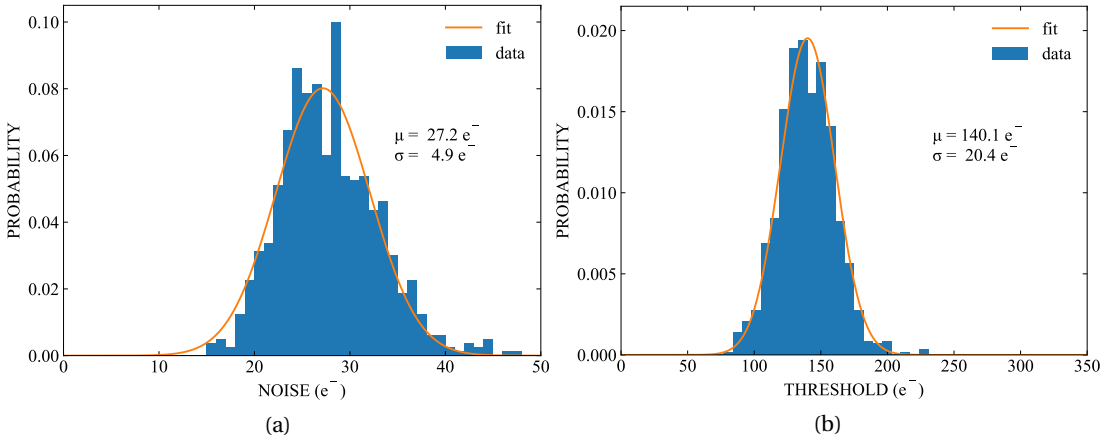


Figure 4.34: Distribution of (a) ENC and (b) threshold for a sample irradiated at 50 Mrad after an annealing of ≈ 30 days at room temperature.

The detection efficiency of the DPTS has been measured with test beams. The results discussed in the following have been obtained with the 10 GeV hadron beam provided by the CERN PS facility. For these measurements, DPTS chips have been placed in the middle of a beam telescope composed of six reference planes, three upstream and other three downstream. The measured detection efficiency and FHR as a function of the nominal charge threshold are shown in Fig. 4.35 for different voltages applied to the sensor substrate. Detection efficiencies $\gtrsim$ 99 % have been measured for charge thresholds $\lesssim$ 150 e^- regardless of the substrate voltage. In the shown range of charge thresholds, the monitored FHR is below the sensitivity limit of the measurements, i.e. below the lowest measurable FHR for the number of acquired oscilloscope captures, for most of the applied substrate voltages.

The detection efficiency and FHR of chips which received different levels of TID and NIEL are shown in Fig. 4.36 as a function of the nominal charge threshold. For all these measurements, the substrate voltage has been kept at −2.4 V. Apart from the sample which received the largest

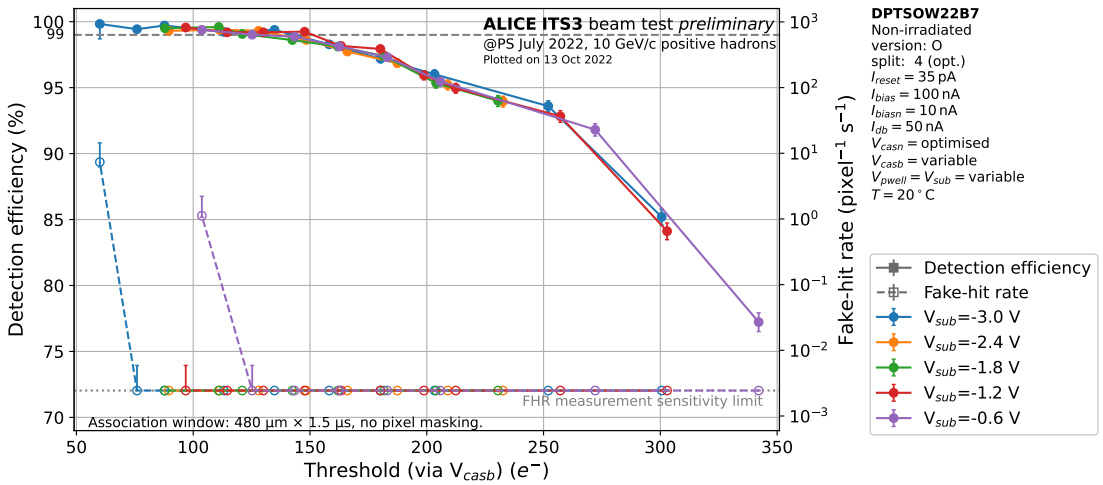


Figure 4.35: Detection efficiency and FHR as a function of the nominal charge threshold for different substrate voltages.

level of NIEL (10^{15} 1 MeV n_{eq} cm^{-2}), all the other chips still feature detection efficiencies $\gtrsim 99\,\%$ for charge thresholds $\lesssim 150\,e^-$. A detection efficiency close to 99 % can still be achieved with the chip irradiated to 10^{15} 1 MeV n_{eq} cm^{-2}, however, due to a reduction of the collected charge, it has been observed only for a charge threshold slightly lower than $100\,e^-$. Regarding the FHR, it starts increasing significantly for charge thresholds below $\approx 100\,e^-$. For the chip irradiated to the largest level of TID, the onset of the FHR is shifted to higher charge thresholds. These measurements have been performed at room temperature, while cooling is usually required to reduce the sensor leakage and the noise it induces to acceptable levels. This indicates significant margins and opens perspectives for achieving full efficiency at even higher levels of irradiation.

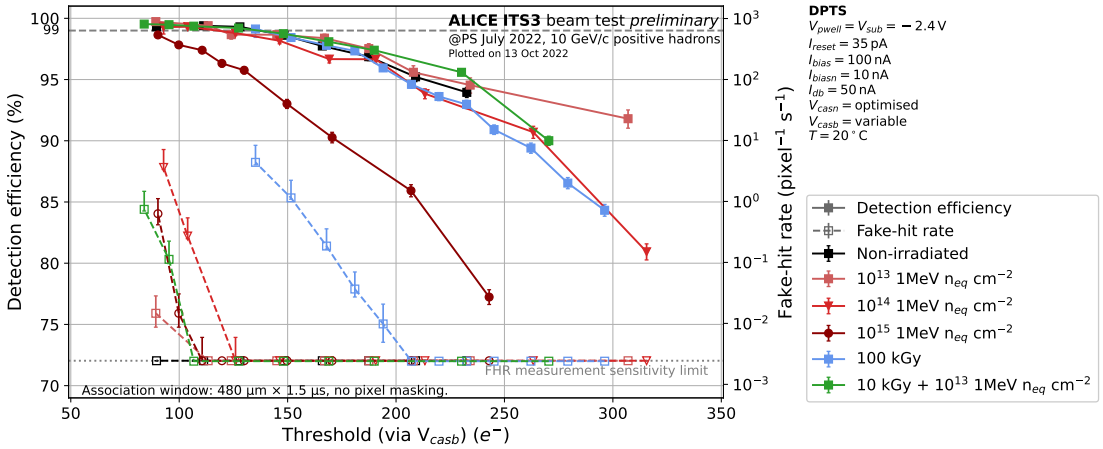


Figure 4.36: Detection efficiency and FHR as a function of the nominal charge threshold for different levels of irradiation [103].

The beam telescope allows to determine the hits on the DUT with a resolution of ≈ 2.4 µm. The detection efficiency can therefore be analyzed with sub-pixel resolution. Fig. 4.37 shows the detection efficiency of the chip irradiated to 10^{15} 1 MeV $n_{eq}\,cm^{-2}$ of NIEL for a nominal charge threshold of ≈ 150 e^- over the area of a single pixel. The values are also shown in a 1D plot as a function of the distance to the pixel center. As expected, the detection efficiency is lower in the pixel corners where the charge is generated further away from the collection electrode and thus more prone to get trapped by the radiation-induced defects during the collection process. For chips irradiated to large NIEL fluences, the detection efficiencies mainly drop in these regions, causing the average value to degrade.

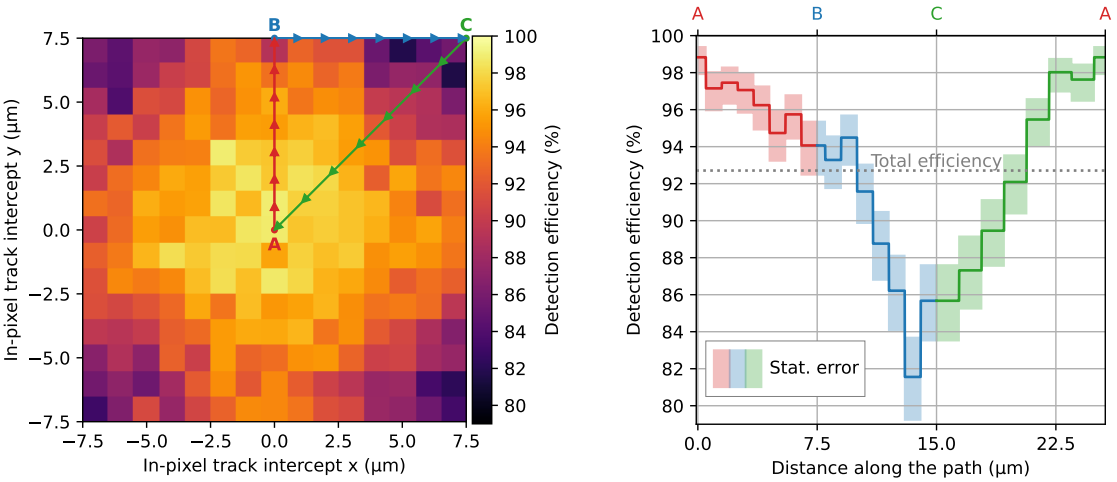


Figure 4.37: In-pixel detection efficiency of a chip irradiated to 10^{15} 1 MeV $n_{eq}\,cm^{-2}$ with a nominal charge threshold of ≈ 150 e^- [103].

4.4 Conclusions

To cope with the demand of high-granularity low-mass vertex detectors for future HEP experiments, the possibility of developing monolithic sensors in a sub-100 nm technology was explored in the framework of the EP-R&D program at CERN. The TPSCo 65 nm imaging technology was considered as a possible candidate for these applications. The chapter focused on two test structures developed to validate this technology for HEP applications, namely the APTS and DPTS.

The goal of the APTS chip is to study the analog properties of the sensor in this technology. It integrates a small matrix of 4 × 4 pixels, each connected to an output pad via a buffer chain to enable the simultaneous off-chip visualization of signals from all the electrodes. Similar process modifications to those implemented in the TowerJazz 180 nm imaging technology to accelerate the charge collection and improve the sensor tolerance to NIEL were applied also in this case. The structure validated the effectiveness of these process modifications as they considerably reduce charge sharing regardless of the pixel pitch. The sensor features a capacitance which, including the input contribution of the readout circuitry, can reach a value

of ≈ 2 fF and sub-nanosecond collection times, proving the acceleration of the charge. The latter should lead to better time resolution and this now is the object of a specific study.

Combining the pixel outputs with a digital readout, the DPTS chip aims to characterize the sensor over a larger sensitive area as it features a matrix of 32 × 32 pixels with a pitch of 15 µm. The digital readout is an asynchronous event-driven logic which encodes in time the events' information such as the addresses of the hit pixels and their ToT. The front-end topology is derived from the one implemented in the 180 nm technology to obtain a more compact solution. A summary of the front-end performance is given in Tab. 4.4.

Table 4.4: Front-end specifications with a threshold of 140 e^-.

Parameter		**Value**
Area		42 μm^2
Power consumption		12 nW
In-time threshold (for a 1 µs time window)		200 e^-
Time jitter	at the in-time threshold	150 ns
	for high charges ($\gtrsim$ 1200 e^-)	10 ns
Threshold dispersion	unirradiated with I_{RESET} = 10 pA	15.8 e^-
	unirradiated with I_{RESET} = 35 pA	20.3 e^-
	10^{15} 1 MeV n_{eq}/cm^2 and 1 Mrad with I_{RESET} = 35 pA	20.7 e^-
	50 Mrad with I_{RESET} = 35 pA	20.4 e^-
ENC	unirradiated with I_{RESET} = 10 pA	15.4 e^-
	unirradiated with I_{RESET} = 35 pA	22.1 e^-
	10^{15} 1 MeV n_{eq}/cm^2 and 1 Mrad with I_{RESET} = 35 pA	24.5 e^-
	50 Mrad with I_{RESET} = 35 pA	27.2 e^-

Measurements with test beams showed that the readout chain, including the front-end, the digital readout and the sensor optimized for faster charge collection, can achieve a detection efficiency $\gtrsim$ 99 % even after irradiation up to 10^{15} 1 MeV $n_{eq}\,cm^{-2}$ and 1 Mrad. It is important to stress that this post-irradiation performance has been achieved at room temperature, which is unprecedented. This proved that efficient particle sensors can be made with the TPSCo 65 nm imaging technology, an important achievement for the EP-R&D effort. Cooling down the sensor might allow to reach tolerance to larger NIEL levels, potentially up to 10^{16} 1 MeV $n_{eq}\,cm^{-2}$. This is currently under investigation, and if proven, this technology could basically cover the requirements of HEP experiments for the next decade.

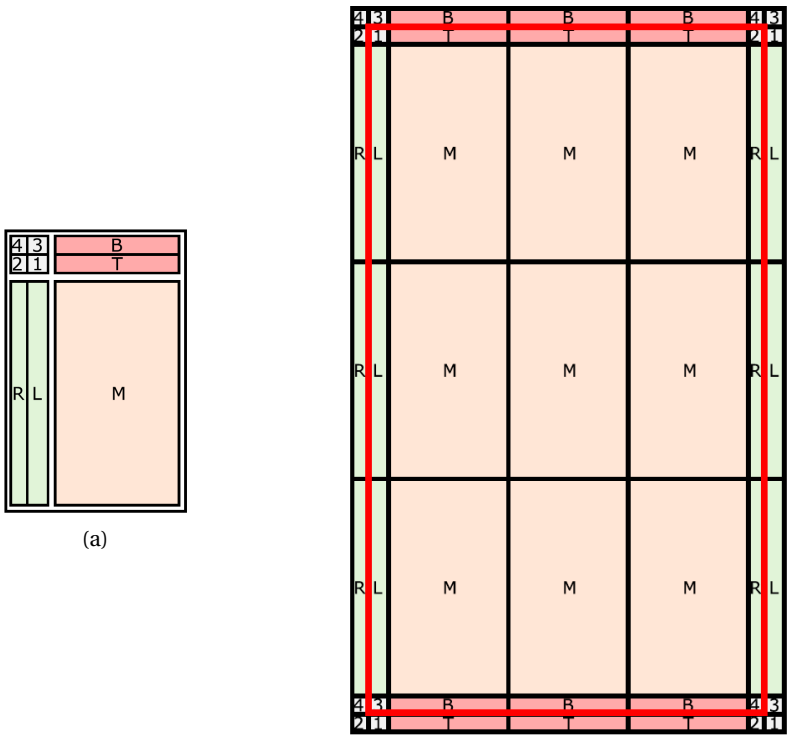
5 Design of monolithic stitched sensors in the TPSCo 65 nm imaging technology

In vertex detectors for HEP experiments, a large fraction of the material is given by the mechanical infrastructure to tile the pixel sensors and provide supplies and control signals. Large-scale sensors are preferred as they facilitate the coverage of large sensitive areas. Recent advancements in the manufacturing processes led to the development of a technique called stitching which, as opposed to a standard manufacturing process, allows producing chips larger than a reticle, the only limit being the wafer size. One of the main objectives of the EP-R&D programme at CERN is to prove the possibility of realizing wafer-scale stitched sensors for HEP applications. A key aspect of the ALICE ITS upgrade is indeed the replacement of its three innermost layers with large-scale stitched sensors able to cover their entire sensitive area [33]. In this system, tiling several sensors would no longer be needed and the distribution of power and control signals would be performed in the sensor itself, achieving a significant reduction of the detector material and improvement of its accuracy. The developments described in the previous chapter proved that efficient and radiation-tolerant particle sensors can be realized with the TPSCo 65 nm imaging process. This technology offers the stitching technique, furthermore, it uses wafers with a diameter of 300 mm, compatible with the ALICE ITS upgrade. To prove the feasibility of stitching for particle detection and gain knowledge about this technique, two different wafer-scale monolithic stitched sensors have been designed in a dedicated engineering run. These are the MOnolithic Stitched Sensor (MOSS) and the MOnolithic Stitched sensor with Timing (MOST). The MOSS features a size of 14 mm × 259 mm and a synchronous strobed readout whereas the MOST a size of 2.5 mm × 259 mm and an asynchronous event-driven readout. A primary concern in designing such large sensors is to obtain a high yield since, due to their large area, they are likely to present manufacturing defects. It is therefore important to ensure that these do not compromise the entire chip. The two sensors deal with this issue using different strategies.

This chapter describes the MOSS and MOST prototypes focusing on their main challenges and design aspects. Two different sections are dedicated to the two structures, namely section 5.2 for the MOSS and 5.3 for the MOST. First, a brief explanation of the stitching technique is given in section 5.1.

5.1 Stitching

The principle of the stitching technique [104] is shown in Fig. 5.1. The design reticle in Fig. 5.1a gets subdivided in sub-frames that correspond to sub-frames of the photomasks. During the manufacturing process, the photomasks are selectively exposed over the wafers onto adjacent locations according to a pre-established pattern, as shown in Fig. 5.1b. The various sub-units are connected together by joining their interconnections at the abutment boundaries, referred to as stitching regions. In order to achieve a well defined overlap among the sub-units and properly interconnect them, a very accurate translation and alignment of the wafers and photomasks at each exposure is required. Indeed, a special set of conservative design rules have to be respected in the stitching regions to ensure these interconnections. The peripheral structures along the outer edges of the core array and at the four corners are designed with dedicated sub-frames of the reticle. With this technique, the entire wafer can be covered and a device as large as the entire wafer can be realized. To ensure a high manufacturing yield, because of the large area of the final structure, it is strongly recommended to adhere to conservative design rules not only in the stitching regions but also within the sub-units themselves unless measures to mitigate the impact of defects are taken.



(a) (b)

Figure 5.1: Principle of the stitching technique: (a) design reticle with sub-frames (b) sub-frames on the wafer for the realization of the stitched device whose dicing lane is indicated with the red line.

5.2 MOSS - MOnolithic Stitching Sensor

The primary goals of the MOSS chip are [105]:

- to explore the feasibility of the stitching technique for particle detection quantifying the obtainable yield.
- to learn how to distribute supplies and signals in a wafer-size chip.
- to study performance parameters such as IR drops, leakage currents and spreads of characteristics.

As a first step to prove stitching, the chip exercises this technique in only one direction and it is composed of ten repeated sensor units completed on the two sides by two smaller endcap regions. A concept diagram of the chip is shown in Fig. 5.2. The repeated sensor unit is subdivided in two half units with pixel arrays of different pitches. The top half unit contains four matrices of 256×256 pixels with a pitch of 22.5µm. The bottom one contains four matrices of 320 × 320 pixels with a pitch of 18µm. Both implement the same peripheral and in-pixel circuitry, therefore, their matrices are characterized by different densities of circuits. Each half unit is a standalone powering and functional region with its own periphery and I/O pads and, therefore, it can operate independently from all the others. If a chip-killing defect occurs, its half unit, representing 1/20 of the sensor, can be switched off while keeping all the others operating. Furthermore, a possible dependence of the occurrence of manufacturing defects on the density of circuits in the matrix can be studied.

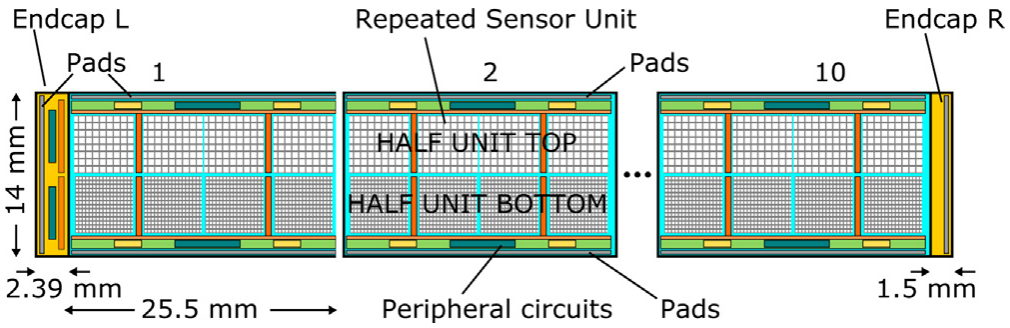


Figure 5.2: Concept diagram of the MOSS chip.

5.2.1 Front-end implementation

The analog front-end implemented in the MOSS chip is a derivation of the one developed for the DPTS chip. The schematic of this circuit is shown in Fig. 5.3 and the only difference with the circuit in the DPTS is the replacement of the diode connected NMOS transistor in the first branch with a PMOS transistor (M3), whose gate is connected to a tuneable voltage bias named V_S. As explained in section 4.3.1, the amplifier itself adjusts the potential on the

collection electrode and sets the input transistor in saturation. To increase the sensor reverse bias, the voltage of the p-well containing the circuitry and of the p-type substrate have to be lowered. Applying a voltage to the bulk of the NMOS transistors beyond the supply, even though within the maximum absolute ratings of the device, might compromise their reliability. As a high yield is a primary goal for the MOSS chip, the voltages of the p-type substrate and p-well containing the in-pixel circuitry are planned to be kept close to ground, at least in the initial phase of the MOSS testing. The introduced bias V_S overcomes partially this limitation and offers the possibility to increase the reverse bias of the sensor. In fact, it allows to shift the voltages in the input branch and, for a given setting of all the other biases, to obtain a larger voltage on the electrode, up to $\approx 900\,\mathrm{mV}$.

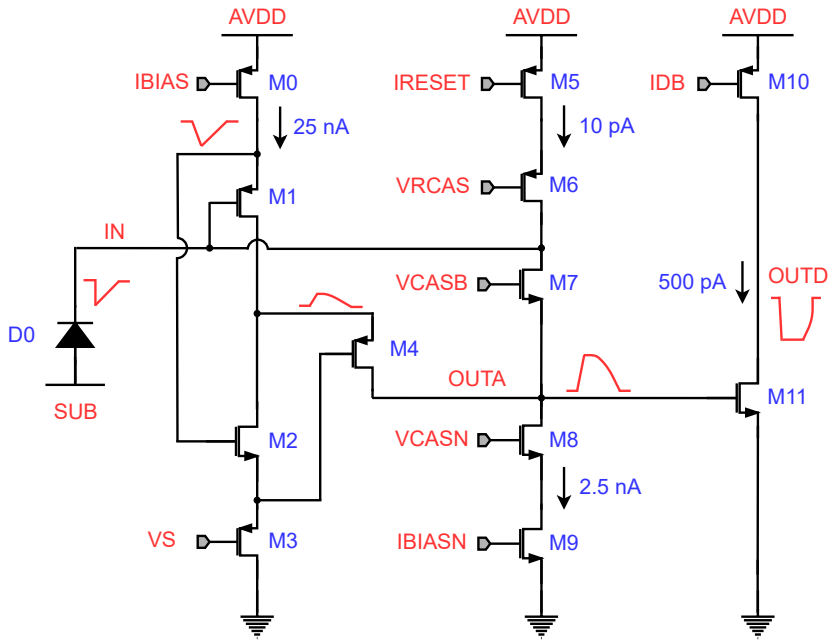


Figure 5.3: Schematic of the front-end in the MOSS chip.

Thanks to the larger pixel pitches compared to the DPTS, some of the devices critical for mismatch, as for example the transistor implementing the I_{RESET} current source, have been enlarged. Furthermore, the layout of the circuit has been designed to be very conservative meaning that all the widths and distances are drawn with large safety factors compared to their minimum values allowed by the design rules. The large-pitch pixel enables an even more conservative layout with respect to the smaller one. The main performance metrics of the front-end obtained with parasitic-extracted simulations are reported in Tab. 5.1. For these simulations, the bulk of the NMOS transistors is set to ground. The used current levels are noted on the schematic. Compared to the front-end in the DPTS, the I_{BIAS} and I_{BIASN} currents are set to larger nominal values, namely to 25 nA and 2.5 nA. Due to the absence of a negative back bias on the NMOS transistors, the headroom of the input transistor reduces and the larger current levels ensure that it has good operating margins. The sensor has been modelled with a capacitance of 5 fF which should be a representative value for the simulated

biasing conditions. The larger input capacitance used in these simulations compared to the ones of the DPTS implementation results in a lower charge gain of the amplifier. Due to this, the threshold dispersion does not considerably improve despite the enlargement of the transistors' size enabled by the larger pixel pitches. Similarly, the ENC does not significantly change even though larger current levels have been set in the simulations.

Table 5.1: Simulated performance metrics of the front-end in the MOSS chip.

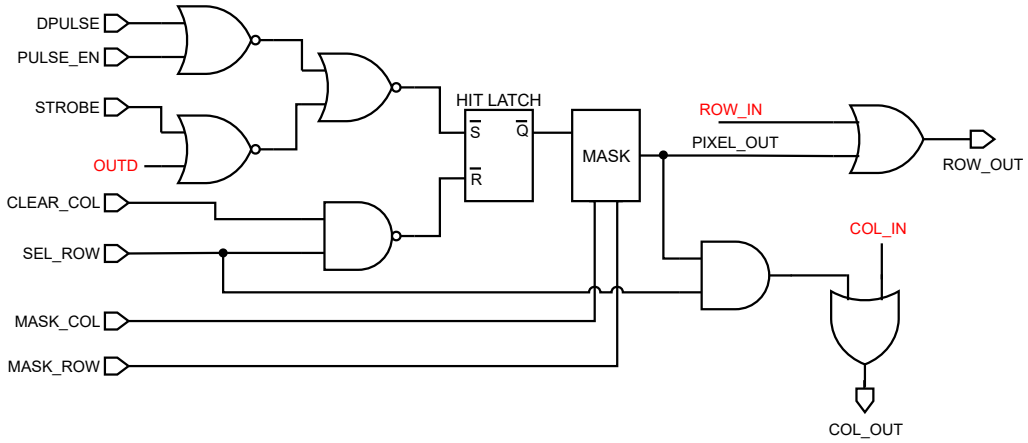
Gain	Charge Threshold	Peaking Time	Phase Margin	Thr. Disp.	ENC
$\approx 0.5\,mV/e^-$	$\approx 150\,e^-$	$\approx 1\,\mu s$	$\approx 60^\circ$	$\approx 11.5\,e^-$	$\approx 12\,e^-$

Taking advantage of the multiple matrices, different versions of the front-end have been implemented. For a fair comparison, these differ from a baseline solution only for the enlargement of single devices. Specifically, the width of the input transistors M1, of the common-source transistor M2 and of the input transistor of the discriminator M11 have been selectively doubled in these variants. No RTS noise has been observed during the measurements of the DPTS chip. With a larger number of pixels, the MOSS chip can provide a better characterization of the RTS noise and, thanks to the first two front-end variants, a possible dependence of the RTS noise on the size of the main amplifying devices (M1-M2) can be observed. The variant with a larger input transistor in the discriminator possibly provides insights into the sizing of this transistor. The latter, indeed, requires careful optimization since it is one of the main sources of threshold dispersion and increasing its size would also increase the output capacitance of the amplifier, resulting in a lower gain and thus larger threshold dispersion.

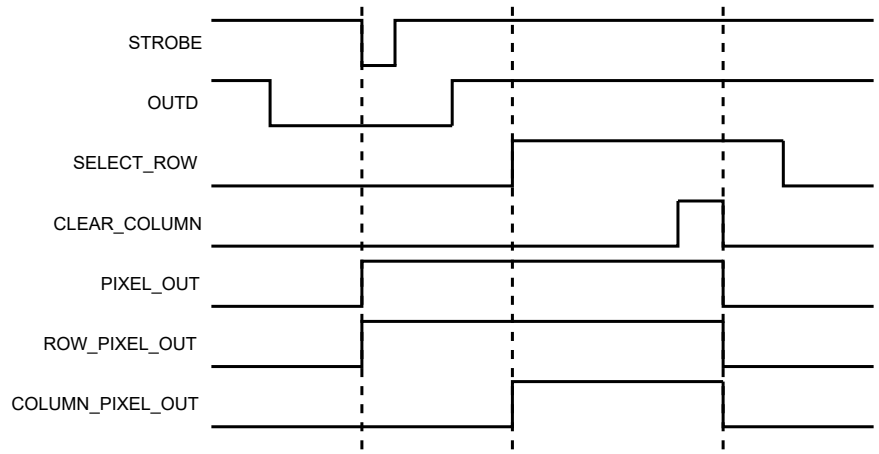
5.2.2 In-pixel digital logic

The MOSS chip features a synchronous strobed readout. The hits in the matrix are read out upon transmission of a global STROBE signal to all the pixels. The in-pixel digital logic is responsible for storing a possible hit in a memory element and transmit its state to the periphery if the STROBE signal is applied. A peripheral circuit reads the hits and then clears the state memory of the hit pixels. The block diagram of the in-pixel digital logic is shown in Fig. 5.4a. A hit is stored in a latch if the STROBE signal is asserted while the output of the discriminator in the analog front-end (OUTD) is active. Both signals are active low. After the hit is latched, the information is propagated to the periphery with an OR chain shared among all the pixels in a row. In the periphery on the side of the matrix, a logic monitors the state of all the rows and, if a hit occurs in one of these, it asserts a signal called SEL_ROW. Hits on multiple rows are processed sequentially with a position-based priority. The SEL_ROW signal allows the hit information to be propagated in another OR chain shared, this time, among all the pixels in a column. In the periphery at the bottom of the matrix, a logic monitors the state of all the columns and, in combination with the row-steering block, establishes the address of the hit pixels. Once this is done, it asserts the CLEAR_COL signal which resets the state latch of the hit pixels and enables them to record subsequent events. A timing diagram of the in-pixel digital

logic is shown in Fig. 5.4b. The logic also gives the possibility to select a pixel and write a hit in its state latch, by-passing the analog front-end. This operation is referred to as digital pulsing to distinguish it from the analog pulsing, the process of obtaining a hit through the analog front-end. Analog pulsing is also possible and is implemented with a circuitry, not shown in the figure, which capacitively inject a tuneable charge into the collection electrode. Furthermore, the logic gives the possibility to mask individual pixels. The generation of the STROBE signal is triggered externally by the transmission of a command to the chip, which can be sent with a frequency of up to 500 kHz. The logic has been optimized to require the minimum amount of gates and, as the analog section, it has been designed with a conservative layout.



(a)



(b)

Figure 5.4: MOSS in-pixel digital logic: (a) block diagram (b) timing diagram.

5.2.3 Architecture

The architecture of a half unit is shown in Fig. 5.5. It is divided in four regions, one per matrix, each with its own analog biasing block, control logic and readout circuits. A top level

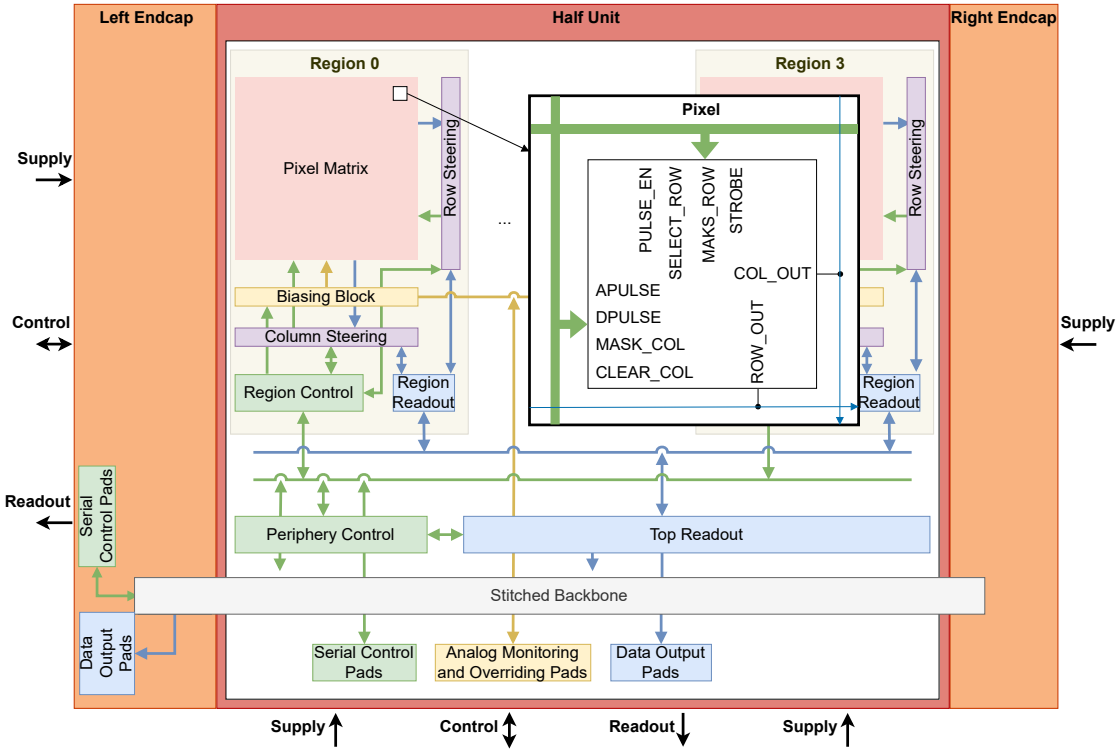


Figure 5.5: Architecture of a half unit of the MOSS chip.

control and readout module communicates with the regional blocks. The flow of the readout data is indicated in blue. The row- and column-steering blocks encode the address of the hit pixels and communicate it to the region readout. The outputs from the four regions are merged in the top readout which is responsible for transmitting them off-chip. The digital readout operates with a 40 MHz clock and transmits the hit data off-chip via an 8-bit port with a custom protocol. The control circuitry, whose flow is reported in green, allows to configure the chip. It also works with a 40 MHz clock and communicates with a serial custom protocol. The periphery control is responsible for configuring the top readout. The control module in the regions is responsible for configuring the row- and column-steering blocks which, in turn, drive the signals in the matrix. It also configures the analog biasing block. The latter provides the biases to the front-end and the in-pixel pulsing circuitry. In order to do so, it integrates a bandgap circuit which generates a reference for 8-bit DACs, in current or voltage, one for each required bias. The biases, whose propagation is reported with the yellow arrows, can be monitored or overridden with dedicated interface pads in case of malfunction of one of the circuits in the biasing block. The half units can optionally be connected to a bus which allows them to be configured and read out, apart from their local interface pads, also from a common control interface and data transmission port in the left endcap. This bus travels over the entire length of the chip and is called stitched backbone. It features an independent power supply and is the only block shared among all the half units on the top or bottom, allowing to study the long-range on-chip communication across the stitching regions. As in the matrix,

also the peripheral circuits are designed for high yield. A custom standard cell library with conservative layouts has indeed been developed for the implementation of the digital blocks in the periphery.

Each half unit has two independent power supplies, one for the analog circuits and another for the digital ones. The entire chip has therefore forty core supplies plus other two for the stitched backbones running over its entire length in the top half units and bottom half units. The power pads are distributed within the half units on the long side of the chip, but also along its short sides in the endcaps. As the repeated sensor units are physically identical, to be able to connect them to different pads in the endcaps, the power lines running parallel to the long side of the chip are made to rotate in each one of them. The rotation allows to use a different power net in every unit. This principle is called line hopping and can be more easily understood with the diagram shown in Fig. 5.6. Only the external sensor units are efficiently powered from the pads in the endcaps. Due to their long distance from the load, the power pads of the other sensor units in the endcaps feature a large series resistance. These pads, however, have been implemented for testing purposes as they allow to monitor the voltage in the middle of the matrices and also verify the stitching.

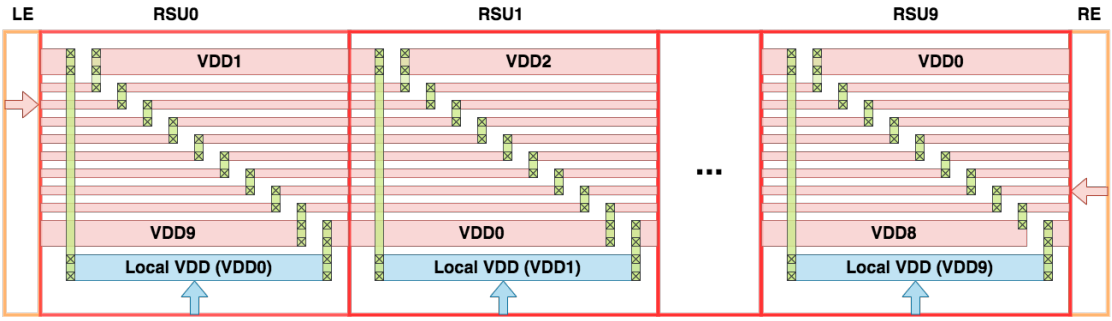


Figure 5.6: Schematic diagram of the line hopping method used in the MOSS chip.

Referring back to the front-end description, the circuit conducts a static current of $\approx 25\,\mathrm{nA}$ in nominal conditions. Taking into account also the pulsing circuit and some decoupling capacitors integrated in the analog section of the pixel, the total static current of the latter is $\approx 30\,\mathrm{nA}$. As the supply voltage is 1.2 V, the static power consumption of the analog section of the pixel is $\approx 36\,\mathrm{nW}$, which results in an analog power density of $\approx 11.11\,\mathrm{mW\,cm^{-2}}$ for the fine-pitch pixels and of $\approx 7.11\,\mathrm{mW\,cm^{-2}}$ for the larger ones. The static power consumption of the digital section is not completely negligible compared to these numbers. Indeed, the leakage current of the digital gates is $\approx 2.53\,\mathrm{nA}$ in the fine-pitch pixels and $\approx 3.58\,\mathrm{nA}$ in the larger ones. The higher leakage current in the large-pitch pixels is given by an enlargement of the transistors which should make them supposedly more reliable. The static digital power density is therefore $\approx 0.94\,\mathrm{mW\,cm^{-2}}$ and $\approx 0.85\,\mathrm{mW\,cm^{-2}}$ over the fine- and large-pitch matrices, respectively. Upon a hit, the current in the front-end dynamically increases. The output of the discriminator, indeed, gets flipped by its input transistor when the current it conducts exceeds the threshold value (I_{DB}) defined by the pull-up current source. Once this happened,

the current in the discriminator branch is limited by the pull-up element to the threshold value itself, which is held throughout the entire ToT duration of the analog waveform. For this reason, the analog dynamic energy associated to a hit depends on the collected charge. Considering the charge released by a MIP and that it is collected most of the times by a single pixel, the average dynamic energy consumed by the front-end for each hit is $\approx 0.5\,\mathrm{pJ}$. For the hit rate of the innermost layer of the ALICE ITS upgrade, i.e. $2.2\,\mathrm{MHz\,cm^{-2}}$, the dynamic power consumption of the front-end is therefore estimated to be $\approx 0.0011\,\mathrm{mW\,cm^{-2}}$. As for the digital readout, the dynamic power consumption is caused by the transmission of the STROBE signal to all the pixels and the propagation of signals to store the hits, transmit them out of the matrix and reset the hit pixels. Given its extracted capacitance, the energy required to transmit a pulse on the global STROBE line in the small-pitch matrix is estimated to be $\approx 0.78\,\mathrm{nJ}$. Considering again the specifications of the ALICE ITS upgrade, the particles' collision rate is 100 kHz. Asserting the STROBE signal with this frequency results in a dynamic power density over the fine-pitch matrix of $\approx 0.235\,\mathrm{mW\,cm^{-2}}$. The energy required to transmit a hit out of the matrix depends on the position of the hit pixel. Assuming a spatially uniform particle flux, i.e. that each pixel has equal probability of being hit, the overall energy to store and transmit a hit to the periphery, reset the pixel and encode its address is on average $\approx 15.4\,\mathrm{pJ}$ in the fine-pitch matrix. In this case, the dynamic power consumption to perform these operations with a hit rate of $2.2\,\mathrm{MHz\,cm^{-2}}$ is $\approx 0.034\,\mathrm{mW\,cm^{-2}}$. The break down of the power density over the matrices is shown in Tab. 5.2. If the data are transmitted to the left endcap, an additional power density of $\approx 0.95\,\mathrm{mW\,cm^{-2}}$ from the stitched backbone has to be added to these values. The largest contributor to the power is by far the static consumption of the front-end. The dynamic power consumption is dominated by the contribution given by the propagation of the STROBE signal which is somewhat lower in the large-pitch matrix due to a lower number of gates connected to this line.

Table 5.2: Power density over the matrices of the MOSS chip.

		Fine pitch	Large pitch
		Power density [$\mathrm{mW\,cm^{-2}}$]	Power density [$\mathrm{mW\,cm^{-2}}$]
Analog	Static	11.11	7.11
	Dynamic (@600 $\mathrm{e^-}$)	0.0011	0.0011
Digital	Static	0.94	0.85
	Dynamic	0.034	0.032
	STROBE (@100 kHz)	0.235	0.165
Total		12.32	8.16

The layout of the wafer-scale MOSS chip is shown in Fig. 5.7. For a better visualization, it is depicted in Fig. 5.8 with one single repeated sensor unit and a zoom on the two types of matrices over an area of 4×4 pixels. The sensor has the same geometry as in the DPTS chip, i.e. it features a 1.14 µm octagonal n-well for the collection electrode, distanced 1.93 µm from the p-well containing the readout circuitry. Furthermore, it implements the process modification with the additional low-dose n- implant cut along the pixel edges. The cut is 2.5 µm wide. As

highlighted in the figure, the pixels are arranged in double rows with the bottom pixels having the analog section on top of the collection electrode and digital part below it, and the top pixels having the parts in the opposite order. This arrangement allows to share the routing lines running along the rows which can thus be routed more conservatively. The overall chip has an area of 14 mm × 259 mm, contains ≈ 6.72 million pixels and ≈ 736.3 million transistors.

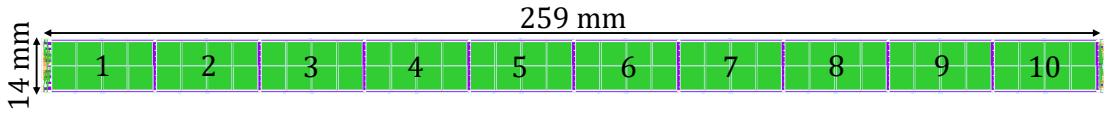


Figure 5.7: Layout of the MOSS chip.

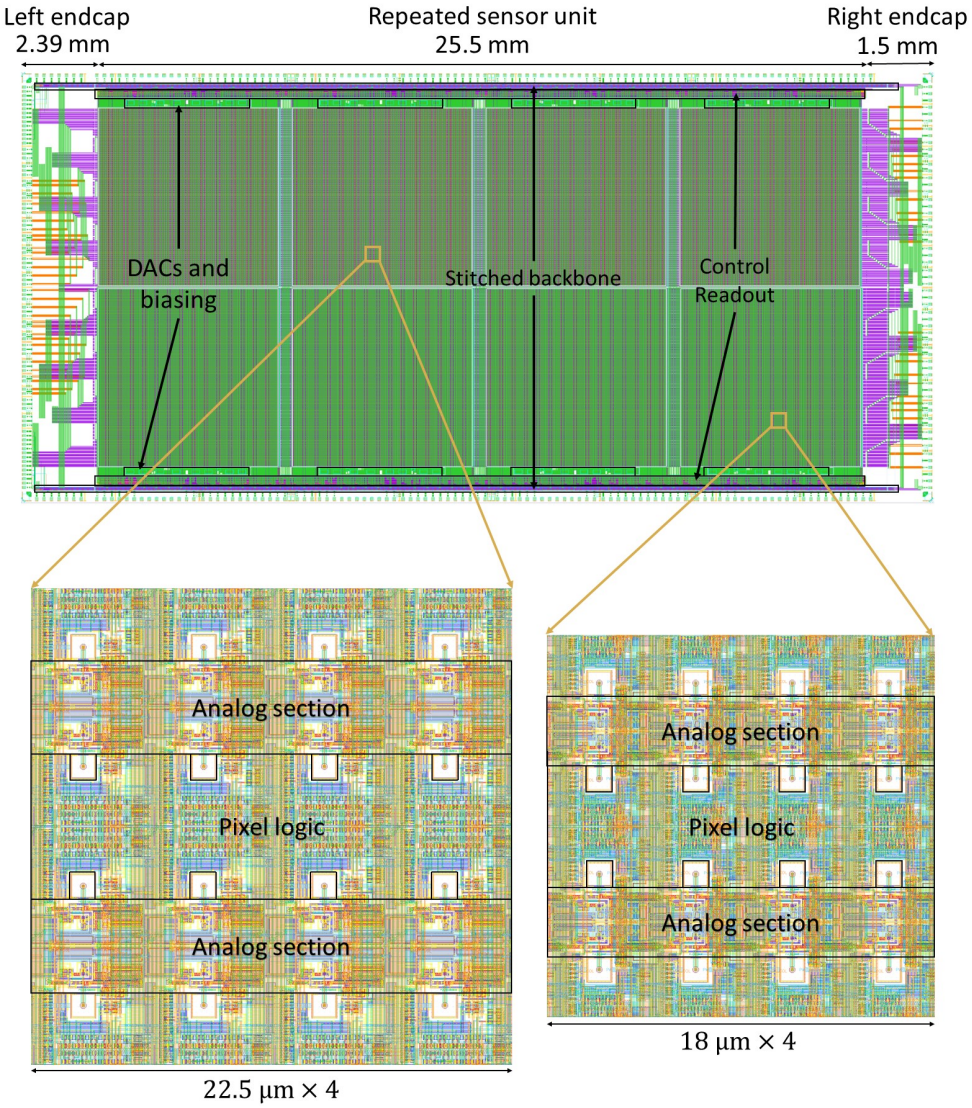


Figure 5.8: Layout of the MOSS chip with one repeated sensor unit.

5.3 MOST - MOnolithic Stitching sensor with Timing

The primary goals of the MOST chip are:

- to investigate the yield of a wafer-scale stitched sensor when in-pixel density is preserved.
- to verify a low-power asynchronous serial readout.
- to explore the performance of long-range (several cm) data transmissions.

As the MOSS, the MOST also exercises the stitching technique in one direction and it is composed of ten repeated identical units completed on the two sides by smaller endcaps. A high-level diagram of the chip is shown in Fig. 5.9. Each repeated unit features 4 matrices of 352 × 64 pixels with a pitch of 18µm. The entire chip has therefore 40 matrices which are connected together by wafer-scale busses. The hits, indeed, are transmitted asynchronously to the endcap at the bottom of the matrices on digital data busses, each of which is shared among pixels in the same columns of the 40 matrices. In order to do so, the in-pixel circuitry uses a shift register controlled by a free-running oscillator started upon detection of a hit. In the bottom endcap, a circuitry appends a column ID to the output of each data bus. A binary tree of OR gates then merges the outputs from all the columns into 4 signals which are transmitted off-chip with differential CML drivers. Apart from the column busses, the readout circuitry in the matrix is independent for groups of pixels. In case of a chip-killing defect, the group where it occurs can be power gated and disconnected from the global readout nets while keeping the rest of the matrix functional. This logic aims to guarantee a high yield even though the in-pixel circuitry is implemented with the density allowed by the standard design rules.

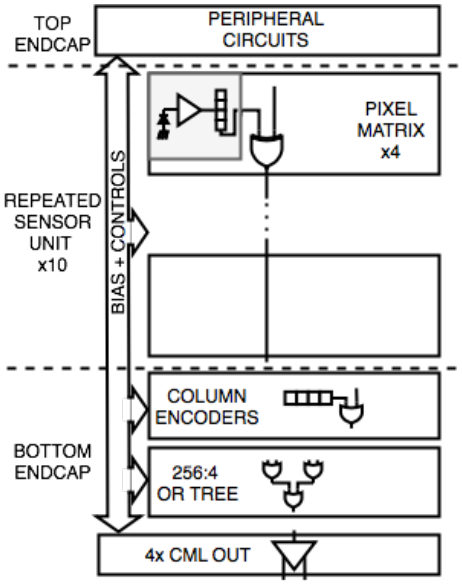


Figure 5.9: High-level diagram of the MOST chip.

5.3.1 Front-end implementation

The analog front-end integrated in the MOST is derived from the circuit developed for the DPTS chip. The adopted modifications allow increasing the input voltage on the electrode, as in the MOSS. The schematic of this circuit is shown in Fig. 5.10. In the MOST, the p-well of the in-pixel circuitry and the p-type substrate are connected to the ground of the analog supply rather than to dedicated nets. Therefore in this chip, differently to the MOSS, back bias, detrimental for yield, cannot be applied. As a consequence, the sensor reverse bias can only be enhanced by increasing the voltage on the collection electrodes thanks to the implemented modifications. As shown in the figure, these entail the addition of a PMOS transistor (M5) with the gate connected to a tuneable voltage bias, named V_S, between the amplification stage (reported in the dashed box) and its ground. In this circuit, all the voltages in the amplifier can be shifted up through the V_S bias. The analog supply also can be increased and the transistor M5 can be configured to ensure that the core devices of the amplifier are not exposed to voltages beyond their maximum ratings. This circuit gives the possibility to obtain larger input voltages compared to the implementation in the MOSS at the expense of a larger power consumption. It is worth to highlight once again that this is the only way to enhance the sensor reverse bias. The discriminator still implements a current comparator. The current forced by the transistors M11 and M12 is defined by their V_{GS} and set through the amplifier output baseline and the bias voltage V_S to be lower than the $I_{DB}/2$ current forced instead by the transistor M13. In DC, therefore, the OUTD node is close to ground and only a

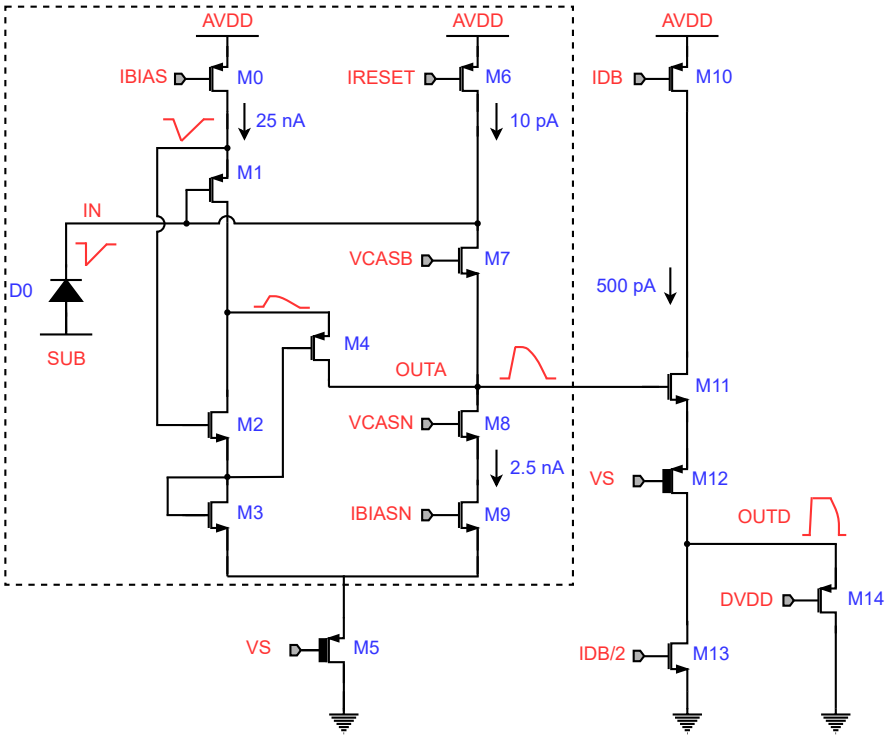


Figure 5.10: Schematic of the front-end in the MOST chip.

small current flows in the discriminator branch. As the signal on OUTA rises upon a particle hit, the current drawn by the transistors M11 and M12 increases, eventually exceeding $I_{DB}/2$ and charging the output node. The transistor M14 is used to clip the OUTD signal to the digital supply DVDD and ensure that it does not go beyond the maximum voltage sustainable by the following digital circuitry. The transistor M10 is used instead to limit the dynamic current in the discriminator to twice the threshold value, i.e. I_{DB}, and avoid high current spikes. Large input voltages can only be obtained by setting a high supply level, which mainly drops across the transistors M5 and M12. These two devices are thus designed with thick gate oxides. With a supply of 3 V, the input voltage can go up to ≈ 2.6 V. The main performance metrics of the front-end obtained through parasitic-extracted simulations with these voltages and the currents noted on the schematic are reported in Tab. 5.3. For these simulations the sensor is modelled with a capacitance of 1 fF. The larger threshold dispersion compared to the one simulated in the DPTS is due to the additional devices.

Table 5.3: Simulated performance metrics of the front-end in the MOST chip.

Gain	Charge Threshold	Peaking Time	Phase Margin	Thr. Disp.	ENC
$\approx 1\,\mathrm{mV/e^-}$	$\approx 150\,\mathrm{e^-}$	$\approx 1\,\mu\mathrm{s}$	$\approx 55°$	$\approx 15\,\mathrm{e^-}$	$\approx 10.5\,\mathrm{e^-}$

5.3.2 Digital readout

The MOST chip features an asynchronous event-driven serial readout. Upon a particle hit, the in-pixel digital logic instantaneously sends a stream of pulses corresponding to the address of the hit pixel to the periphery at the bottom of the matrices on a digital data bus. This logic is partly independent for each pixel and partly shared among groups of four adjacent pixels in the same column. The logic confined to each pixel is shown in Fig. 5.11. The goal of this circuit is to generate a digital pulse, shorter than the one produced by the front-end discriminator, whenever a hit occurs. The discriminator is thus connected to a logic which generates a pulse with a width equal to dT upon each leading edge of its output. In order to do so, the pulse generator uses a delay element which inverts and shifts the discriminator output by a duration of dT and provides its output with the original signal to a NOR gate

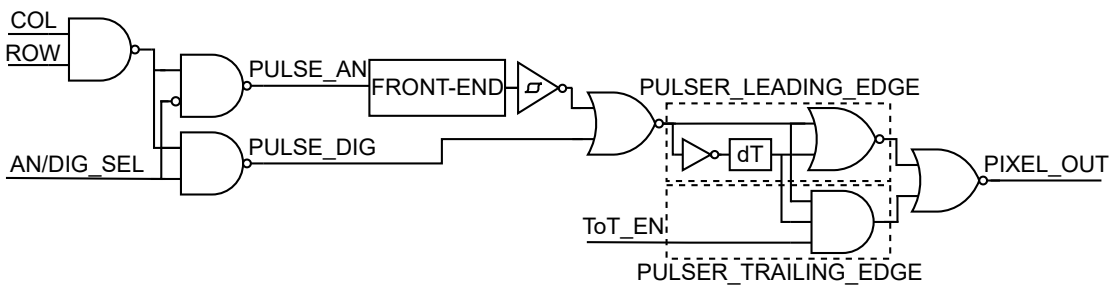
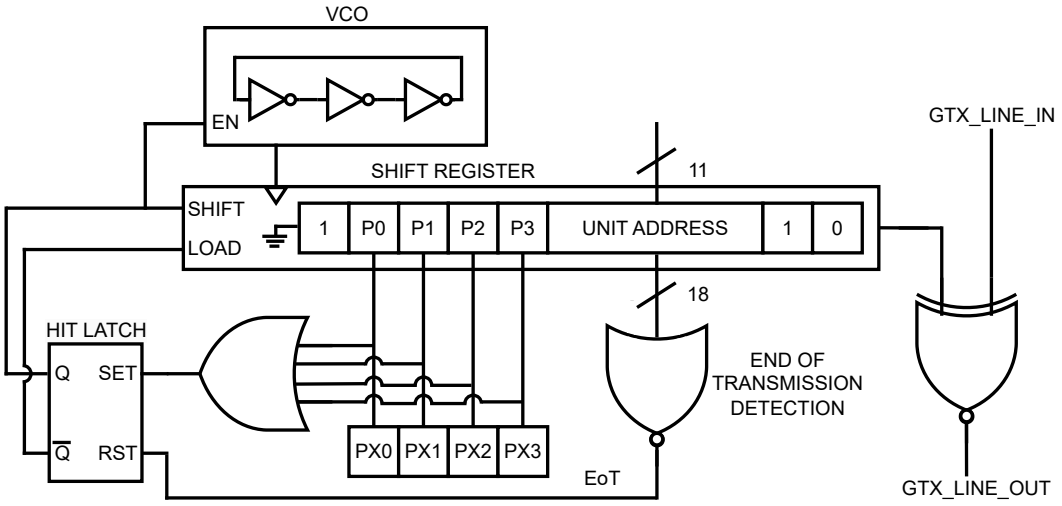


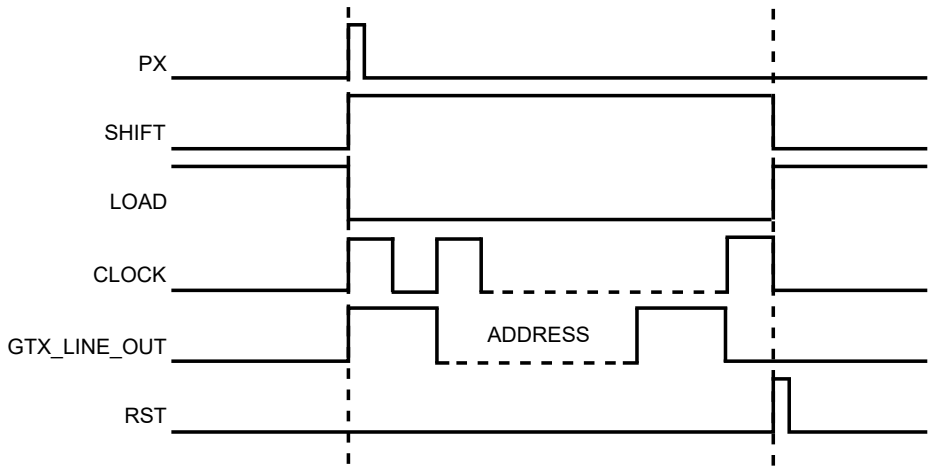
Figure 5.11: MOST in-pixel digital readout circuit.

which generates said pulse. These signals are also provided to a tri-state AND gate which can optionally be activated to obtain the same pulse on the trailing edge of the discriminator output. In this way, the following logic would react to both edges of the discriminator output enabling an off-chip reconstruction of the ToT information. As in the MOSS, the in-pixel logic gives the possibility to perform analog and digital pulsing of single pixels. These operations are implemented by the NAND gates placed before the front-end block. The analog pulsing trigger signal goes through an additional circuit in the analog domain which allows tuning the capacitively injected charge.

The in-pixel logic shared among groups of four pixels is shown in Fig. 5.12a. This is the circuitry actually responsible for transmitting the address of the hit pixel on a digital data bus. The previous logic, represented by the boxes labelled as PX, is connected to a 4-input OR gate. The latter combines the outputs of the four pixels into a single line and therefore generates a pulse whenever a hit occurs in one of these. A latch following the OR gate stores the hit. Once this happens, a shift register switches from working in load mode to shift mode. In load mode, the bits identifying the address of the pixels are loaded into the register. In shift mode, the content of the register is serially injected into a global transmission line through an XNOR gate. In order to do so, the latch also activates a ring oscillator which generates the clock required by the shift register to operate. The ring oscillator is composed of current-starved inverters whose bias can be externally adjusted to vary the oscillation period and so the bit time of the output sequence, which is 1 ns in nominal conditions. Once the transmission is over, a NOR gate detects that the shift register is empty and flips its output, indicating the end of the transmission. This signal is used to reset the state latch and thus predispose the group logic to acquire subsequent hits. A timing diagram of the group logic is shown in Fig. 5.12b. An additional gate, not shown in the figure, allows to select a group and disable its digital logic in case it generate an excessive noise rate by preventing the state latch from being set.



(a)



(b)

Figure 5.12: Digital readout circuit of a group of pixels in the MOST matrix: (a) logic and (b) timing diagram.

The groups of pixels make up the active area of the entire chip according to the following rules:

- 4 consecutive groups in a column form a unit and transmit their outputs on 4 separate global lines.
- groups that are unit(s) apart transmit their outputs on the same global line.
- the stacking of 22 units form a sub-column.
- 64 parallel sub-columns form a sub-matrix.
- the stacking of 40 sub-matrices, 4 in each repeated sensor unit, form the entire active area of the chip.

The readout hierarchy up to the sub-matrix level is shown in Fig. 5.13. The interleaved connections to the global lines are implemented to avoid collision of transmissions on these lines in case of charge sharing, as the access to them is fully asynchronous and not arbitrated. The group logic transmits the address of the hit pixel in the group, its unit, sub-matrix and repeated sensor unit. The coordinate of the group in the unit and of the column are resolved positionally and attached in the bottom endcap.

5.3.3 Power gating

To guard from possible chip-killing defects, the groups of pixels are not supplied directly by the main power but instead connected to it through switches. In case of a harmful defect, sets of groups can be disconnected from the power by turning off their power switches via dedicated control signals. The control for the analog power switches is different from the one

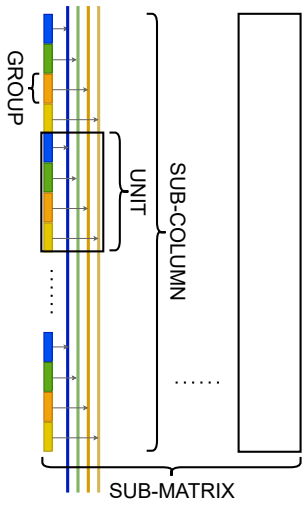


Figure 5.13: Hierarchy of the readout in the MOST chip.

of the digital power switches. In the case of the analog circuitry, the granularity of the power gating is a row of groups, i.e. a band of 4 pixels. Regarding the digital circuitry, two control signals are distributed per sub-column and these control the power of the even and odd pixel groups. The granularity of the digital power gating is therefore of half a sub-column. A sketch of the power gating strategy is shown in Fig. 5.14.

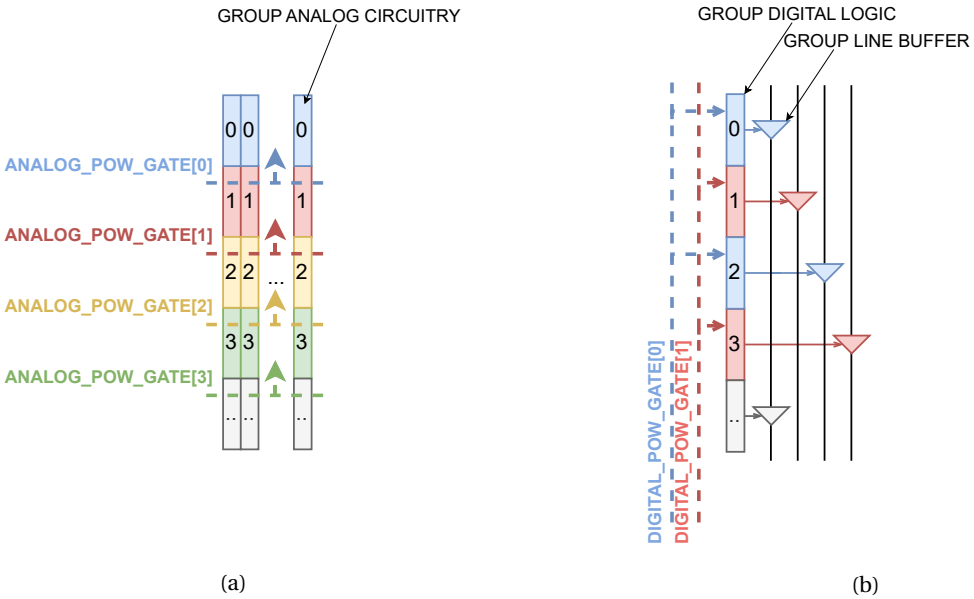


Figure 5.14: Strategy of the power gating for the (a) analog and (b) digital circuitry of the pixel groups in the MOST matrix.

5.3.4 Architecture

The overall architecture of the MOST chip is shown in Fig. 5.15. All the figures below refer to the chip rotated by 90° for space reasons. The chip can be configured through the control modules. These are disposed parallel to the long edge of the chip. To drive the configuration signals running along the columns of the matrices, as the controls of the digital power switches, a control module is interposed between two pairs of matrices. The control modules operate with a 40 MHz clock and communicate with the same serial custom protocol as in the MOSS chip. Furthermore, in all the repeated sensor units, they can be accessed through a common control interface in the top endcap thanks to a bus, the stitched backbone, which crosses the stitching regions and travels over the entire length of the chip. The control interface in the top endcap is the only access point to configure the chip.

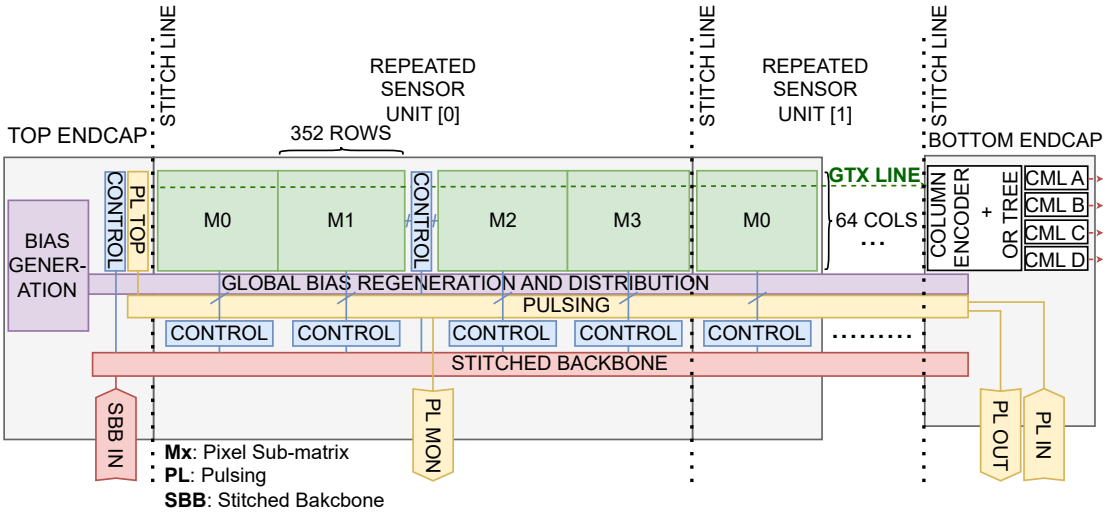


Figure 5.15: Architecture of the MOST chip.

The analog biasing circuit of the pixel matrix is composed of a logic distributed over the entire chip. A high-level scheme of this circuit is shown in Fig. 5.16. For each required bias, an external reference is provided to the chip through a dedicated pad in the top endcap. In this region, a circuit generates from the external reference a pair of voltages V_{GN} and V_{GP} which represent it. These signals are routed over the entire length of the chip. For each sub-matrix, a circuit, starting from these voltages, locally restores the reference quantity. The logic used to generate the distributed voltages from the reference value and vice versa prevents, to a large extent, the propagated quantity from being affected by the voltage drops on the power and ground lines. Furthermore, independently for each bias and sub-matrix, a local adjustment of up to ± 15 % can be applied to the propagated quantity to compensate for potentially large variations. For each row of pixel groups, a block scales the local quantity to the magnitude required by the pixel circuits and distributes the bias to the matrix. In case a row of pixel groups has to be power gated due to a chip-killing defect, its biasing block disconnects the lines distributed to the pixels from the global circuitry. This avoids that also other sections of the matrix are affected by the same fault.

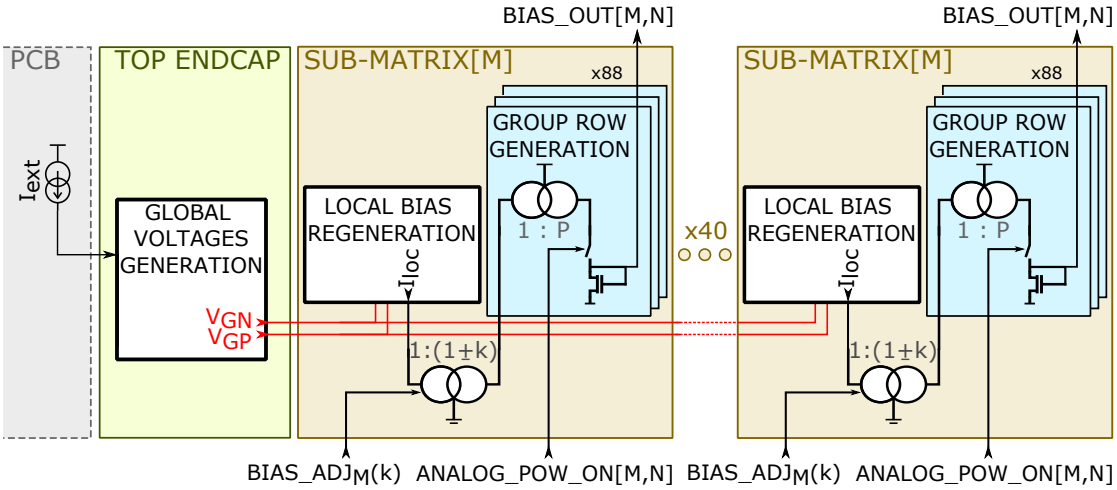


Figure 5.16: Biasing scheme of the MOST chip.

The pulsing operations in the pixels are triggered externally by sending a signal to the chip. The interface pad dedicated to this signal is placed in the bottom endcap. It is then propagated to the row of the pixel to be pulsed through a buffer chain. This logic is labelled as "pulsing" in the chip architecture in Fig. 5.15 and, as the stitched backbone and biasing block, is distributed over the entire length of the chip. Even though its basic purpose is to propagate the trigger of the in-pixel pulsing circuitry, it is implemented also to allow a timing characterization of the long-range signal transmissions across the stitching regions. A diagram of this logic is shown in Fig. 5.17. Each unit of the buffer chain is placed in correspondence of a row of pixel groups. To be able to track the delay of the buffer chain, a pad in the middle of each repeated sensor unit allows to monitor some of its internal nodes. Once it reaches the top endcap, the buffer chain folds back and continues all the way down to the bottom endcap where it transmits the propagated pulse off-chip via an interface pad. The signal can be injected into each row to trigger a pulsing operation in the pixels where a dedicated control line distributed along the columns is asserted. This can be done with the trigger pulse travelling towards the top endcap or in the opposite direction. In the top endcap, the propagated pulse can also be used to trigger a block which injects a configurable sequence of bits into the global transmission lines of the matrix. This block allows to test transmissions of arbitrary signals over the entire length of the chip. The logic is designed to enable the characterization of the delay and jitter of every component of the buffer and readout chain. As the hits are transmitted asynchronously to the periphery and out of the chip, a timing characterization of the readout chain would allow to perform an off-line reconstruction of the arrival time of the particles on the pixels.

The circuits in the chip periphery are designed with robust layouts. The same custom standard cell library with conservative layouts developed for the MOSS has been used also in this chip for the implementation of the digital peripheral blocks. The chip features four power domains: one for the digital blocks, one for the CML buffers in the bottom endcap and two for the analog circuits. Indeed, the analog power supply in the matrix can be increased to accommodate

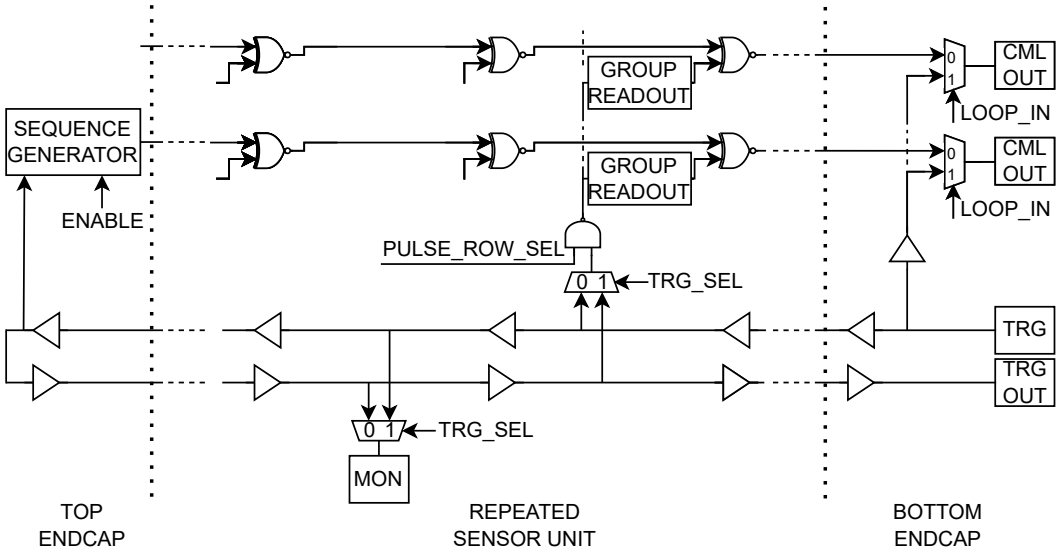


Figure 5.17: Diagram of the pulsing logic in the MOST chip.

large input voltages and hence sensor reverse biases. On the other hand, part of the analog periphery works with a supply constantly fixed at 1.2 V. The pads which provide the matrices' power are distributed regularly in the chip, both on the long sides and in the external endcaps. The most power-hungry component in the matrix is the analog front-end. The total current consumption of the analog section of the pixel is $\approx 28\,\mathrm{nA}$. With a power supply of 3 V, high enough to obtain a sufficiently large sensor reverse bias, its static power consumption is $\approx 84\,\mathrm{nW}$, which results in an analog power density of $\approx 26\,\mathrm{mW\,cm^{-2}}$. As for the other front-end implementations, the current in the discriminator increases during a hit to the threshold value which is held throughout the entire ToT duration. For the charge released by a MIP, the dynamic energy required by the front-end for a hit is $\approx 1.35\,\mathrm{pJ}$ and its dynamic power consumption is $\approx 0.003\,\mathrm{mW\,cm^{-2}}$ for the hit rate of $2.2\,\mathrm{MHz\,cm^{-2}}$ expected on the innermost layer of the ALICE ITS upgrade. For the same hit rate and a uniform particle flux over the matrix, the dynamic power consumption of the digital logic is $\approx 1.03\,\mathrm{mW\,cm^{-2}}$ and is mainly required for the transmission of the hits to the periphery. This operation indeed requires a dynamic energy of $\approx 440\,\mathrm{pJ}$ whereas the in-pixel processing only $\approx 26.2\,\mathrm{pJ}$. Due to an average leakage current of $\approx 7\,\mathrm{nA}$ per pixel, the static digital power consumption is comparable to the dynamic one and is $\approx 2.6\,\mathrm{mW\,cm^{-2}}$. A summary of these values is reported in Tab. 5.4.

Table 5.4: Power density over the matrices of the MOST chip.

		Power density [$\mathrm{mW\,cm^{-2}}$]
Analog	Static	26
	Dynamic (@600 $\mathrm{e^-}$)	0.003
Digital	Static	2.6
	Dynamic	1.03
Total		29.63

The layout of the MOST chip with one single repeated sensor unit and a zoom on a matrix is shown in Fig. 5.18. The sensor features the low-dose n- implant cut along the pixel edges and has the same geometry as in the MOSS. As highlighted in the figure, the analog section of the pixel is flipped in every other row in a way that pairs of pixels in the same column can share routing lines. The overall chip has an area of 2.5 mm × 259 mm, contains ≈ 0.9 million pixels and ≈ 160.2 million transistors.

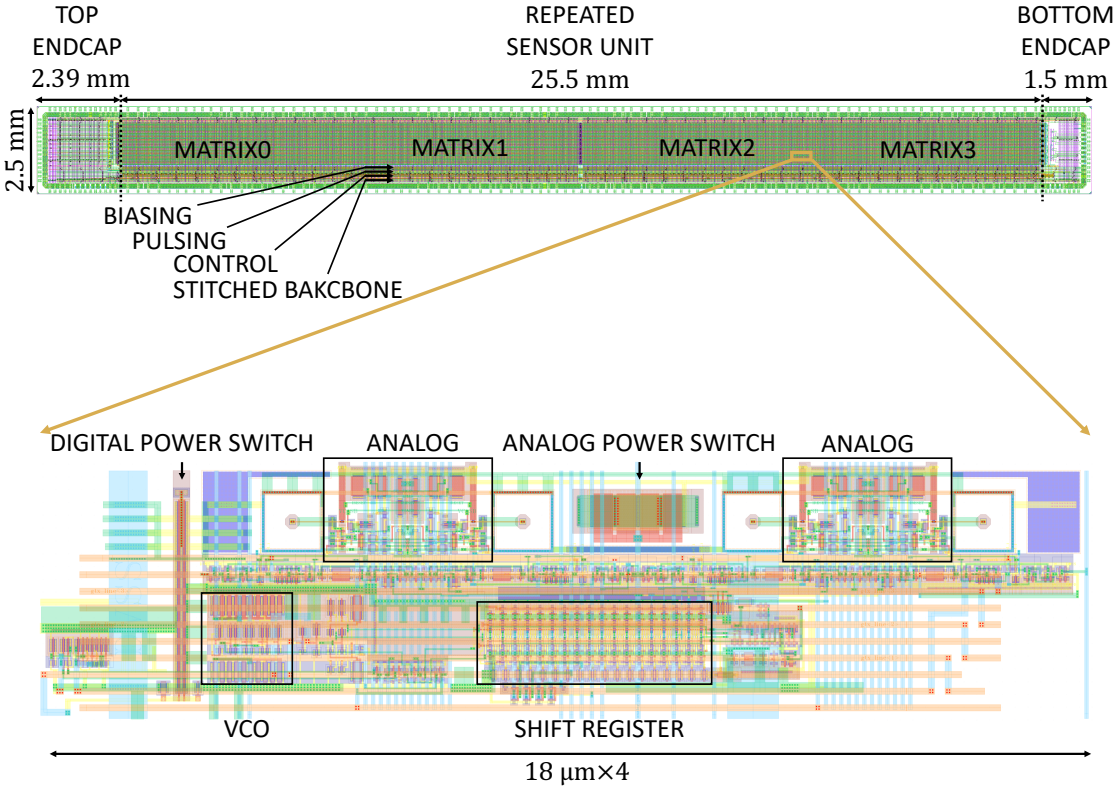


Figure 5.18: Layout of the MOST chip with one repeated sensor unit.

5.4 Summary

In this chapter, the main design aspects of two stitched sensors in the TPSCo 65 nm imaging technology, the MOSS and MOST, have been presented. The primary goal of these prototypes is to prove the feasibility of realizing wafer-scale monolithic stitched sensors.

The MOSS measures 14 mm × 259 mm and implements a synchronous strobed readout. The chip is composed of 10 repeated sensor units, in turn divided into a top and bottom half unit, capped on the two sides by a left and right endcap. Each of these 20 regions is a standalone powering and functional unit including its own periphery and I/O pads. This segmentation allows to turn off independently each one of the half units, representing 1/20 of the chip, in case it has a chip-killing defect, while keeping the others functional. Both the top and bottom units contain 4 matrices of equal size. However, the matrices in the top units feature

256 × 256 pixels with a pitch of 22.5 µm whereas in the bottom ones 320 × 320 pixels with a pitch of 18 µm. Implementing the same logic, the top and bottom matrices are characterized by different circuit densities. It is thus possible to investigate a possible dependence of the occurrence of manufacturing defects on the circuit densities in the matrices. The half units can be configured and read out through their local interface pads. Optionally, this can be done through a common interface and data transmission port in the left endcap thanks to a bus travelling over entire length of the chip called stitched backbone. The latter features an independent power supply and is the only block shared among all the top or bottom half units. All the circuits in the chip are designed with a robust layout, i.e. with widths and distances larger than the minimum values allowed by the design rules. The front-end developed for the DPTS has been used also for this structure. In this circuit, a modification allowing to increase the sensor reverse bias through its settings has been implemented. Reducing the substrate voltage to increase the sensor reverse bias is detrimental for yield. It is thus planned to be kept close to ground, at least in the initial testing phase of the prototype. The modification allows to partially overcome this limitation.

The MOST measures 2.5 mm × 259 mm and implements an asynchronous event-driven readout. As the MOSS, it is composed of 10 repeated sensor units completed on the two sides by two endcaps, referred to, this time, as top and bottom endcap. Each repeated unit features 4 matrices of 352 × 64 pixels with a pitch of 18 µm. The hits are transmitted asynchronously to the bottom endcap on digital data busses, each of which is shared among pixels in the same column of the different matrices. The digital readout logic in the matrices is divided into independent identical blocks distributed over groups of four adjacent pixels in the same column. To transmit the hits to the bottom endcaps, the group logic uses a shift register controlled by a ring oscillator started upon detection of a hit in one of its four pixels. In the bottom endcap, a circuitry appends an ID to the outputs of each data bus and merges them into 4 signals transmitted off-chip with differential CML drivers. The MOST is a single powering and functional module. Part of the biasing and pulsing circuitry is indeed shared among all the repeated sensor units. A bus running along the entire length of the chip called, consistently to the MOSS, stitched backbone allows to communicate with control modules in each repeated sensor unit through an interface port in the top endcap. This port is the only access point to configure the chip. The peripheral blocks have conservative layouts. In the matrix, instead, the normal circuit density allowed by the standard design rules has been preserved. The circuitry of the pixel groups is connected to the power grid through switches, one for the analog section and another for the digital logic. In case of fault, the circuitry of sets of groups can be power gated by turning off their switch. For the analog circuitry, the granularity of the power gating is a row of groups, i.e. a band of 4 out of 14080 rows of pixels. The digital circuitry can instead be power gated with a granularity of 1/2 out of 64 columns. Similarly to the MOSS, the front-end used in this structure is based on the topology developed for the DPTS, modified to allow increasing the sensor reverse bias through its settings. The modifications adopted in this case enable sensor reverse biases even beyond the values allowed by the MOSS implementation at the expense of a larger power consumption.

6 Conclusions and outlook

Integrating the sensitive layer and readout in the same silicon die, monolithic sensors present several advantages over the more largely used hybrid sensors such as easier detector assembly, lower cost, lower power consumption and lower material budget. This work was carried out in the EP-R&D program at CERN and focused on the development of readout solutions for monolithic active pixel sensors targeting future HEP experiments. In these, to cope with a higher luminosity and improve the detector accuracy, the pixel sensors are required to feature higher granularities, dissipate lower power and withstand larger levels of irradiations.

The optimization of the analog front-end circuit integrated in the MALTA2 chip, a monolithic sensor of 10.12 mm × 20.2 mm featuring a matrix of 224 × 512 pixels with a pitch of 36.4 µm targeting the requirements of the outer layer of the ATLAS ITk upgrade, was shown. The sensor is manufacutured in the TowerJazz 180 nm imaging process. The pixels implement a small collection electrode which therefore offers a small capacitance (<5 fF), beneficial for high analog performance. Process modifications were implemented to accelerate the charge collection and reduce charge sharing among pixels. These improve also the NIEL tolerance of the sensor while preserving its small capacitance. The front-end is a continuously active open-loop amplifier followed by a high-gain common-source discriminator stage. MALTA2 samples were extensively characterized to evaluate the front-end performance before and after irradiation and the main characterization results were shown. The circuit has, with a power consumption $\lesssim 1\,\mu W$, an ENC and a threshold dispersion respectively of $6.5\,e^-$ and $6\,e^-$. Thanks to these values, charge thresholds of $100\,e^-$ can be achieved and maintained even after irradiation with neutrons up to $3{\cdot}10^{15}\,1\,\text{MeV}\,n_{eq}\,cm^{-2}$ and 3 Mrad or with X-rays up to 100 Mrad despite a somewhat larger ENC and threshold dispersion. At these thresholds, measurements with test beams show a detection efficiency of the sensor $\gtrsim 99\,\%$ before irradiation. The efficiency drops to $\approx 35.6\,\%$ for samples irradiated at $3 \cdot 10^{15}\,1\,\text{MeV}\,n_{eq}\,cm^{-2}$ and 3 Mrad and measured at a temperature of $-20\,°C$.

The possibility of developing monolithic sensors in a smaller node technology, as the TPSCo 65 nm imaging process, was explored, also in view of the ALICE ITS upgrade. Several test structures were submitted to validate this technology for HEP applications. The APTS was

developed to characterize the analog behaviour of the sensor. It contains a small matrix of 4×4 pixels, each of them connected to an output pad through a buffer chain to enable the simultaneous off-chip visualization of signals from all the electrodes. Similar process modifications to those implemented in the TowerJazz 180 nm imaging technology were applied also in this case. Measurements on this chip proved that the sensor can reach a capacitance, including the input contribution of the readout circuitry, in the order of $\approx 2\,\mathrm{fF}$ and sub-nanosecond collection times. Furthermore, they validated the effectiveness of the process modifications. The DPTS was developed to evaluate the efficiency of the sensor over a larger sensitive area as it includes a matrix of 32×32 pixels with a pitch of 15 µm. It also implements a novel readout circuitry. The analog front-end integrated in this structure is an evolution of the topology developed in the previous technology to obtain a more compact solution. The digital readout is an asynchronous event-driven logic which time-encodes the events' information such as address of the hit pixels and ToT. Electrical tests on the chip showed that the front-end has, with a power consumption of 12 nW, an ENC and a threshold dispersion respectively of $\approx 15.4\,\mathrm{e}^-$ and $\approx 15.8\,\mathrm{e}^-$. These values improve with a larger power consumption. Measurements with test beams showed on samples irradiated up to $10^{15}\,1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$ a detection efficiency $\gtrsim 99\,\%$ for charge thresholds $\lesssim 100\,\mathrm{e}^-$ at room temperature, which is an unprecedented result. These measurements allowed the validation of the full readout chain, including the novel front-end, the asynchronous digital readout and the sensor design optimized for faster charge collection.

The stitching technique allows to fabricate chips as large as the entire wafer. Large scale sensors facilitate the coverage of large areas and, if able to cover the entire sensitive area of a detector, eliminate the need of tiling several sensors reducing the material budget. This is targeted in the ALICE ITS upgrade. To prove the feasibility of the stitching technique for HEP applications, two wafer-scale sensors were designed in the TPSCo 65 nm imaging technology, the MOSS and MOST. The former features a size of 14 mm × 259 mm and a strobed readout. The latter, instead, features a size of 2.5 mm × 259 mm and an asynchronous event-driven readout. A primary concern in designing such large systems is to obtain an acceptable yield and prevent that a single manufacturing defect, likely to occur due to the very large area, jeopardizes the entire chip. In the MOSS, to cope with this issue, the circuits are designed with conservative layouts. Furthermore, the system is composed of 20 independent units. In case of a chip-killing defect, each one of these can be turned off while keeping the rest of the chip operational. The MOST is instead a single functional module. Its peripheral blocks are also designed with a robust layout. In the matrix, instead, the circuits are implemented with normal design rules to check the occurrence of manufacturing defects in this case. If faults occur, the pixels can be power gated. For the analog circuitry, the granularity of the power gating is a band of 4 out of 14080 rows, whereas for the digital one it is 1/2 out of 64 columns. The topology of the front-end developed for the DPTS was used as a solution for these chips. Modifications were implemented in this circuit to introduce the possibility of increasing the reverse bias of the sensor through the front-end settings.

As the stitching technique has never been attempted before on such complex systems in HEP, the successful validation of these chips would represent a crucial milestone for the developments of future monolithic sensors. Nevertheless, significant progress can still be achieved. Both in the MOSS and MOST, the power pads are distributed regularly along all the sides of the chips. For a seamless integration into an experiment, the stitched sensors should be serviced only from one of their short sides. For the ALICE ITS upgrade, it is planned to confine all the connections to the external world out of the detector acceptance, where more material can be added without degrading the physics performance of the experiment. With the present power consumption and routing resources provided by the used technology, supplying the sensors only from one of their short sides would lead to large power drops, up to hundreds of mV on the opposite side of the chip. Such large drops would hamper the proper functionality of the circuitry. To avoid this, innovative schemes for the supply distribution in which its voltage drop is regulated along the structure could be implemented. However, reducing the power consumption would be in general beneficial for several system aspects. In the stitched developments of this work, with a current of $\approx 25\,\mathrm{nA}$ in the amplification stage and $\lesssim 1\,\mathrm{nA}$ in the discriminator, the analog front-end dominates the power consumption of the matrices. Its power consumption has to be therefore reduced, potentially by exploring alternative solutions. As highlighted in the thesis, this can also be achieved with a higher Q/C ratio or for a given charge, as it is dictated by the sensor process, a smaller sensor capacitance. In fact, a larger voltage at the input of the circuitry can be used to reduce its DC biasing current while obtaining the same, or an even larger, dynamic current which provides gain. With a sufficiently large input signal, a simple inverter can be used to amplify and digitize the charge released by the traversal of a particle. This circuit provides the ultimate power consumption as it dissipates in DC only a small leakage current. Even if the sensor capacitance becomes negligible, the input capacitance of the circuit itself would limit the Q/C ratio. In this case, a $2\times$ minimum-size inverter requires $\approx 5000\,\mathrm{e}^-$ to flip in a 65 nm technology. This value reduces to $\approx 800\,\mathrm{e}^-$ in a 28 nm technology and to $\approx 100\,\mathrm{e}^-$ in a 5 nm technology [106]. Although other topologies can be considered, this illustrates the benefit of finer linewidth technologies to improve the power-to-performance ratio of the analog front-end circuit.

In conclusion, integrated circuits and sensors have revolutionized the way HEP experiments are built, and this trend is set to continue. Furthermore, the latest breakthroughs in image sensors for commercial use, such as stitching, have the potential to enable the development of larger area detectors, but they also pose new challenges that must be addressed through further innovation.

Chip gallery

Figure 1: MALTA2 chip.

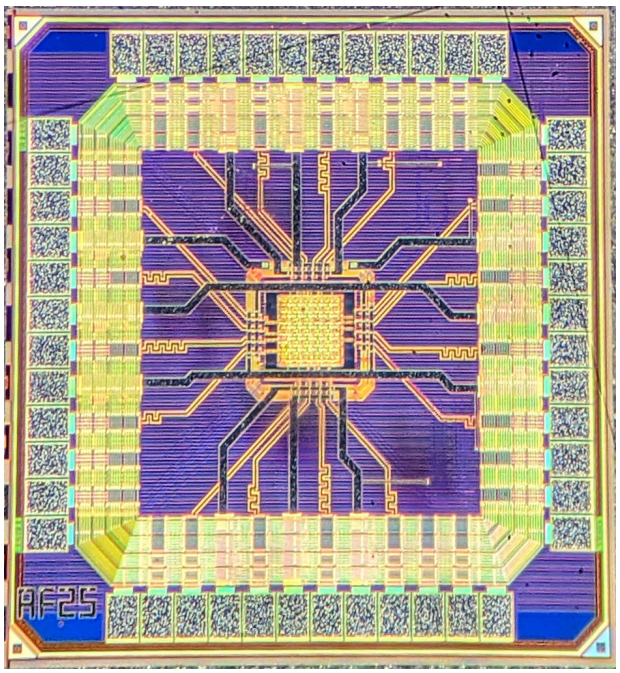


Figure 2: APTS chip.

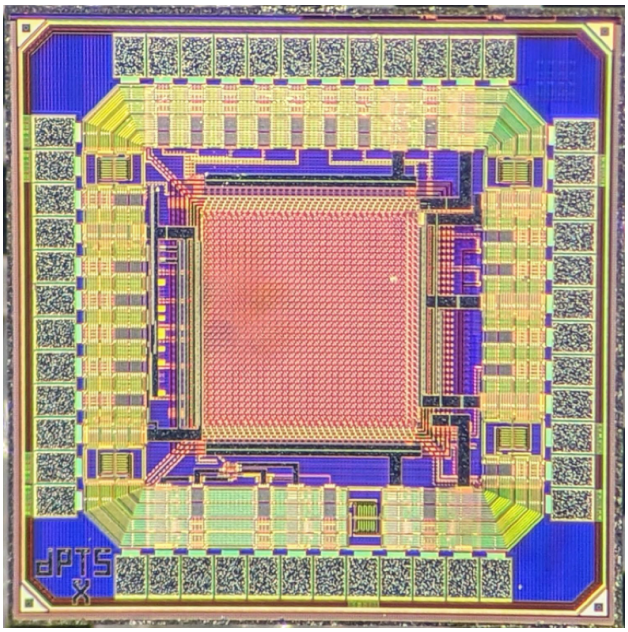


Figure 3: DPTS chip.

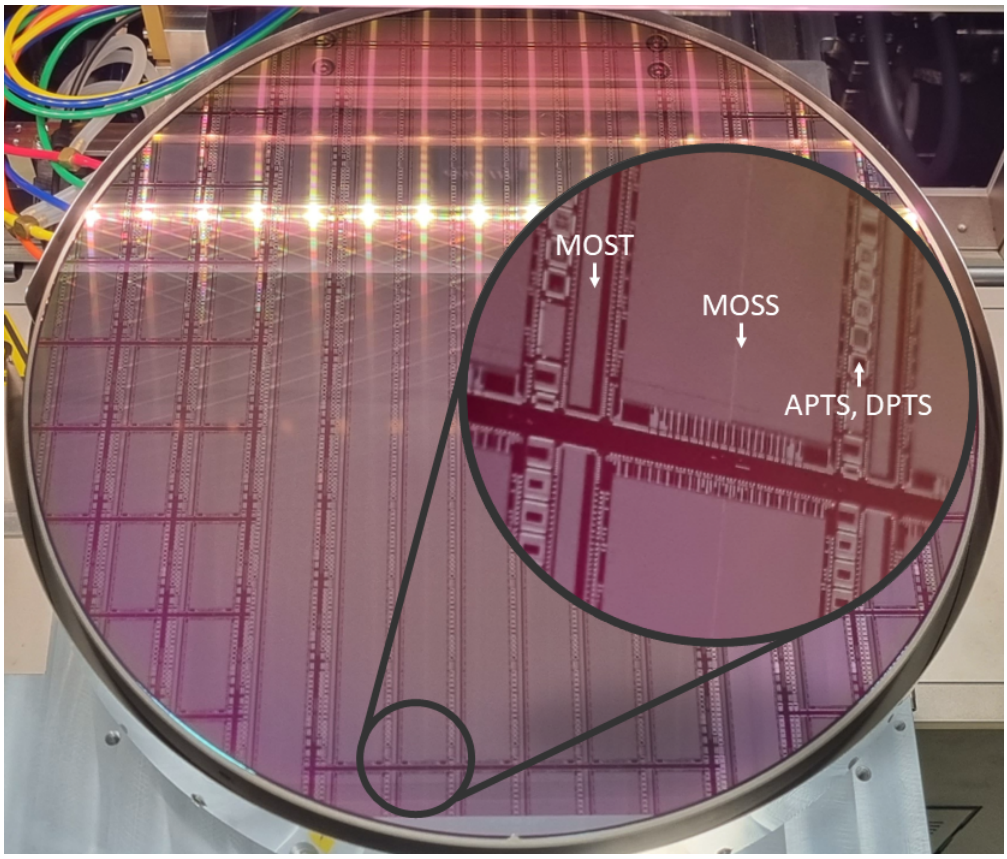


Figure 4: Wafer with the MOSS and MOST stitched sensors and other small-scale chips (including the APTS and DPTS prototypes).

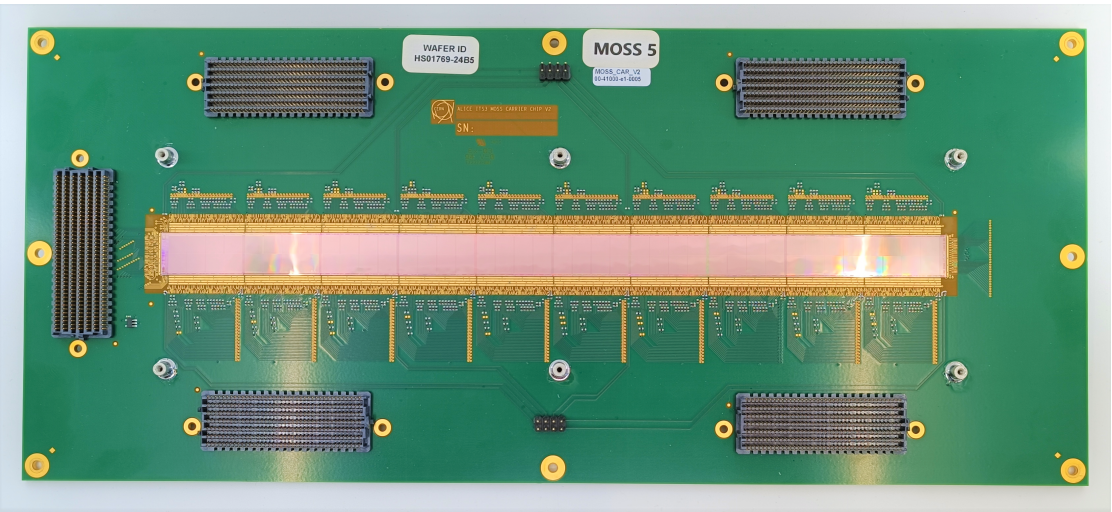


Figure 5: MOSS chip on a carrier board.

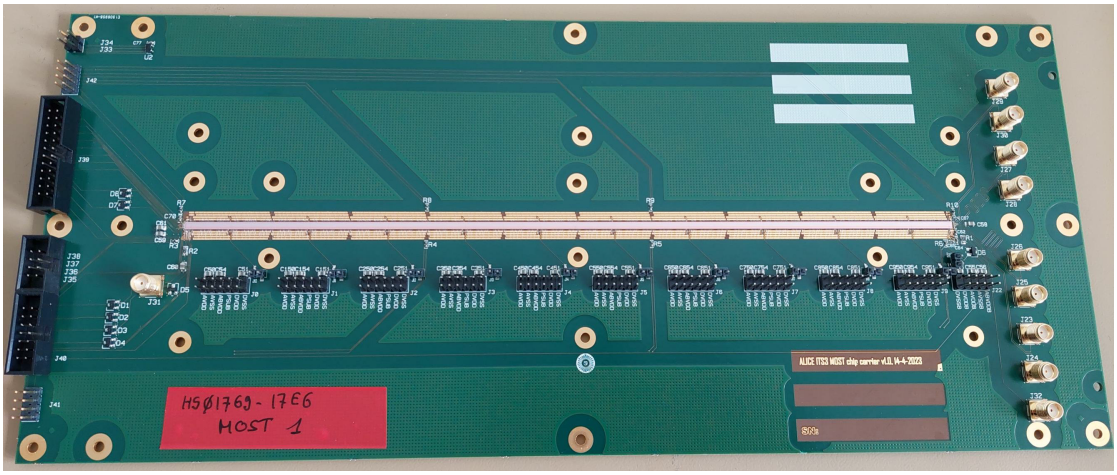


Figure 6: MOST chip on a carrier board.

Bibliography

[1] C. Wilson, “On the Cloud Method of Making Visible Ions and the Tracks of Ionizing Particles”, *Nobel Lecture*, [Online]. Available: https://www.nobelprize.org/prizes/physics/1927/wilson/lecture/.

[2] D. A. Glaser and D. C. Rahm, “Characteristics of Bubble Chambers”, *Phys. Rev.*, vol. 97, pp. 474–479, 2 1955. DOI: 10.1103/PhysRev.97.474.

[3] C. D. Anderson, “The Apparent Existence of Easily Deflectable Positives”, *Science*, vol. 76, no. 1967, pp. 238–239, 1932. DOI: 10.1126/science.76.1967.238.

[4] C. M. G. Lattes et al., “Processes involving charged mesons”, *Nature*, vol. 159, pp. 694–697, 1947. DOI: https://doi.org/10.1038/159694a0.

[5] G. Charpak and F. Sauli, “High-accuracy, two-dimensional read-out in multiwire proportional chambers”, *Nucl. Instrum. Methods Phys. Res.*, vol. 113, no. 3, pp. 381–385, 1973. DOI: https://doi.org/10.1016/0029-554X(73)90503-X.

[6] J. Kemmer and G. Lutz, “New structures for position sensitive semiconductor detectors”, *Nucl. Instrum. Methods Phys. Res.*, vol. 273, no. 2, pp. 588–598, 1988. DOI: https://doi.org/10.1016/0168-9002(88)90061-7.

[7] C. J. S. Damerell et al., “A vertex detector for SLD”, *Nucl. Instrum. Methods Phys. Res.*, vol. 275, no. 3, pp. 484–493, 1989. DOI: https://doi.org/10.1016/0168-9002(89)90735-3.

[8] E. H. M. Heijne et al., “The silicon micropattern detector: A dream?”, *Nucl. Instrum. Methods Phys. Res.*, vol. 273, no. 2, pp. 615–619, 1988. DOI: https://doi.org/10.1016/0168-9002(88)90065-4.

[9] L. Rossi et al., "Pixel detectors: from fundamentals to applications". Springer, 2006, ISBN: 978-3-540-28333-1.

[10] The ATLAS Collaboration, “ATLAS detector and physics performance: Technical Design Report, 1”, CERN, Geneva, Tech. Rep., 1999. [Online]. Available: https://cds.cern.ch/record/391176.

[11] G. R. Lynch and O. I. Dahl, “Approximations to multiple coulomb scattering”, *Nucl. Instrum. Methods Phys. Res.*, vol. 58, no. 1, pp. 6–10, 1991. DOI: https://doi.org/10.1016/0168-583X(91)95671-Y.

[12] H. Kolanoski and N. Wermes, "Particle detectors". Oxford University Press, 2020, ISBN: 978-0-19-885836-2.

[13] V. A. J. Van Lint, "The physics of radiation damage in particle detectors", *Nucl. Instrum. Methods Phys. Res.*, vol. 253, no. 3, pp. 453–459, 1987. DOI: https://doi.org/10.1016/0168-9002(87)90532-8.

[14] L. Evans and P. Bryant, "LHC Machine", *J. Instrum.*, vol. 3, no. 08, S08001, 2008. DOI: 10.1088/1748-0221/3/08/S08001.

[15] The ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", *J. Instrum.*, vol. 3, no. 08, S08003, 2008. DOI: 10.1088/1748-0221/3/08/S08003.

[16] The CMS Collaboration, "The CMS Experiment at the CERN LHC", *J. Instrum.*, vol. 3, no. 08, S08004, 2008. DOI: 10.1088/1748-0221/3/08/S08004.

[17] The ALICE Collaboration, "The ALICE experiment at the CERN LHC", *J. Instrum.*, vol. 3, no. 08, S08002, 2008. DOI: 10.1088/1748-0221/3/08/S08002.

[18] The LHCb Collaboration, "The LHCb Detector at the LHC", *J. Instrum.*, vol. 3, no. 08, S08005, 2008. DOI: 10.1088/1748-0221/3/08/S08005.

[19] C. De Melis, "The CERN accelerator complex. Complexe des accélérateurs du CERN", 2016. [Online]. Available: https://cds.cern.ch/record/2119882.

[20] D. Brandt, "CAS - CERN Accelerator School: Intermediate Accelerator", CERN, Geneva, 2006. [Online]. Available: http://cds.cern.ch/record/603056.

[21] The ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Physics Letters B*, vol. 716, no. 1, pp. 1–29, 2012. DOI: https://doi.org/10.1016/j.physletb.2012.08.020.

[22] The CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Physics Letters B*, vol. 716, no. 1, pp. 30–61, 2012. DOI: https://doi.org/10.1016/j.physletb.2012.08.021.

[23] "High-Luminosity Large Hadron Collider (HL-LHC): Technical design report", CERN, Geneva, Tech. Rep., 2020. DOI: 10.23731/CYRM-2020-0010.

[24] "The Inner Detector of the ATLAS experiment at the LHC". [Online]. Available: https://atlas.cern/Discover/Detector/Inner-Detector.

[25] H. Pernegger, "The Pixel Detector of the ATLAS experiment for LHC Run-2", *J. Instrum.*, vol. 10, no. 06, p. C06012, 2015. DOI: 10.1088/1748-0221/10/06/C06012.

[26] The ATLAS Collaboration, "The Silicon Microstrip Sensors of the ATLAS SemiConductor Tracker", *Nucl. Instrum. Methods Phys. Res.*, vol. 578, no. 1, pp. 98–118, 2007. DOI: 10.1016/j.nima.2007.04.157.

[27] The ATLAS Collaboration, "The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance", *J. Instrum.*, vol. 3, no. 02, P02013, 2008. DOI: 10.1088/1748-0221/3/02/P02013.

[28] The ATLAS Collaboration, "Alignment of the ATLAS Inner Detector in Run-2", *Eur. Phys. J. C*, vol. 80, 2021. DOI: https://doi.org/10.1140/epjc/s10052-020-08700-6.

[29] The ATLAS Collaboration, "Technical Design Report for the ATLAS Inner Tracker Pixel Detector", CERN, Geneva, Tech. Rep., 2017. DOI: 10.17181/CERN.FOZZ.ZP3Q.

[30] "Expected tracking and related performance with the updated ATLAS Inner Tracker layout at the High-Luminosity LHC". [Online]. Available: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2021-024/.

[31] The ALICE Collaboration, "Technical Design Report for the Upgrade of the ALICE Inner Tracking System", *J. Phys.*, vol. 41, no. 8, p. 087 002, 2014. DOI: 10.1088/0954-3899/41/8/087002.

[32] M. Mager et al., "ALPIDE, the Monolithic Active Pixel Sensor for the ALICE ITS upgrade", *Nucl. Instrum. Methods Phys. Res.*, vol. 824, pp. 434–438, 2016. DOI: https://doi.org/10.1016/j.nima.2015.09.057.

[33] L. Musa, "Letter of Intent for an ALICE ITS Upgrade in LS3", CERN, Geneva, 2019. DOI: 10.17181/CERN-LHCC-2019-018.

[34] J. Zhu et al., "Systematic experimental study on stitching techniques of CMOS image sensors", *IEICE Electronics Express*, vol. 13, no. 15, pp. 20 160 441–20 160 441, 2016. DOI: 10.1587/elex.13.20160441.

[35] C. Gargiulo, "Space Frame and Cold Plate Characterization", CERN, Geneva, Tech. Rep., 2016. [Online]. Available: https://indico.cern.ch/event/518710/.

[36] D. A. van den Ende et al., "Mechanical and electrical properties of ultra-thin chips and flexible electronics assemblies during bending", *Microelectronics Reliability*, vol. 54, no. 12, pp. 2860–2870, 2014. DOI: https://doi.org/10.1016/j.microrel.2014.07.125.

[37] The ALICE Collaboration, "First demonstration of in-beam performance of bent Monolithic Active Pixel Sensors", *Nucl. Instrum. Methods Phys. Res.*, vol. 1028, p. 166 280, 2022. DOI: https://doi.org/10.1016/j.nima.2021.166280.

[38] M.L Prydderch et al., "A 512 × 512 CMOS Monolithic Active Pixel Sensor with integrated ADCs for space science", *Nucl. Instrum. Methods Phys. Res.*, vol. 512, no. 1, pp. 358–367, 2003. DOI: https://doi.org/10.1016/S0168-9002(03)01914-4.

[39] C. Granja et al., "Quantum imaging monitoring and directional visualization of space radiation with timepix based SATRAM spacecraft payload in open space on board the ESA Proba-V satellite", *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, pp. 1–2, 2014. DOI: 10.1109/NSSMIC.2014.7431205.

[40] F. Cannillo et al., "Monolithic active pixel sensor for dosimetry application", *ESSDERC '03*, pp. 63–66, 2003. DOI: 10.1109/ESSDERC.2003.1256811.

[41] A. Zang et al., "The Dosepix detector, an energy-resolving photon-counting pixel detector for spectrometric measurements", *J. Instrum.*, vol. 10, no. 04, p. C04015, 2015. DOI: 10.1088/1748-0221/10/04/C04015.

[42] A. P. H. Butler et al., “Bio-medical X-ray imaging with spectroscopic pixel detectors”, *Nucl. Instrum. Methods Phys. Res.*, vol. 591, no. 1, pp. 141–146, 2008. DOI: https://doi.org/10.1016/j.nima.2008.03.039.

[43] I. Schilling et al., “Measuring the Beam Energy in Proton Therapy Facilities Using ATLAS IBL Pixel Detectors”, *Instruments*, vol. 6, no. 4, 2022. DOI: 10.3390/instruments6040080.

[44] K. Abe et al., “Design and performance of the SLD vertex detector: a 307 Mpixel tracking system”, *Nucl. Instrum. Methods Phys. Res.*, vol. 400, no. 2, pp. 287–343, 1997. DOI: https://doi.org/10.1016/S0168-9002(97)01011-5.

[45] I. Sedgwick et al., “LASSENA: A 6.7 Megapixel, 3-sides Buttable Wafer-Scale CMOS Sensor using a novel grid-addressing architecture”, *Proc. of IISW*, 2013. [Online]. Available: https://imagesensors.org/2013-papers/.

[46] M Esposito et al., “DynAMITe: a wafer scale sensor for biomedical applications”, *J. Instrum.*, vol. 6, no. 12, p. C12064, 2011. DOI: 10.1088/1748-0221/6/12/C12064.

[47] Particle Data Group, “Passage of Particles Through Matter”, 2022. [Online]. Available: https://pdg.lbl.gov/2022/reviews/rpp2022-rev-passage-particles-matter.pdf.

[48] R. M. Sternheimer, “The Density Effect for the Ionization Loss in Various Materials”, *Phys. Rev.*, vol. 88, pp. 851–859, 4 1952. DOI: 10.1103/PhysRev.88.851.

[49] L. V. Landau, “On the energy loss of fast particles by ionization”, *Collected Papers of L. D. Landau*, pp. 417–424, 1965. DOI: https://doi.org/10.1016/B978-0-08-010586-4.50061-4.

[50] H. Bichsel, “Straggling in thin silicon detectors”, *Rev. Mod. Phys.*, vol. 60, pp. 663–699, 3 1988. DOI: 10.1103/RevModPhys.60.663.

[51] S Meroli et al., “Energy loss measurement for charged particles in very thin silicon layers”, *J. Instrum.*, vol. 6, no. 06, P06013, 2011. DOI: 10.1088/1748-0221/6/06/P06013.

[52] U. Fano, “Ionization Yield of Radiations. II. The Fluctuations of the Number of Ions”, *Phys. Rev.*, vol. 72, pp. 26–29, 1 1947. DOI: 10.1103/PhysRev.72.26.

[53] D. A. Neamen, "Semiconductor physics and devices: basic principles". Springer, 2011, ISBN: 0073529583.

[54] Wikipedia, “P-n junction”, [Online]. Available: https://en.wikipedia.org/wiki/P-n_junction.

[55] S. Ramo, “Currents Induced by Electron Motion”, *Proc. of the IRE*, vol. 27, pp. 584–585, 9 1939. DOI: 10.1109/JRPROC.1939.228757.

[56] W. Hoffelner, "Irradiation Damage in Nuclear Power Plants". Springer, 2015, ISBN: 978-1-4614-5588-2.

[57] B. L. Gregory et al., “Injection Dependence of Transient Annealing in Neutron-Irradiated Silicon Devices”, *IEEE Trans. Nucl. Sci.*, vol. 14, no. 6, pp. 116–126, 1967. DOI: 10.1109/TNS.1967.4324783.

[58] M. Garcia-Sciveres and N. Wermes, "A review of advances in pixel detectors for experiments with high rate and radiation", *Reports on Progress in Physics*, vol. 81, no. 6, 2018. DOI: 10.1088/1361-6633/aab064.

[59] E. Cabruja et al., "Special bump bonding technique for silicon pixel detectors", *Nucl. Instrum. Methods Phys. Res.*, vol. 576, no. 1, pp. 150–153, 2007. DOI: https://doi.org/10.1016/j.nima.2007.01.143.

[60] I. Peric et al., "The FEI3 readout chip for the ATLAS pixel detector", *Nucl. Instrum. Methods Phys. Res.*, vol. 565, no. 1, pp. 178–187, 2006. DOI: https://doi.org/10.1016/j.nima.2006.05.032.

[61] A. Starodumov et al., "High rate capability and radiation tolerance of the PROC600 readout chip for the CMS pixel detector", *J. Instrum.*, vol. 12, no. 01, p. C01078, 2017. DOI: 10.1088/1748-0221/12/01/C01078.

[62] M. Prathapan et al., "Towards the large area HVCMOS demonstrator for ATLAS ITk", *Nucl. Instrum. Methods Phys. Res.*, vol. 936, pp. 389–391, 2019. DOI: https://doi.org/10.1016/j.nima.2018.11.022.

[63] C.A. Marin Tobon, "PADRE pixel read-out architecture for Monolithic Active Pixel Sensor for the new ALICE Inner Tracking System in TowerJazz 180 nm technolog", Ph.D. dissertation, Universitat Politècnica de València, 2017. DOI: 10251/86154.

[64] T. Wang et al., "Depleted fully monolithic CMOS pixel detectors using a column based readout architecture for the ATLAS Inner Tracker upgrade", *J. Instrum.*, vol. 13, no. 03, p. C03039, 2018. DOI: 10.1088/1748-0221/13/03/C03039.

[65] K. Moustakas et al., "CMOS monolithic pixel sensors based on the column-drain architecture for the HL-LHC upgrade", *Nucl. Instrum. Methods Phys. Res.*, vol. 936, pp. 604–607, 2019. DOI: https://doi.org/10.1016/j.nima.2018.09.100.

[66] J. R. Pierce, "Physical sources of noise", *Proc. of the IRE*, vol. 44, no. 5, pp. 601–608, 1956. DOI: 10.1109/JRPROC.1956.275123.

[67] Y. Tsividis and C. McAndrew, *Operation and modeling of the MOS transistor*. Oxford University Press, 2011, ISBN: 978-0195170153.

[68] Alan L. McWhorter, "1/f noise and related surface effects in germanium", Ph.D. dissertation, Massachusetts Institute of Technology, 1955. DOI: http://hdl.handle.net/1721.1/12061.

[69] Z. Y. Chang and W. Sansen, "Effect of $1/f$ noise on the resolution of CMOS analog readout systems for microstrip and pixel detectors", *Nucl. Instrum. Methods Phys. Res.*, vol. 305, no. 3, pp. 553–560, 1991. DOI: https://doi.org/10.1016/0168-9002(91)90156-K.

[70] A. Rivetti, *CMOS: Front-End Electronics for Radiation Sensors*. CRC Press, 2015, ISBN: 9781138827387.

[71] E. Gatti and P. F. Manfredi, "Processing the Signals From Solid State Detectors in Elementary Particle Physics", *Rivista del Nuovo Cimemto*, vol. 9, pp. 1–146, 1986. DOI: 10.1007/BF02822156.

[72] W. Snoeys, "Monolithic pixel detectors for high energy physics", *Nucl. Instrum. Methods Phys. Res.*, vol. 731, pp. 125–130, 2013. DOI: https://doi.org/10.1016/j.nima.2013.05.073.

[73] J. R. Schwank et al., "Radiation Effects in MOS Oxides", *IEEE Trans. Nucl. Sci.*, vol. 55, no. 4, pp. 1833–1853, 2008. DOI: 10.1109/TNS.2008.2001040.

[74] F. Faccio and G. Cervelli, "Radiation-induced edge effects in deep submicron CMOS transistors", *IEEE Trans. Nucl. Sci.*, vol. 52, no. 6, pp. 2413–2420, 2005. DOI: 10.1109/TNS.2005.860698.

[75] F. Faccio et al., "Radiation-Induced Short Channel (RISCE) and Narrow Channel (RINCE) Effects in 65 and 130 nm MOSFETs", *IEEE Trans. Nucl. Sci.*, vol. 62, no. 6, pp. 2933–2940, 2015. DOI: 10.1109/TNS.2015.2492778.

[76] W. Deng, "Design and Development of a Stitched Monolithic Active Pixel Sensor Prototype for the ALICE ITS Upgrade", Ph.D. dissertation, Central China Normal University, 2022.

[77] W. Snoeys et al., "Layout techniques to enhance the radiation tolerance of standard CMOS technologies demonstrated on a pixel detector readout chip", *Nucl. Instrum. Methods Phys. Res.*, vol. 439, no. 2, pp. 349–360, 2000. DOI: https://doi.org/10.1016/S0168-9002(99)00899-2.

[78] G. Anelli et al., "Radiation tolerant VLSI circuits in standard deep submicron CMOS technologies for the LHC experiments: practical design aspects", *IEEE Trans. Nucl. Sci.*, vol. 46, no. 6, pp. 1690–1696, 1999. DOI: 10.1109/23.819140.

[79] F. Faccio, "Radiation Effects on Electronics", Fifth Barcelona Techno week, 2021.

[80] I. Berdalovic et al., "Monolithic pixel development in TowerJazz 180 nm CMOS for the outer pixel layers in the ATLAS experiment", *J. Instrum.*, vol. 13, no. 01, p. C01023, 2018. DOI: 10.1088/1748-0221/13/01/C01023.

[81] I. Berdalovic et al., "MALTA: a CMOS pixel sensor with asynchronous readout for the ATLAS High-Luminosity upgrade", *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, pp. 1–4, 2018. DOI: 10.1109/NSSMIC.2018.8824349.

[82] E.J. Schioppa et al., "Measurement results of the MALTA monolithic pixel detector", *Nucl. Instrum. Methods Phys. Res.*, vol. 958, p. 162 404, 2020. DOI: https://doi.org/10.1016/j.nima.2019.162404.

[83] M. Dyndal et al., "Mini-MALTA: radiation hard pixel designs for small-electrode monolithic CMOS sensors for the High Luminosity LHC", *J. Instrum.*, vol. 15, no. 02, P02005, 2020. DOI: 10.1088/1748-0221/15/02/P02005.

[84] F. Piro et al., "A 1-μW Radiation-Hard Front-End in a 0.18-μm CMOS Process for the MALTA2 Monolithic Sensor", *IEEE Trans. Nucl. Sci.*, vol. 69, no. 6, pp. 1299–1309, 2022. DOI: 10.1109/TNS.2022.3170729.

[85] W. Snoeys et al., "A process modification for CMOS monolithic active pixel sensors for enhanced depletion, timing performance and radiation tolerance", *Nucl. Instrum. Methods Phys. Res.*, vol. 871, pp. 90–96, 2017. DOI: https://doi.org/10.1016/j.nima.2017.07.046.

[86] M. Munker et al., "Simulations of CMOS pixel sensors with a small collection electrode, improved for a faster charge collection and increased radiation tolerance", *J. Instrum.*, vol. 14, no. 05, p. C05013, 2019. DOI: 10.1088/1748-0221/14/05/C05013.

[87] H. Hillemanns et al., "Radiation hardness and detector performance of new 180nm CMOS MAPS prototype test structures developed for the upgrade of the ALICE Inner Tracking System", *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, pp. 1–5, 2013. DOI: 10.1109/NSSMIC.2013.6829475.

[88] M. Munker, "Test beam and simulation studies on high resistivity cmos pixel sensors", Ph.D. dissertation, University of Bonn, 2018. DOI: urn:nbn:de:hbz:5n-51913.

[89] C. Gao et al., "A novel source-drain follower for monolithic active pixel sensors", *Nucl. Instrum. Methods Phys. Res.*, vol. 831, pp. 147–155, 2016. DOI: https://doi.org/10.1016/j.nima.2016.03.074.

[90] M. J. M. Pelgrom et al., "Matching properties of MOS transistors", *IEEE J. Solid-State Circuits*, vol. 24, no. 5, pp. 1433–1439, 1989. DOI: 10.1109/JSSC.1989.572629.

[91] H. Hillemanns et. al., "Total ionizing dose effects in 180 nm CMOS structures for monolithic active pixel sensors for the ALICE Inner Tracking System detector upgrade", ALICE internal note, 2013.

[92] T. Kugathasan et al., "Monolithic pixel development in 180 nm cmos for the outer pixel layers in the atlas experiment", *Proc. Sci.*, vol. TWEPP-17, p. 047, 2018. DOI: 10.22323/1.313.0047.

[93] R. Cardella et al., "Lapa, a 5 gb/s modular pseudo-lvds driver in 180 nm cmos with capacitively coupled pre-emphasis", *Proc. Sci.*, vol. TWEPP-17, p. 038, 2018. DOI: 10.22323/1.313.0038.

[94] L. Perktold and J. Christiansen, "A multichannel time-to-digital converter ASIC with better than 3 ps RMS time resolution", *J. Instrum.*, vol. 9, no. 01, p. C01060, 2014. DOI: 10.1088/1748-0221/9/01/C01060.

[95] L. Snoj et al., "Computational analysis of irradiation facilities at the JSI TRIGA reactor", *Appl. Radiat. Isotopes*, vol. 70, no. 3, pp. 483–488, 2012. DOI: https://doi.org/10.1016/j.apradiso.2011.11.042.

[96] "EP R&D: Programme on Technologies for Future Experiments". [Online]. Available: https://ep-rnd.web.cern.ch/.

[97] W. Snoeys et al., "Optimization of a 65 nm CMOS Imaging Technology for Monolithic Sensors in High Energy Physics", *Proc. Sci.*, vol. Pixel2022, p. 083, 2023. DOI: 10.22323/1.420.0083.

[98] J. Hasenbichler, "Development of Novel Pixel CMOS Sensors Optimised for Time Resolution", Ph.D. dissertation, Vienna University of Technology, 2021. DOI: 10.34726/hss.2021.63185.

[99] A. Dorda et al., "Measurements of total ionizing dose effects in TPSCo 65 nm and influence of NMOS bulk bias", *J. Instrum.*, vol. 18, no. 02, p. C02036, 2023. DOI: 10.1088/1748-0221/18/02/C02036.

[100] I. Sanna et al., "Characterization of first prototypes fabricated in 65 nm CMOS technology for the ALICE Inner Tracker System upgrade", *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, 2023.

[101] B. Lim et al., "Thin monolithic pixel sensors with fast operational amplifier output in a 65 nm imaging technology for ALICE ITS3", *Nucl. Instrum. Methods Phys. Res.*, vol. 1055, p. 168 478, 2023. DOI: https://doi.org/10.1016/j.nima.2023.168478.

[102] L. Cecconi et al., "Design and readout architecture of a monolithic binary active pixel sensor in TPSCo 65 nm CMOS imaging technology", *J. Instrum.*, vol. 18, no. 02, p. C02025, 2023. DOI: 10.1088/1748-0221/18/02/C02025.

[103] G. Aglieri Rinella et al., "Digital Pixel Test Structures implemented in a 65 nm CMOS process", *arXiv*, 2023. DOI: 10.48550/ARXIV.2212.08621.

[104] Tower Semiconductor Ltd, "Stitching design rules for forming interconnect layers", US Patent 6225013B1, 2001. [Online]. Available: https://patents.google.com/patent/US6225013B1/en.

[105] G. Aglieri Rinella, "Developments of stitched monolithic pixel sensors towards the ALICE ITS3", *Nucl. Instrum. Methods Phys. Res.*, vol. 1049, p. 168 018, 2023. DOI: https://doi.org/10.1016/j.nima.2023.168018.

[106] A. Marchioro, "Transistors below 10 nm, and what to do with them in HEP", EP-ESE Electronics Seminars at CERN, Geneva, 2019. [Online]. Available: https://indico.cern.ch/event/788033/.

Francesco Piro

LinkedIn: https://www.linkedin.com/in/francesco-piro/ francesco.piro93@gmail.com

EDUCATION

- **École Polytechnique Fédérale de Lausanne - EPFL** — Lausanne, Switzerland
  PhD Candidate in Microsystems and Microelectronics — Feb 2020 - Oct 2023
  - Thesis: Front-end circuits for monolithic CMOS sensors targeting high-energy physics applications.
- **University of Naples Federico II** — Naples, Italy
  MSc in Electronics Engineering, 110/110 cum laude — Oct 2015 - Dec 2017
  - Thesis: Development of a pixel detector in a SiGe BiCMOS technology for time-of-flight measurements.
- **University of Naples Federico II** — Naples, Italy
  BSc in Electronics Engineering, 110/110 — Sep 2012 - Oct 2015
  - Thesis: FPGA implementation and experimental verification of a digital circuit for background identification of a video stream.

EXPERIENCE

- **CERN, Experimental Physics Department, Micro-Electronics group** — Geneva, Switzerland
  PhD student and Senior Fellow, Analog IC Designer — Feb 2020 - Present
  - Joined the EP-R&D program on monolithic pixel sensors for future high-energy physics experiments, focusing in particular on the upgrade of the ALICE Inner Tracking System.
- **CERN, Experimental Physics Department, ATLAS collaboration** — Geneva, Switzerland
  Marie Curie Fellow, Analog IC Designer — Mar 2018 - Dec 2019
  - Joined the MALTA project on the development of monolithic pixel sensors targeting the upgrade of the ATLAS Inner Tracker.
- **UNIGE Department of Nuclear and Particle Physics** — Geneva, Switzerland
  Intern Student, Analog IC Designer — Jun 2017 - Dec 2017
  - Joined the TT-PET project on the development of a monolithic pixel sensor for time-of-flight measurements in SiGe BiCMOS technology with a target time resolution below 100 ps.

SKILLS AND COMPETENCES

Technical Skills:
- Circuit analysis and simulation.
- Layout of integrated circuits.
- Circuit physical verification.
- Measurement instrumentation.

Tools: Cadence Virtuoso Analog Design Suite, Calibre, Assura, Python, MATLAB, Cadence Digital Suite

Languages: English: proficient | Italian: native

PUBLICATIONS

- F. Piro et al., "A 1-µW Radiation-Hard Front-End in a 0.18 µm CMOS Process for the MALTA2 Monolithic Sensor," in *IEEE Trans. Nucl. Sci.*, June 2022, doi: 10.1109/TNS.2022.3170729.
- F. Piro et al., "A Compact Front-End Circuit for a Monolithic Sensor in a 65 nm CMOS Imaging Technology," in *IEEE Trans. Nucl. Sci.*, September 2023, doi: 10.1109/TNS.2023.3299333.
- Co-author in more than ten conference and journal papers (ORCID: https://orcid.org/0000-0002-6862-275X).