UNIVERSITÀ DI PISA

Department of Physics
Masters Degree in Physics

Characterization of monolithic CMOS pixel sensors for charged particle detectors and for high intensity dosimetry

Candidate:
Eleonora Ravera

Supervisors:
Prof. Francesco Forti
Prof.ssa Maria Giuseppina Bisogni

Academic year 2021/2022

Contents

Introduction

Since the 1980s, when the fabrication of device with very small electrodes (50-100 µm) became a practical possibility, pixel detectors have been widely employed for imaging and tracking charged particles in the vertex region of experiments at accelerators. Thanks to their excellent spatial resolution, today even better than 10 µm, they allow for true three dimensional space-point determination even at high particle fluxes and in particular for the identification of secondary vertices of short-lived particles such as τ and B mesons. Requirement imposed by accelerators are stringent and they will become even more so with the increase of luminosity; in this scenario CMOS Monolithic Active Pixel Sensors (MAPS), based on the technology of CMOS cameras, are being developed to improve the performance of the hybrid pixel detectors, which currently constitute the state-of-art for large scale pixel detector, in particular by reducing the amount of material, power consumption and pixel dimension. Indeed, while hybrid pixels are made by two parts, the sensor and the electronics, welded together through microconnections, the MAPS integrate them all on the same wafer.

Experiments such as ALICE at LHC and STAR at RHIC have already introduced the CMOS MAPS technology in their detectors. ALICE Tracking System (ITS2), upgraded during the LHC long shut down in 2019-20, was the first large-area ($\sim$10 m^2) silicon vertex detector based on CMOS MAPS. Thanks to the reduction of the material budget, ITS2, which uses the ALPIDE chip developed by ALICE collaboration, obtained an amazing improvement both in the position measurement and in the momentum resolution, improving the efficiency of track reconstruction for particle with very low transverse momentum (by a factor 6 at $p_T \sim 0.1$ GeV/c). Further advancements in CMOS MAPS technology are being aggressively pursued for the ALICE ITS3 and the Belle II vertex detector upgrades (both foreseen around 2026-27), and by the R&D53 collaboration for the upgrade at HL-LHC, with the goals of further reducing the sensor thickness and improving the readout speed of the devices, while keeping power consumption at a minimum.

Beside tracking, the development of pixel detectors is a very active field with many applications: a noteworthy example of detector originally used in particle physics and later employed for medical imaging, in space detectors and for art authentication, is Medipix,

a hybrid system developed at CERN within the Medipix collaboration. Among medical applications, a possible use of CMOS MAPS could be in dosimetry: in the last few years the search of radiotherapy oncological treatments with high intensity beams (FLASH mode) is requiring new dosimeters, both for the therapies as well as new beam-monitors (especially for focused very high energy electron beams), which are capable of deal with extreme dose rate (up to 40 Gy/s).

I have studied the characteristics of two ALPIDE-like CMOS MAPS chips and tested them under different front end configuration. The first chip, the TJ-Monopix1 from the Monopix series initially developed for ATLAS ITK by the R&D collaboration, is a TowerJazz MAPS fabricated in 180 nm CMOS technology with an active area of 1×2 cm^2 (448×224 pixels) and is one of the prototypes for the Belle II vertex detector upgrade. The second chip, called Main Demonstrator-1, has an active area of 1.28×1.28 cm^2 (512×512 pixels) is produced by LFoundry in 110 nm CMOS technology and designed by the ARCADIA (Advanced Readout CMOS Architectures with Depleted Integrated sensor Arrays) group; it is intended to be a general purpose device with possible use in medical scanners, space experiments, future lepton colliders and also possibly X-ray applications with thick substrates. The main differences between the two chips are in the output signal type and in the readout sequence of the matrix. Concerning the former, TJ-Monopix1 returns an analog output information, that is the time over threshold of the pulse, which can be related with the charge released by the particle in the sensor, while MD1 returns only a digital information; regarding the latter, instead, TJ-Monopix1 has a completely sequential readout, while MD1 roughly combines the information of the hits before the readout in order to reduce the data transmission time.

I have set up the test systems for the two chips in the INFN clean laboratories and characterized the devices electrically and with radioactive sources in terms of threshold, noise, dead time and analog response. The mean minimum stable threshold evolved through different generation of chips and nowadays it is less than 500 e−, allowing thinner sensors with smaller signals: TJ-Monopix1 has proven to be in agreement with this trend, having a threshold of ∼400 e−, to be compared with the 2000 e− signal expected for a minimum ionizing particle in an epitaxial layer of 25 µm. Moreover, since one of the main challenges of MAPS are the differences between pixels due to process parameters variation across the wafer, which make the sensor response nonuniform, I have measured the threshold and noise dispersion across the matrix, which I found to be 40 e− and 2 e− respectively. I have also studied the response of the analog signal recorded by TJ-Monopix1, that is the time over threshold, and performed a calibration of its absolute value using a Fe55 X-ray source. All these measurements are important to verify the design parameters of the chip and to validate the chip simulation.

As conclusion of the measurement campaign, we have tested TJ-Monopix1 at very high intensity using the electron beam of the new ElectronFlash accelerator designed for both medical research and R&D in FLASH-radiotherapy and recently installed at Santa Chiara hospital in Pisa. I have participated in the design of the setup needed for testbeam measurement and I am currently working on the analysis of the data collected.

CHAPTER 1

Use of pixel detectors

This chapter tries to summarize the main applications of pixel detectors, while a description of the technological implementation is presented in chapter 2; in the following sections I will also use terms such as hybrid pixels, monolitich active pixel systems (MAPS), Charge Coupled Devices (CCDs), Depleted Field Effect Transistor sensors (DEPFET) that will discussed later.

The relation between the development of cameras and that of pixel detectors dates back to 1969, when the idea of CCDs, for which Boyle and Smith were awarded the Nobel Prize in Physics in 2009, revolutionized photography allowing light to be captured electronically instead of on film. Even though the CMOS technology already existed at the time the CCDs spread, the costs of productions were too high to allow the diffusion of these sensors for the following 20 years. From that moment on, the fast diffusion of CMOS was mainly due to the less cost than CCD, and the less power supply required. Nowadays CCDs are still prefered over MAPS in astronomy, where the astronomical sources' rate are low enough to cope with slow readout time (tens of ms).

The principal use cases of pixel detectors in physics are particle tracking and imaging: in the former case individual charged particles have to be identified, in the latter instead an image is obtained by the usually un-triggered accumulation of the impinging radiation.

1.1 Tracking in HEP

In the early days of high-energy physics gaseous detector where used for tracking and there was no need to replace them since they had a sufficient spatial resolution (100 µm). Since 1974, with the measurement of the invariant mass of the J/Psi and the affirmation of the quark model, all experiments start to look for better spatial resolutions in order to achieve the possibility of reconstructing short lived particles and measuring their decays length.

Historically[1], the first pixel detector employed in particle physics was a CCD: it was

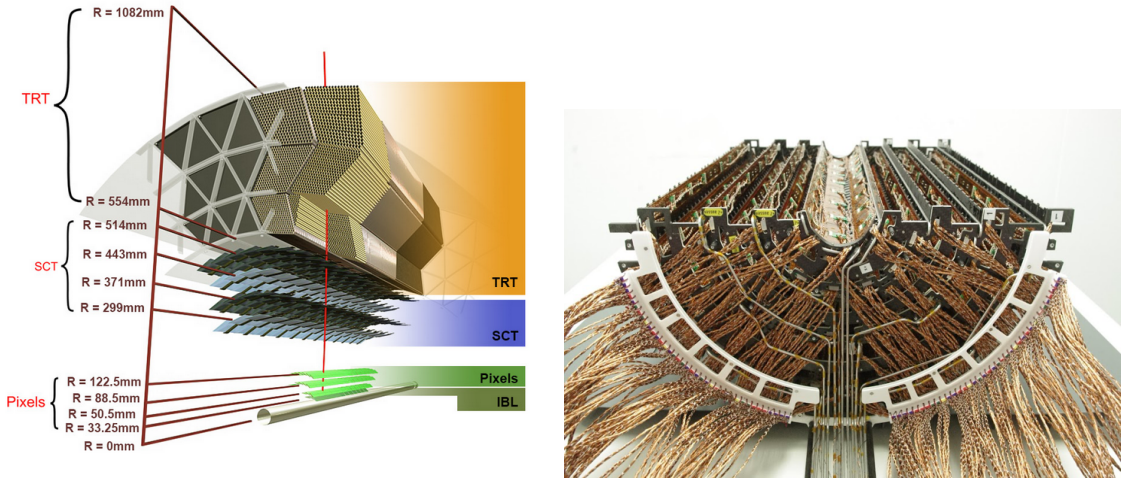
installed in the spectrometer at the CERN's Super Proton Synchrotron by the ACCMOR Collaboration (Amsterdam, CERN, Cracow, Munich, Oxford, RAL) in the mid 1980s, with the purpose of studying the (at the time) recently-discovered charmed particles. The second famous usage of CCDs took place in the SLAC Large Detector at SLAC linear collider in the years 1996-98, where the CCD technology was adopted instead of the microstrip detectors for their excellent spatial resolution (cell size $22\times22\mu m^2$ giving a resolution of $\sim5\,\mu m$) thanks to the sufficient time for readout between two successive collisions (160 ms).

From that period on, particle tracking in HEP experiments has been transformed radically. It became mandatory to build a inner vertex detector, with the following tasks:

- pattern recognition with the identification of particle tracks even in the presence of large backgrounds and pile-up
- measurement of vertices (primary and secondary)
- multi-track and vertex separation in the core of jets
- measurement of specific ionization
- momentum measurement combining with the information from other detectors

The more demanding requirements led to the development of hybrid pixel detectors starting from 1990s: a dedicated collaboration, RD19, was established at CERN with the specific goal of defining a semiconductor micropattern detector with an incorporated signal processing at a microscopic level. In those years a wide set of prototypes of hybrid pixel has been manufactured; among the greatest productions a mention goes to the huge ATLAS and CMS vertex detectors. From the middle of 2013 a second collaboration, RD53, has been established with the new goal of finding a pixel detector suitable for the phase II in future upgrades of those experiments. Requirements imposed by LHC are stingent and they will become even more with the increase of luminosity at HL-LHC: for example, a dose and radiation of 500 Mrad and 10^{16} MeV n_{eq}/cm^2 are exepcted after 5 years of operation. Time resolution, material budget and power consumption are also issues for the upgrade: to distinguish different events from different bunches a time resolution better than 25 ns for a bunch crossing frequency of 40 MHz is required, a material budget lower than 2% X_0 and a power consuption lower than $500\,mW/cm^2$ are required.

Even if the collaboration is specifically focused oFn the design of hybrid pixel readout chips (aiming to 65 nm tecnique), also other options have been taken in account per quanto riguarda the sensor. Among the solutions proposed to improve radiation robustness of the sensor, 3D silicon detector, invented by Sherwood Parker in 1995, are very promising. In 3D sensors the electrode is a narrow column of n-type implanted vertically across the bulk instead of being implanted on the wafer's surface. The charge produced by the impinging particle is then drifted transversally within the pixel, and, as the mean path between two electrode can be soufficent low, the trap probability is not an issue. Even if 3D detector are adequately radiation hard and are a possible solution for hybrid pixel modules, especially



(a) ATLAS tracker detector

(b) CMS barrel pixel detector

Figure 1.1

in the innermost pixel detector layer, their fabrication process is currently low volume, making them unlikely to cover large areas. Another promising possibility is to use fast Monolitich Active Pixels systems which could allow the reduction of material budget and improve spatial resolution.

1.1.1 Hybrid Pixels at LHC

ATLAS

ATLAS is one of two general-purpose detectors at the LHC and has the largest volume detector ever constructed for a particle collider (46 m long and 25 m in diameter). The Inner Detector (ID) consists of three different systems all immersed in a magnetic field parallel to the beam axis whose main components are: the pixel, the micro-strips and transition radiation trackers (fig.1.1a). Concerning the pixel detector, they installed a 3-layer hybrid pixel detector in 2007 and an additional one inserted within the original detector envelope and therefore called insertable B-layer (IBL) in 2014. 92 million pixels are divided in 4 barrel layers and 3 disks in each end-cap region, covering a total area of $1.9\,\mathrm{m}^2$ and having a 15 kW of power consumption.

As stated by the ATLAS collaboration the pixel detector is exposed by an extreme particle flux: "By the end of Run 3[1], the number of particles that will have hit the innermost pixel layers will be comparable to the number it would receive if it were placed only a few kilometres from the Sun during a solar flare". Considering that the particle density will increase even more with HL-LHC, radiation hardness is definitively a target to achieve. Also the complexity of the readout will be raised, as the number of pixels will be increased of a factor about 7, passing from 92 milions to 6 billion, then a readout logic will have to meet the high amount of data.

Hybrid pixels will be used at the start of high-luminosity application, although an

[1]Run 3 start in June 2022

active development of monolitich devices is ongoing for possible future use in the outer pixel layers. The TJ-Monopix1 which I will describe in chapter 3 is part of this development.

Regarding the sensor, a valueable option is using 3D pixels, which have already proved themselves in ATLAS, for the IBL, where they were introduced in a limited acceptance range; the sensor chosen for the upgrade will be bonded with ITkPix, the first full-scale 65 nm hybrid pixel-readout chip developed by the RD53 collaboration.

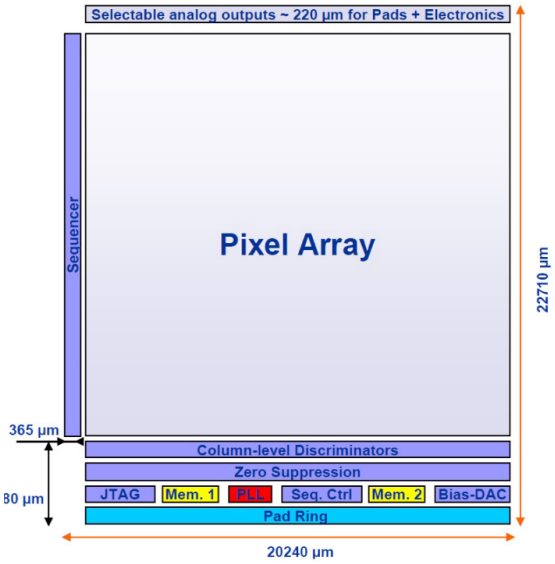
CMS

The CMS hybrid pixel detector (fig.1.1b) has been upgraded in 2017, when, with the replacement of a piece of the beam pipe, a layer has been added to the detector at 3 cm from it. 124 million pixels are divided between the barrel pixel detector (BPIX) and the forward disks (FPIX), with sensors which are different from each other and produced by different foundryes. The sensors have an area equal to 100 µm by 150 µm and have been produced on 285 µm to 300 µm thick wafers.

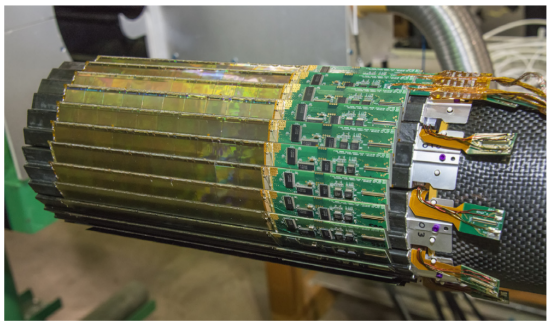
The time resolution is 25 ns, the rate capability 600 MHz/cm^2, and the information coming from the detector are stored on chip for the Level-1 trigger latency ($\sim$4 µs). The upgrade ROIC was redesigned for the outer 3 layers, replacing analog signal readout with on-chip ADCs which record the pulse height information or with digital readout at higher rate.

LHCb

LHCb is a dedicated heavy-flavour physics experiment that exploits pp interactions at 14 TeV at LHC. It was the last experiment to upgrade the vertex detector, the Vertex Locator (VELO), replacing the silicon-strip with 26 plane pixel detector (beacause of the fixed target geometry) in May 2022. As the instantaneous luminosity in Run3 is increased by a factor $\lesssim$10, much of the readout electronics and of the trigger system have been developed in order to cope with the large interaction rate. To place the detector as close as possible to the beampipe and reach a better track reconstruction efficiency and resolution, the VELO has a surprising feature: during the injection of LHC protons it is parket at 3 cm from the beams and only when the stability is reach it is moved at $\sim$5 mm. Readout speed is a priority for the detector that use a triggerless readout at 40 MHz collision rate, producing 20 Gbps per ROIC. The Velopix, which is the hybrid system designed for LHCb, is made by bonding sensors, each measuring 55$\times$55 micrometers and 200 µm-thick, to a 200 µm-thick ASIC specially developed for LHCb and coming from the Medipix family, which can handles hit rates up to 900 MHz per chip. Since the detector is operated under vacuum near the beam pipe, the heat removal is particularly difficult and evaporative CO2 microchannel cooling are used.



(a) Block-diagram of the ULTIMATE-2 sensor of the MIMOSA family



(b) The HFT pixel detector of STAR

Figure 1.2

1.1.2 Monolitich Active Pixels

MIMOSA

MIMOSA [2][3] (standing for Minimum Ionizing MOS Active pixel sensor), designed in 2008, prefigured the architecture of MAPS for coming vertex detector being the first large scale sensor to be employed as detector. MIMOSA-26 equiped the final version of EUDET high resolution beam telescope both at CERN-SPS and at DESY while the MIMOSA-28 devices are used for the first MAPS-based vertex detector at the STAR experiment at RHIC. MIMOSA-26 is fabricated in a 350 nm CMOS technology, and a module features 1152 columns, split into 18 indipendent groups, and 576 rows, with square pixels having a side of 18.4 µm lenght; the epitaxyal layer is not fully depleted and the charge collection is mostly by diffusion, resulting in charge sharing between pixels and collection time bigger than 100 ns. The chip is an Active Pixel Sensor (APS) and therefore it incorporates on pixel the amplification, while the signal discrimination and zero-suppression logic are placed at the End of Column: the readout is done in a rolling shutter mode with a frame integration time that can be lowered down to 85 ms, moreover a memory allows to store up to six hits.

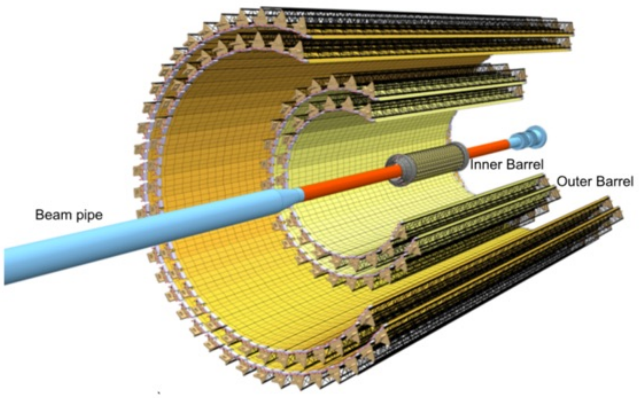
The EUDET telescope, equipped with six sensor planes, requires high granularity and thin pixel detectors in order to achieve an excellent track resolution (around 2 µm) even at the rather low particle energies of 6 GeV. The STAR experiment at the Relativistic Heavy Ion Collide (RHIC) accelerator at the Brookhaven National Laboratory (BNL) is the first to include MAPS in the vertex detector[4]. The main tracking detector in STAR is a TPC with radii 60-190 cm embedded in a 0.5 T solenoidal magnetic field, that provides a pointing resolution of approximately 1 mm. The pixel detector, PXL,

is a part of a 3-detector system, the Heavy Flavor Tracker (HFT), and has been added to the pre-existing STAR apparatus just before the 2014 Run in order to improve the impact parameter resolution and to enable the direct reconstruction of hadronic decays of heavy flavor mesons and baryons. The Heavy Flavor Tracker (fig.1.2b) is composed by the Silicon Strip Detector, the Intermediate Silicon Tracker and the Pixel Detector (PXL); the first one is placed at 22 cm from the beam pipe and consists of double sided strips with 95 µm inter-strip pitch, the second one, placed at 14 cm, is made of single sided silicon pads with 600 µm×6 mm pitch and the last one made by two layes is placed at 2.8 cm and 8 cm fabricated with ULTIMATE-21.2a (also known as MIMOSA-28), a successor of MIMOSA-26 sensor, with pitch 20.7 µm and thinned down to 50 µm (fig.1.2a). An area of 0.16 m^2 is covered by 400 MAPS sensor, corresponding to 356 milions of pixels divided into array size of 928 × 960. Each pixel includes circuitry for readout, amplification, and Correlated Double Sampling for signal extraction and noise subtraction and the frame integration time is 185.6 µs; after the subtraction the signal to noise ratio is ~ 30, with a noise between 10-12 electrons and a signal of 1000 e^-. Thanks to the Heavy Flavor Tracker system and the Pixel Detector, STAR achieved a track pointing resolution of 46 µm for 750 MeV/c kaons, and better than 30 µm for particle momenta bigger than 1 GeV/c: this performance enabled the study of D-meson production with a high significance signal.

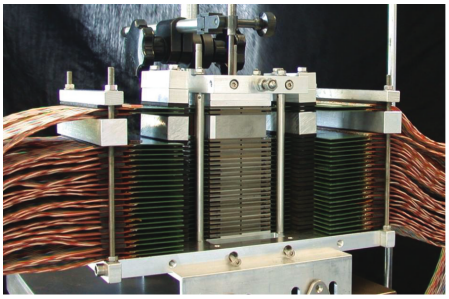
Another possible application of MIMOSA has been proposed by ALICE, which has been studing the possibility of exploiting the extreme granularities of MAPS for calorimeter application[5]. In a such calorimeter, the energy measurement would come out from the counts of particles traversing the active layers, resulting then in a digital calorimeter. A prototype forward calorimeter (FoCAL), fabricated with the MIMOSA23 chips and containing 39 million pixels devided in 24 layers, alternated with 24 layers of tungsten (1.3b), have been tested with electron beams and exhibited an energy resolution better than standard hadronic calorimeters, with a stochastic terms of $30\%/\sqrt{E(GeV)}$, a constant term of 2.8% and noise term of 0.063 GeV.

ALPIDE at ALICE

The ALICE (A Large Ion Collider Experiment) tracking detector consists of the Inner Tracking System (ITS), the gaseous Time Projection Chamber (TPC) and the Transition Radiation Detector, all embedded in a magnetic field of 0.5 T. The ITS (fig.1.3a) is made by six layers of detectors, two for each type, from the interaction point outwards: Silicon Pixel Detector, Silicon Drift Detector and Silicon Strip Detector. Contrary to the others LHC experiments, ALICE tracker in placed in a quite different environments, which enables the usage of a MAPS-based detector: the dose assorbed by the tracker is expected to be smaller by two order of magnitude and the rate of interactions is few MHz instead of 40 MHz, even though the number of particles coming out from each interaction is very high (the Silicon Pixel Detector is invested by a density of particles of $\sim$100 cm^{-2}). The reconstruction of very complicated events with a large number of particle is then a challenge, hence to segment and to minimize the amount of material, which may cause secondary interaction futher complicating the event topology, is considered a viable strategy.



(a) ALICE ITS scheme



(b) FoCAL prototype fabricated with MAPS

Figure 1.3

ITS2, upgraded during the LHC long shut down in 2019-20, was the first large-area ($\sim$10 m^2 covered by 2.5 Gpixels) silicon vertex detector based on CMOS MAPS. The detector employes the ALPIDE chip, developed by ALICE collaboration, fabricated in the 180 nm CMOS Imaging Sensor process of TowerJazz, whose design takes full advantage of process feature which allows full circuitry within the pixel matrix. Thanks to the reduction in the material budget, ITS2 obtained an amazing improvement both in the position measurement and in the momentum resolution, especially improving the efficiency of track reconstruction for particle with very low transverse momentum (by a factor 6 at $p_T \sim 0.1$ GeV/c). Further advancements in CMOS MAPS technology are being aggressively pursued for the ALICE ITS3 vertex detector upgrade (foreseen around 2026-27), with the goals of reducing the sensor thickness and improving the readout speed (which now is completely asynchronous) of the devices, while keeping power consumption at a minimum.

BelleII

Due to the high background level coming from the nanobeam used at SuperKEKB in order to achieve a such high luminosity ($6 \times 10^{35} cm^2/s$), silicon strip cannot be used in the inner layer of the tracker. The occupancy is too high to allow the usage of strips up to 40 mm from the beam pipe. Moreover for a precise reconstruction of B-decay vertices, the usage of thin detector is mandatory at the low energy (4-7 GeV) of the beam, in order to minimize the multiple scattering of particles. The current Vertex Detector of BelleII, VXD, is made of a Pixel Detector (PXD), fabricated with 2 layers of DEPFET-based pixels, and 4 layers of a double-sided silicon strip detectors (SVD)[6]. Due to the small capacitance of the collection node, DEPFET presents a high signal-to-noise ratio (in 30-50) thanks to the low instrinsic noise and to the large signal achieved with he fully depleted bulk: pixels are thinned to 75 µm in the active region, then a Minimum Ionizing Particle (MIP) is expected to create a signal of $\sim 6000\, e^-$ (eq.2.1), while the typical noise of DEPFET is around $200\, e^-$. The ASIC read out is still based on a rolling shutter logic, with an integration time of 20 µs. In order to reduce the data-storage memory PXD hits

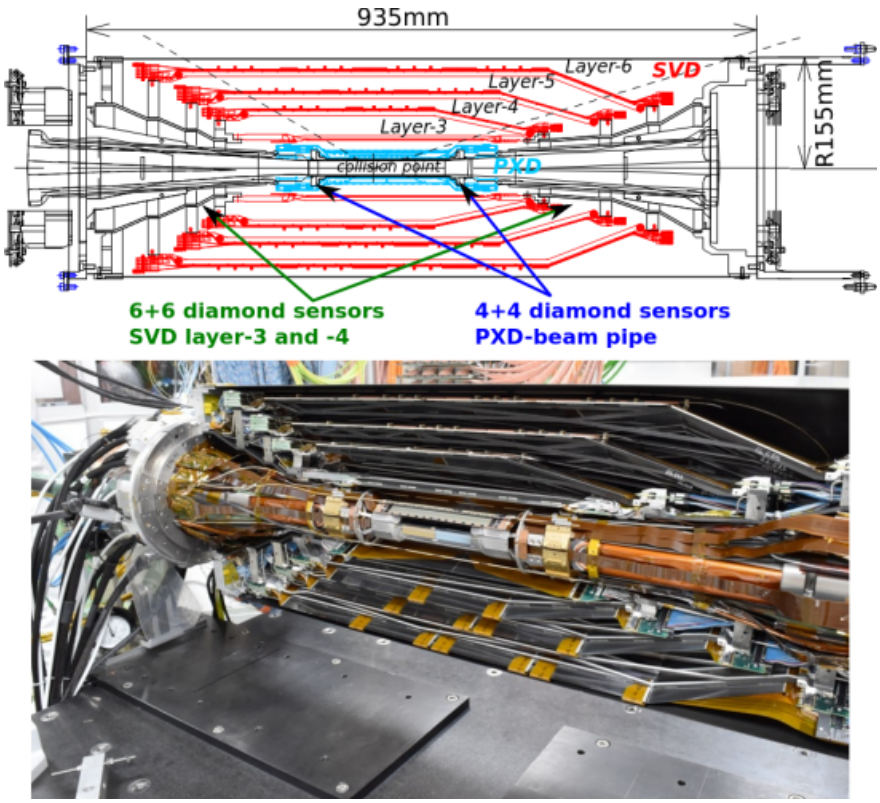


Figure 1.4: BelleII vertex detector

are only used to improve spatial resolution of tracks: the SVD informations are used by the High Level Trigger to look for regions of interest in the pixel ladders just by extrapolating back the tracks found in the tracker detector, and this method allows to store only data belonging to these areas; the PXD hits are then used in offline track fit to improve the vertex resolution.

CMOS MAPS have been proposed for the replacement of VXD during the Long Shut Down 2 foreseen around 2026-27; the new vertex detector, VTX, should be made of 5 layers fabricated with the optimized Belle II pixel sensor (OBELIX), a detector based on TJ-Monopix (see at chapter 3). The main advantages VTX should bring are a significant improvment in the track and vertex resolution (14 µm before upgrade, $\lesssim$10 µm expected after upgrade), a reduction in the material budget, a higher background tolerance because of the smaller sensor than strips dimension.

1.2 Other applications

Pixel detectors are widely used also for photon detection: they can be used as single photon counter or integrating and collecting the charge released by more impinging particles. The utilisation in the first case is similar to the tracking one, except that the requirements are less tight, so much that two noteworthy of microchips originally meant for detectors in particle physics at the LHC, and later employed in other fields are Medipix and Timepix. They are read-out chips developed by the Medipix Collaborations since early 1990s. For two decades, different Medipix generations have been produced, having a rough correlation

with the feature size used: for example, Medipix2 (1999) used 250 nm feature size CMOS while Medipix3 (2005) 130 nm. For photons imaging other materials with higher atomic charge than silicon could be prefered, as a high photon absorption efficiency is needed: it was for this reason that Medipix2 was bump bonded to identically segmented sensors of both Si and GaAs.[7]

The applications in scientific imaging vary from astrophysics and medical imaging to more exotic domains as studies of protein dynamics, material science, art authentication and archaeology. One of the most important employment of Medipix is as X-ray single photon counting in industrial and medical radiography and in 3D computed tomography[2]. Regarding the latter, for example, thanks to a New-Zealand company, the MARS Bioimaging detector has been fabricated, which is capable of resolving the photons energy and produce 3D coloured images. Besides tracking in HEP (I have already cited the use of Timepix3 is in the beam telescope of the LHCb VELO), an important use of Timepix is in dosimetry. A small-Timepix detector with the dimension of a USB can also be found at the International Space Station, where it is exploited for radiation, principally made of haevy-ion, monitoring.

1.2.1 Applicability to FLASH radiotherapy

A possible new application of pixels detector is dosimetry or beam monitoring of charge particles in high intensity radiotherapy (RT). Recently[3] a promising method for radiotherapy at ultra high dose rate (at least 40 Gy/s) and for this reason called FLASH-RT[8], instead of CONV-RT (0.03 Gy/s), came out. However, finding dosimeters, as well as beam monitor, suitable at ultra high dose rate is still an open issue since almost all standard online dosimeters have shown saturation problems. In table 1.1 are listed the typical value of operation which distinguish between CONV-RT and FLASH-RT.

Radiotherapy

The radiological treatment is a common method used in 60% of tumors both as palliative care and as treatment. It can be given before, after or during a surgery, (Intra operative radiation therapy-IORT) and many different types of radiations (photons, electrons, protons and ions, which mainly are carbon ions) can be used to irradiate the affected tissues. Exploiting the ionizating energy loss, that can be parametrized by the Linear Energy Transfer (LET), a biological damage can be delivered to the tissue: while protons, α particles and light ions are high LET radiations with values in range 100 keV/µm to 200 keV/µm, x-rays and gamma-rays are low LET radiations with values in range 0.2 keV/µm to 2 keV/µm. If x-ray photons, with energy in 4-25 MeV are used, the ionization is caused by the Compton electrons and pair production, and is more in the superficial layers of the tissue due to the

[2]The analysis of the direction dependence of X-ray absorption is perform, for example, in order to obtain an image in Computed Tomography (CT)

[3]The first evidences has been observed on a mice experiments in 1966 and in 2014 by the group of Favaudon and Vozenin. After this, many test on cats and pigs have been performed, and also there has been a clinical trial on a cutaneous tumor-patient

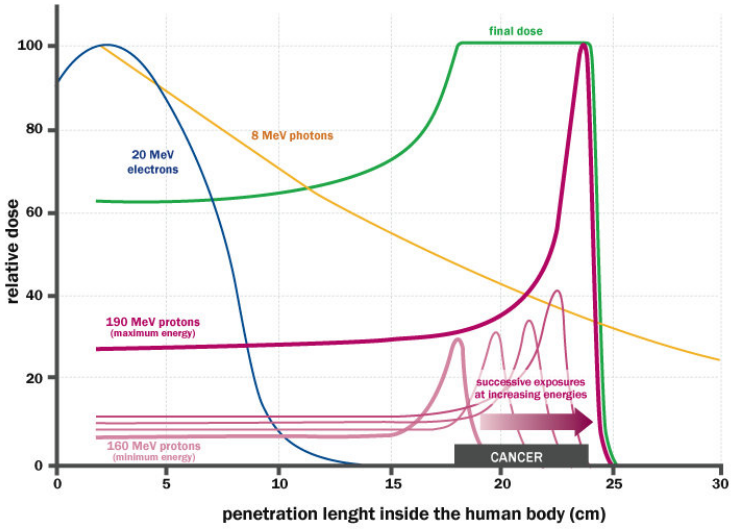


Figure 1.5: The Spread Out Bragg Peak (SOBP) curve (green), which is a constant dose distribution, is obtained from the superposition of many Bragg peak of hadrons with different energy.

	CONV-RT	FLASH-RT
Dose rate	0.03 Gy/s	40 Gy/s
Intra pulse dose rate	100 Gy/s	106 Gy/s
Treatment duration	∼minutes	$\lessapprox$500 ms
Dose Per Pulse	0.3 Gy	1 Gy to 10 Gy
Pulse width	3 µs	∼2 µs

Table 1.1: Typical range of values used in a CONV radiotherapeutic treatement made with pulsed beam, compared to the range in which the FLASH effect has been observed.

exponential attenuation of the beam. The heavy charged particles energy loss, instead, is strongly localized in the final region of the track, that is the Bragg peak, such as the the treatement typically requires the scanning of the target.

Electrons, instead, of energy in range of a dozen of MeV tend to spread out on a bigger region of a few centimeters in both the diameter and thickness and for this reason using Very High Energy Electrons (VHEE) has been taken into account for irradiation of deeper tissues.

FLASH effect

This treatment takes advantages of biological differences between tumors and healthy tissues: it is characterized by reducing normal tissue toxicity and maintaining equivalent tumor damage. The response to dose can be described by the survival fraction probability,

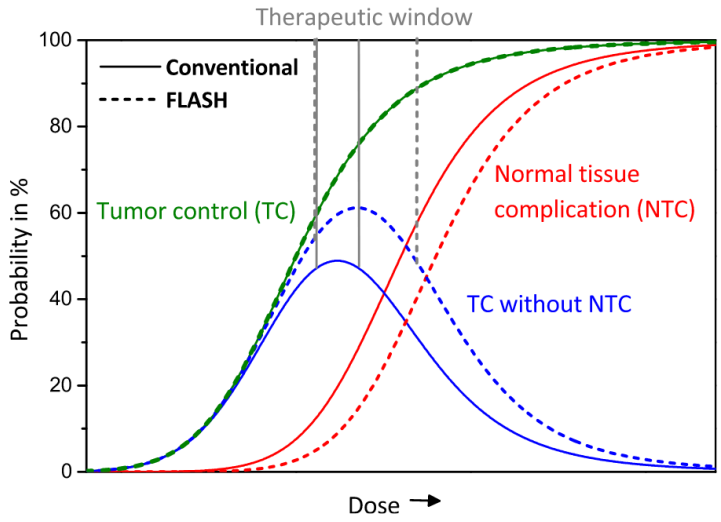


Figure 1.6: Illustration of dependence of Tumor Control Probability (TCP), Normal Tissue Contral Probability (NTCP) and therapeutic window on dose, for CONV-RT ad FLASH-RT.

describing the fraction of surviving cell as a function of the dose:

$$S(D) = S(0)\, e^{-(\alpha D + \beta D^2)} \tag{1.1}$$

where α and β respectively represents the rate of cell killing by single ionizing events and by double hits. Hence, at high doses the density of damages increases and the cells repair becomes more difficult. The FLASH effect is not yet completelly understood and the underlying mechanisms are not yet clear, but many ideas have been proposed; among the most debated there are the fact that the high dose rate would prevent cells from reparing, or that the cell oxygen content and the immune response play a role. However what has been noticed is that the FLASH-RT allows to widen the therapeutic window, defined as below. The Tumor Control Probability (TCP) and the Normal Tissue Complication (NTC) functions parametrize respectively the efficiency of damaging on the tumor after having released a certain dose and the probability of affecting the healthy tissues. The intermediate zone between the increase of the TC and of the NTC is called therapeutic window, and the wider it is and the more effective the treatment.

Dosimetric problems

Since the dosimeters problems are strictly related with the time structure of the beam used for the treatement, let's consider the case of a pulsed electron beam (fig.1.7), where a high quantity of dose ($\gtrsim$1 Gy/pulse) is released in a very short time (macropulses). At small macro-pulse frequency the signal collection time is shorter than the time between pulses, and, consequently, the saturation is influenced only by the Dose Per Pulse (DPP).

Up to now, all online dosimeters have shown saturation problems at high DPP, differently from radiochromic films, which are the standard passive dosimeters and have shown

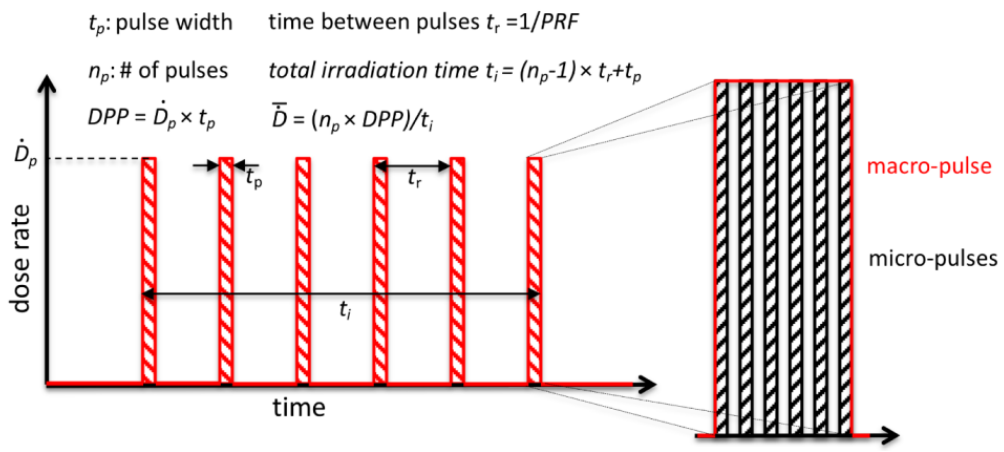


Figure 1.7: Typical beam structure of a beam used in electron radiotherapy

a dose-rate indipendece up to 100 Gy/s. In spite of the response is linear in a wide dynamic range, they do not provide any online dosimetric informations, since the time required to extract the physical value from the reading is not long, requiring a measurement of the film density[4].

Ionization Chambers (ICs), which are the online reference dosimeter also according to law protocols, at high level of radiation (already at dose per pulse two orders of magnitude lower than the ones used for FLASH-RT) show both problems of saturation and recombination. When a high density of ions and electrons is produced in the gas, a high counter electric field opposed to the drift one might be generated; if a neutral region build up, both the recombination of i/e pairs, with a subsequent photoemission and abrupt discharge can happen. A correction factors, k_{sat}, can be introduced for sufficient low level of radiation and in this case a precise dose measurement can still be done: under conventional operation, with Dose Per Pulse lower than 1 mGy the correction factor is $<5\%$.

Concerning the conventional semiconductor, if exposed to high dose rate, they suffer of saturation problems just as the ICs, but the development of fast MAPS devices, with rate capability of 100 MHz/cm^2 or more, might open the possibility of using these detectors as dosimeters at high dose rates. The idea is to use the high readout speed to split the dose per pulse in many buckets to reduce the saturation effect. Indeed a thin planar sensor could allow for the preservation of an enough strong electric field even at high dose rate which, together with the short mean path the e/h must cover to get the electrode (the epitaxial layer typically is $\sim$30 µm), could result in a fast collection of electrons by drift and in a non-saturated response. Besides, also the small capacity typical of MAPS is beneficial for reducing the readout time: a reduced C_d enables for a fast discharge and then for a fast readout. Finally, MAPS devices would provide good time and space resolutions compared to other dosimeter technologies, and since they can be thinned down to about 50 µm, could also be employed in monitoring the beam position with minimal disturbance.

Among other detectors, optical fiber and alanine dosimeters have been proposed for high dose environments and many groups are going on studies on their applicability on

[4]The radiation produces a polymerization of an active layer, resulting in a different density of coloration.

Commercial detector	Detector type	saturation [Gy/p]
PTW TW34045 Advanced Markus EC	ionization chamber	0.3
PTW TM60017 Dosimetry Diode E	silicon diode	0.15
PTW TW60019 microDiamond	diamond	0.15
DoseVue DoseWireTM Series 100	scintillator fiber	11-26

Table 1.2: Results obtain in [9]

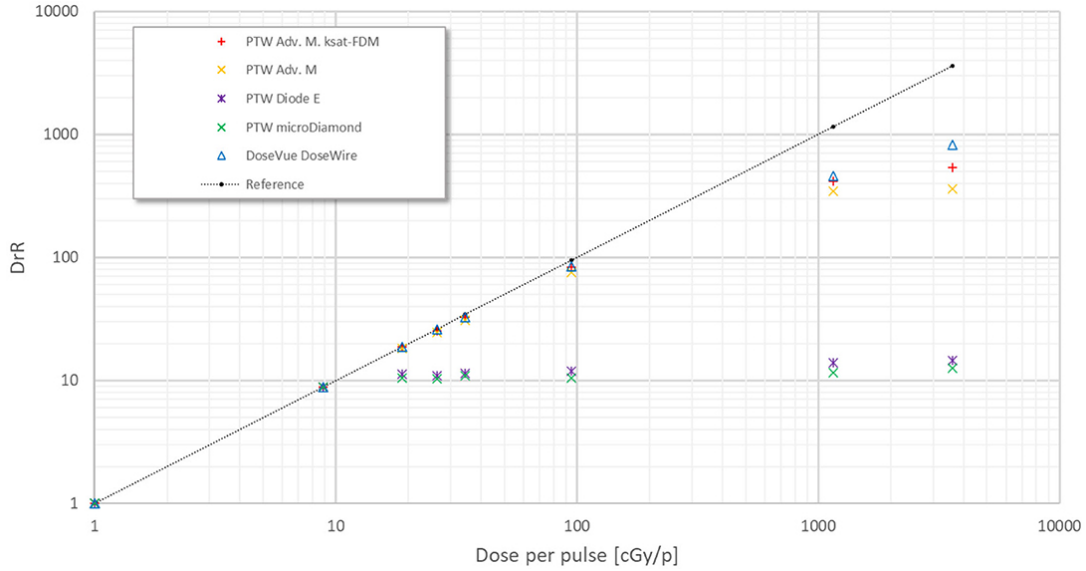


Figure 1.8: Saturation problems underlyed in [9].

FLASH-RT. In reference [9] are presented some results related with saturation problems at high DDP of different types of detectors. The dosimeters tested and their value at which saturation becomes are reported in table 1.2, while in figure 1.8 are reported the measurements. The DrR is then defined as the ratio between the signal response of each dosimeters at a fixed DPP of 1 cGy/p and is called dosimeter reading ratio (DrR):

$$DrR = \frac{R}{R|_{Dp=1cGy/p}} = \frac{R}{R_{ref}} \tag{1.2}$$

Then, for a saturation not affected by saturation problems DrR should be a straight line with a slope equal to 1; since the reference measurements of dose has been performed with radiochromic films (GAFCHROMIC EBT-XD), whose dose indipendece has been tested[5] in range from 0 Gy to 15 Gy, the dotted black line in figure 1.8 represents the reference measurment done with it.

[5]The radiochromic films calibration has been obtained by irradiating the films with dose values in range from 0.01 Gy to 15 Gy, by positioning the films in a polymethylmethacrylate (PMMA) phantom at R_{100} depth, corresponding to 10 cm

CHAPTER 2

Pixel detectors

Pixel detectors are semiconductor detectors which are segmented in two dimensions: this distinguishes them from the strips, such that a single plane of detector already provides both the coordinates of impact of the detected particle. Their operation is based on the p-n junction (fig.2.1).

A p-n junction is built by bringing in contact two n and p doped silicon crystals. In an n-doped crystal some silicon atoms are replaced with valence 5 donor atoms (such as P) which provide loosely bound e^- carriers, while in p-doped crystals some atoms are replaced by valence 3 acceptor (such as B) which absorb existing free electrons, effectively creating a positively charged carrier called "hole". At the boundary, recombination of opposite charge carriers occurs, forming a region, the depletion zone, which is free of charge carriers. The charged donors$^+$ and acceptor$^-$ atoms, that remain ionized in the n-type and p-type regions, constitute a space charge and create an electric field across the junction, causing a drift current in the opposite direction to the diffusion one, through which the junction reaches an equilibrium state. Assuming a constant space change, the electric field is linear and reach a maximum at the boundary of the p and n layers.

2.1 Signal formation

When a charged particle passes through a semiconductor and loses energy by ionization, only a part of that energy is used to generate electron-hole pairs, since another part is used for other processes, as lattice excitation. The average energy needed to create a pair at 300 K in silicon is $w_i = 3.65\,\text{eV}$, that is more than the mean ionization energy because of the interactions with phonons. For a minimum ionizing particle (MIP) the most probable value (MPV) of charge released in the semiconductor is $0.28\,\text{keV}/\mu\text{m}$, hence the number of electrons-hole pairs is:

$$\langle \frac{dE}{dx} \rangle \frac{1}{w_i} \sim 80\,\frac{e/h}{\mu\text{m}} \sim \frac{1.28\,10^{-2}\text{fC}}{\mu\text{m}} \tag{2.1}$$

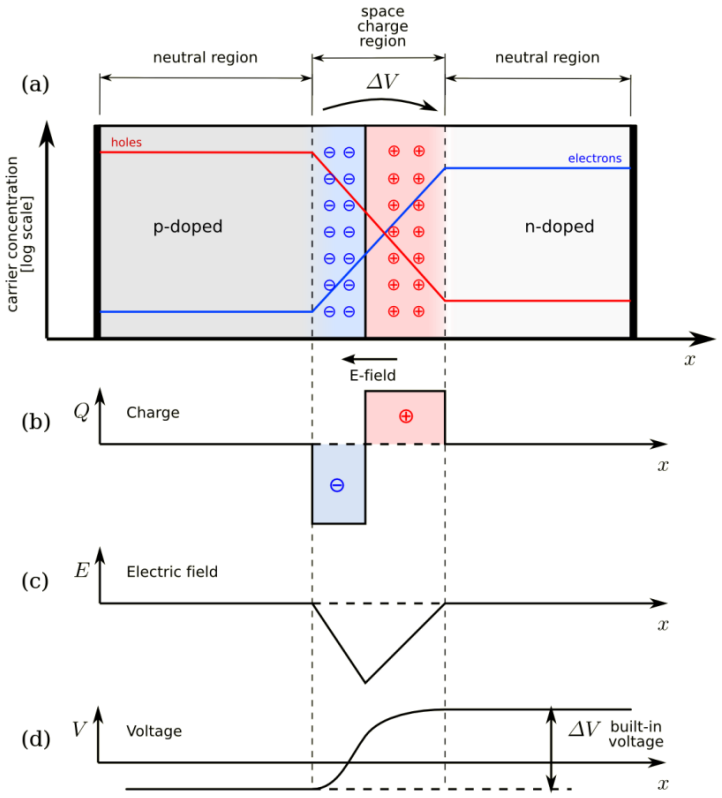


Figure 2.1: The structure of a p-n junction. (a) structure, (b) space charge density, (c) electric field distribution and (d) potential distribution.

The number $N_{e/h}$ of couples generated undergoes fluctuations that usually follow a Poisson distribution; thus the fluctuations of $N_{e/h}$ is equal to $\sigma_{e/h} = \sqrt{N_{e/h}}$. Since the energy loss is not a purely statistical fact, because of the energy the particle can lose is obviously $\leqslant$E and the energy needed for ionization must be $\geqslant$ of a minimum value, the resolution actually is lower by a factor $\sqrt{F}$, where F is called the Fano factor. F is a function of the material and temperature and for silicon is equal to $\sim$0.115.

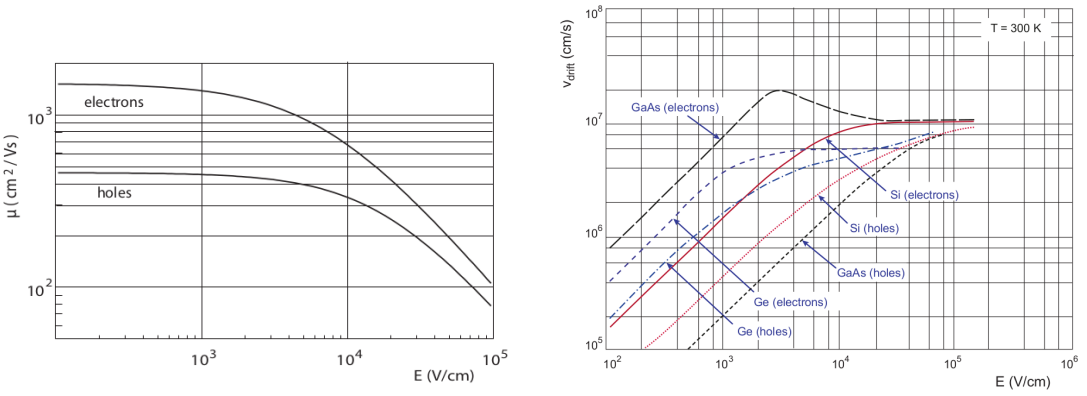
In order to avoid a signal loss, pairs e/h must be produced in the depleted region of the semiconductor, where the probability of recombination with charge carriers is low. For this reason pixel detectors are commonly reverse biased: a positive bias is given to the n electrode and a negative to the p in order to increase the depletion zone. The width of the depletion region depends on the external bias V_{ext}, the resistivity ρ and also with the dopant:

$$d_n \sim 0.55\sqrt{\frac{\rho}{\Omega cm}\frac{V_{ext}}{V}}\mu m \qquad d_p \sim 0.32\sqrt{\frac{\rho}{\Omega cm}\frac{V_{ext}}{V}}\mu m \tag{2.2}$$

Thus, high resistivity wafers ($100\ \Omega cm - k\Omega cm$) are typically preferred because they allow a thicker depletion zone with smaller bias voltage.

The charges created within the sensor are separated by the electric field and collected at their respective electrodes (p for holes and n for electrons)[1]. A signal i_e is generated on

[1]Even if in principle both the electrodes can be used to read the signal, for pixel detectors, where the number of channel and the complexity of readout are high, only one is actually used. In strip and pad



(a) Dependence of the mobility on the electric field.

(b) Drift velocity at room temperature in different semiconductors

Figure 2.2

the electrode e by the drift of these charges, as stated by the Shockley-Ramo's theorem:

$$i_e(t) = -q\, v(t)\, E_{WF,e} \tag{2.3}$$

where $v(t)$ is the instantaneous velocity of the charge q and E_{WF} is the weighting field, that is the field obtained biasing the electrode e with 1V and all the others with 0V. The drift velocity of the charge depends on the electric field and on the mobility of the particle:

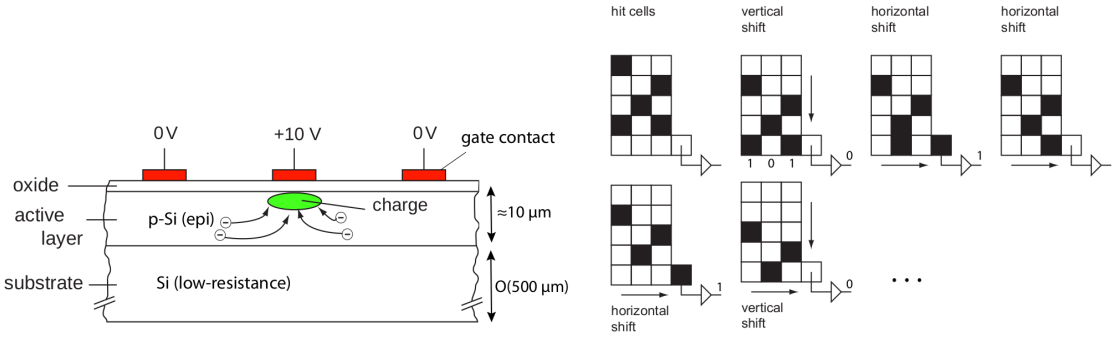
$$v = \mu(E)\, E \tag{2.4}$$

where $\mu(E)$ is a function of the electric field and is linear in E only for small E: at higher values the probability of interactions with optical phonons increases, the mobility drops and this leads to a saturation of the velocity (fig.2.2). Typical values for electrons and holes mobility in silicon at room temperature are $\mu_n \sim 1450\,\mathrm{cm}^2/\mathrm{Vs}$, $\mu_h = 500\,\mathrm{cm}^2/\mathrm{Vs}$.

2.2 Charge Coupled Devices

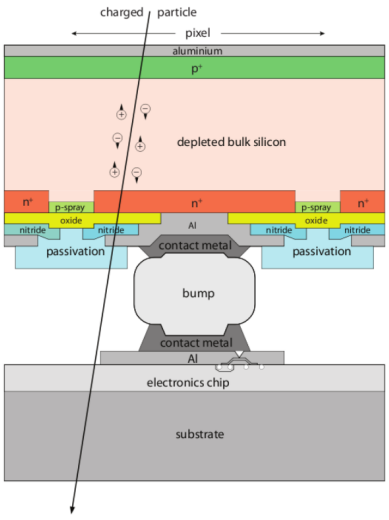
Charge Coupled Devices (fig.2.3a) are one of the first pixel detectors initially developed to detect visible light and then adapted for charged particles and x-rays. In CCDs the charge is created in a very thin active epitaxial layer (typically 10 µm, maximally about 30 µm) and then locally stored in a potential minimum which is created by a MOS structure. The size of the CCD cells is typically in the range 10 µm to 20 µm such that spatial resolutions are of the order of a few micrometers. The collected charges are moved stepwise from electrode to electrode towards the readout note (thus so called 'bucket chain') by applying a potential with a clock with frequency of ∼MHz; the readout chain (fig.2.3b) is completely sequential with only one readout node and this makes the entire process comparatively slow (tens of ms), despite of such high frequency. A particular type of CCD, the pnCCDs, are typically used to detect low energy (<10 keV) x-ray photons for their homogeneous spatial detection efficiency of photons. The pnCCDs have a sideward depletion similar to silicon

detectors, instead, is more common a dual-side readout

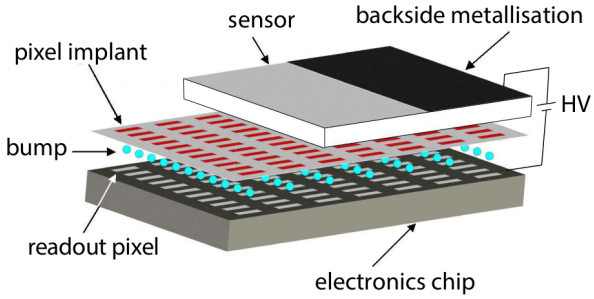


(a) (b)

Figure 2.3: (a) CCD cross section scheme. (b) Readout of a CCD matrix: the charge is moved stepwise in vertical and horizontal directions into the readout node



(a) Hybrid cross section scheme



(b) Hybrid matrix scheme

drift chambers that makes the electric field stronger, compared with the normal CCDs. The pnCCDs designed for photon imaging are often fabricated with high Z materials, to increase absorption efficacy.

2.3 Hybrid pixels

In hybrid pixels, which currently are the state-of-art technology for large scale pixel detectors in most particle physics experiments, the sensors and the electronics are realized on separate wafers and connected together a bump of conductive material, typically In or Sn (fig.**??**). They provide a practical system where the sensor and the ASIC (application specific integrated circuit) can be optimized separately, allowing the fabrication of radiation-hard devices capable of operating at GHz rates. However hybrid pixel have also some disadvantages: the bump-bonding interconnection technology is expensive and delicate; the separate substrates for electronics and sensor lead to an increase in the material budget; the pixel dimension must be long enough for the bump-bonding technology, with a current limit of about 50 µm.

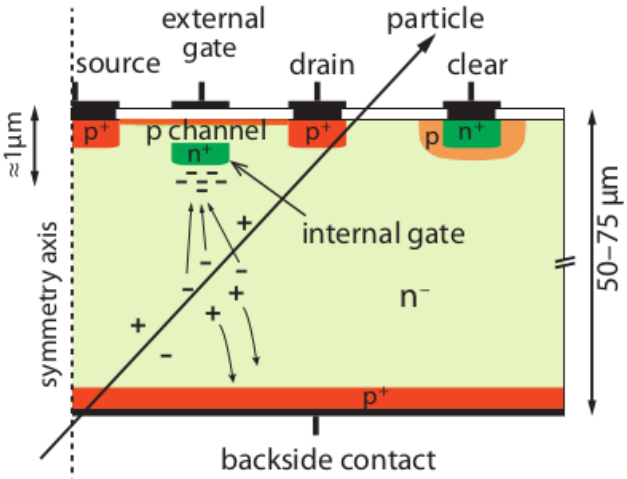


Figure 2.5: DEPFET cross section scheme

2.4 DEPFET

DEPFETs are the an attempt towards the integration of the front end (FE) on the same substrate of the sensor. Each pixel implements a MOSFET (metal-oxide-semiconductor field-effect transistor) transistor (a p-channel in fig. 2.5): the hole current flows from source to drain is controlled by the external gate and the internal gate together. The internal gate is made by a deep $n+$ implant towards which electrons drift after being created in the depletion region; the accumulation of electrons in the region underneath the n implant changes the gate potential and controls the transistor current, resulting in an internal amplification, the removal of the signal charge from the internal gate is called "Clear". DEPFET typically have a good S/N ratio: thanks to the on-pixel amplification, to the thick depletion region. As in CCDs, DEPFET require a serial readout of the pixel signal, and are therefore relatively slow devices, but they can be made very think (50 µm). In recent years, the sensor development was driven by an intensive R&D and prototyping for x-ray imaging and the ILC vertex detector.

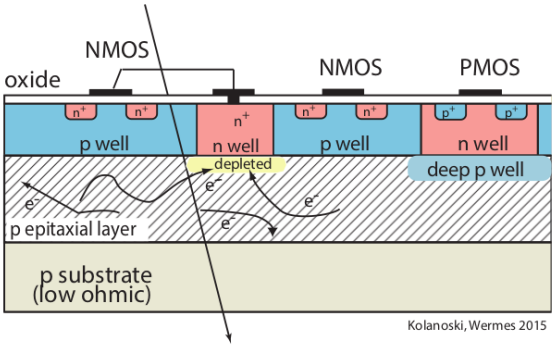


Figure 2.6: MAPS cross section scheme

2.5 CMOS MAPS and DMAPS

Monolithic active pixels (fig 2.6) accommodate on the same wafer both the sensor and the FE electronics, with the second one implanted on top within a depth of about 1 µm below the surface. MAPS have been first proposed and realized in the 1990s and their practical usage has been enabled by the development of the consumer electronics sector, which guarantees the halving of CMOS transistors dimension at least every two years, as stated by the Moore's law. As a matter of fact the dimension of components, their organization on the pixel area and logic density are important issues for the design and for the layout. Compared to CCDs, the readout time is dramatically reduced by the in-pixel amplification and discrimination, typically followed by a sparsified readout selecting only over threshold pixels and avoiding the need for signal transportation over thousands of pixels.

A critical parameter for accelerator experiments is the material budget, which represents the main limit factor for momentum measurement resolution in a magnetic field; since hybrid pixels are thicker ($\sim$ hundreds of µm) than monolithic ones (even less than 100 µm). Using the latter the material budget can be reduced to a third: typical values for hybrid pixels is 1.5 % X_0 per layer, while for monolithic 0.5 % X_0. Compared to MAPS, among other disadvantages of hybrid pixels there is the bigger power consumption, that requires also a bigger cooling system, leading to a futher increase of material. On the other hand MAPS are still in the development phase, and although they have been used in several experiments as discussed in chapter 1, their potential remains to be fully exploited.

Monolithic active pixel can be distinguished between two main categories: MAPS and depleted MAPS (DMAPS). In the initial CMOS MAPS (2.6) an unmodified CMOS process was used that presented the full depletion of the 1 µm to 20 µm epitaxial layer.

The charge is mainly collected by diffusion rather than by drift, making the path of created charges in the bulk longer resulting in relatively slow collection (of order of 100 ns). Moreover, the collection can be partial, especially after irradiation of the detector, when the trapping probability becomes higher. In DMAPS instead, a modified process is employed allowing the creation of a much thicker depletion layer, thus increasing significantly the collected charges. This requires the addition of deep implanted areas, as shown in figure 2.7. The charge released in the epi layer is very small (few thousands of electrons) but the extremely small electrode capacitance allows the formation of a detectable voltage signal V=Q/C.

In figure 2.6 it is shown as example of CMOS MAPS: the sensor implements an n well as collection diode; to prevent the others n wells (which contain PMOS transistor) of the electronic circuit competing in charge collection and to shield the CMOS circuit from the substrate, additional underlying deep p well are needed.

2.5.1 DMAPS: large and small fill factor

There are two different sensor-design approaches (figure 2.7) to DMAPS:

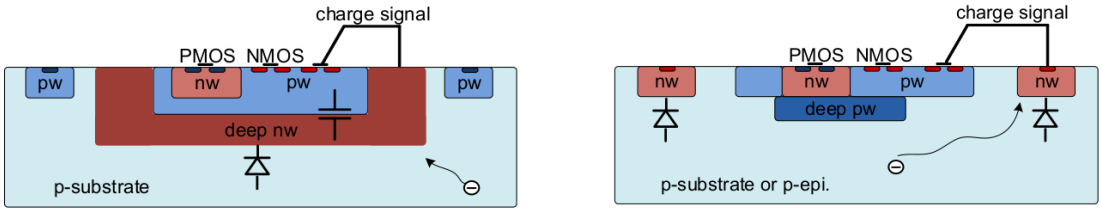


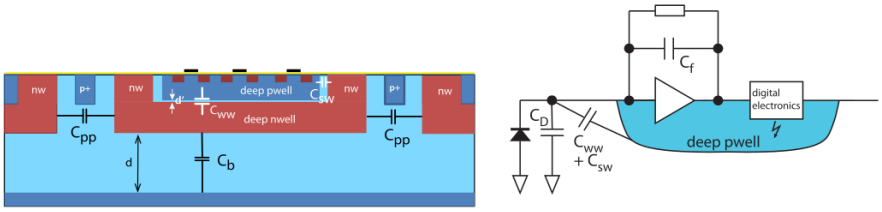
Figure 2.7: Concept cross-section of CMOS MAPS with large and small fill factor

- large fill factor: a large collection electrode that is a large deep n-well and that host the embedded electronics
- small fill factor: a small n-well is used as charge collection node

To implement a uniform and stronger electric field, DMAPS often uses large electrode design that requires multiple wells (typically four including deep n and p wells); with this layout the total capacity of the sensor increases because of the addition of a new term (fig. 2.8), which contributes to the total amplifier input capacity (∼100 fF). In addition to the capacity between pixels (C_{pp}) and between the pixel and the backside (C_b), a non-negligible contribution comes from the capacities between wells (C_{SW} and C_{WW}) needed to shield the embedded electronics. These capacities affect the thermal and 1/f noise of the charge amplifier and the τ_{CSA} too:

$$ENC^2_{thermal} \propto \frac{4}{3}\frac{kT}{g_m}\frac{C_D^2}{\tau_{sh}} \qquad \tau_{CSA} \propto \frac{1}{g_m}\frac{C_D}{C_f} \tag{2.5}$$

where g_m is the transconductance, τ_{sh} is the shaping time. Among the disadvantages coming from this large input capacity there is a coupling between the sensor and the electronics resulting in cross talk noise on neighbouring electrodes; indeed, since digital switching in the FE electronics does a lot of oscillations, this problem is especially connected with the intra-wells capacities. So, larger charge collection electrode sensors provide a uniform elec-



(a) (b)

Figure 2.8: DMAPS scheme with shown the capacity terms (C_{pp}, C_b, C_{WW}, C_{SW}) which contribute to the total detector capacity C_D

tric field in the bulk that results in short drift path and so in good collection properties, especially after irradiation, when trapping probability can become an issue.

The small fill-factor variant, instead, benefits from a small capacity (5 fF to 20 fF), but suffers from a non uniform electric field and from all the issue related to that (slowness and high trapping probability). As we will see, these two different types of sensor require

	small fill factor	large fill factor
small sensor C	$\surd$ ($< 5\,$fF)	$\times$ ($\sim$ 100 200fF)
low noise	$\surd$	$\times$
low cross talk	$\surd$	$\times$
velocity performances	$\surd$	$\times$ ($\sim$100 ns)
short drift paths	$\times$	$\surd$
radiation hard	$\times$	$\surd$

Table 2.1: Small and large fill factor DMAPS characteristics

different amplifier: the large electrode one is coupled with a charge sensitive amplifier, while the small one with a voltage amplifier (sec.2.6).

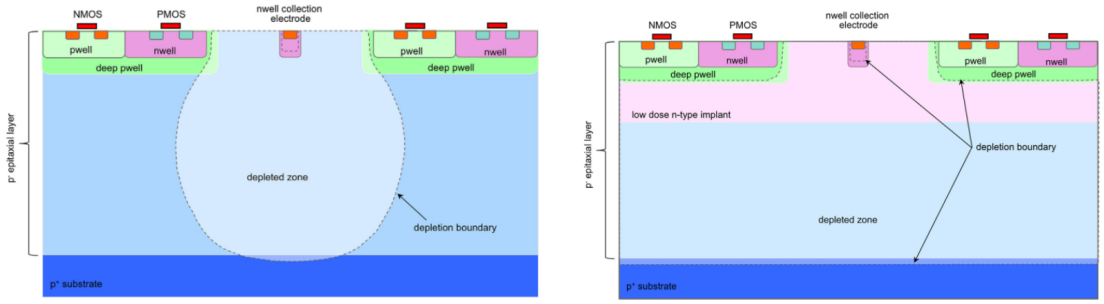
2.5.2 A modified sensor

A process modification, developed by CERN in collaboration with the foundries, which has become the standard solution to combine the characteristics of a small fill factor sensor (small input amplifier capacity) and of a large fill factor sensor (uniform electric field), is the one carried out for the ALICE upgrade for about ten years [10]. A compromise between the two sensors could also be making smaller pixels, but this solution requires reducing the electronic circuit area, so a completely new pixel layout should be thought. The modification consists in inserting a low dose implant under the electrode and one of its advantage lies in its versatility: in fact, not requiring special preliminary adjustments in the sensor and in the electronics, typically both standard and modified sensors are often produced for testing.

Before the process modification, the depletion region extends below the diode towards the substrate, and it does not extend much laterally, even if a high bias is applied to the sensor (fig. 2.9). After the modification, two distinct pn junctions are built: one between the deep p well and the n^- layer, and the other between the n^- and the p^- epitaxial layer, extending to the whole area of the sensor. Since deep p well and the p-substrate are separated by the depletion region, the two p electrodes can be biased separately and this is beneficial to enhance the vertical electric field component. The doping concentration is an optimization parameter: it must be high enough to be larger than that in the epitaxial layer to prevent the punch through between p-well and the substrate, but it must also be low enough to allow the depletion for reasonable bias values.

2.6 Analog front end

After the collection on the electrode, the signal enters the front end amplification circuit (fig.2.10), ready to be shaped and transmitted out of chip. Low noise amplification, fast hit discrimination and an efficient, high-speed readout architecture, consuming as low power as possible, are the goal of the readout integrated electronics (ROIC). The



(a) (b)

Figure 2.9: A modified process for ALICE tracker detector: a low dose n implant is used to create a planar junction. In (a) it is difficult to deplete the epitaxial layer over its full width (b) the pixel is fully depleted.

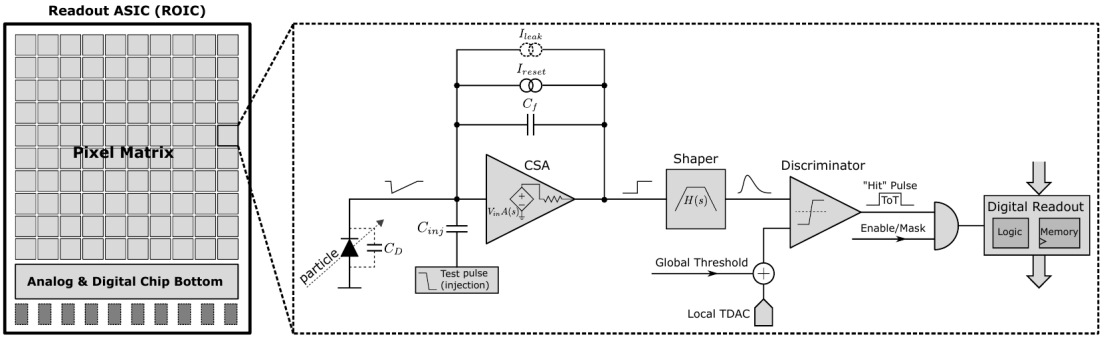


Figure 2.10: Readout FE scheme: in this example the preamplifier is a charge sensitive one (CSA) but changing the capacitive feedback into a resistive one, this can be converted in a voltage or current amplifier.

first generation of pixel ROIC chip were typically designed in 0.25 µm CMOS process and the main part of the processing were done in the chip periphery; over the years the technological development has allowed the decrease in the dimension of the transistors, allowing the usage of 0.13 µm technology. The third generation of ROIC developed in 65 nm is going on in order to meet the challenges of the HL-LHC. The main parts of the analog front end chain are a preamplifier (that often is the only amplification stage) with a reset to the baseline mechanism (often realized by a constant current reset circuit) and a leakage current compensation, a shaper (a band-pass filter) and finally a discriminator. The whole chain must be optimized and tuned to improve the S/N ratio. It is very important both not to have a large noise before the amplification stage to avoid amplifying that noise, and to choose a reasonable threshold of the discriminator to cut noise-hits much as possible.

Depending on whether a capacity or a resistance is used as feedback, respectively a charge or a voltage amplifier is used: if the voltage input signal is large enough and has a sharp rise time, the voltage sensitive preamplifier is preferred. Consequently, this flavor does not work for large fill factor MAPS whose signal is small: $v_{in} = Q/C_D \approx 3\,\text{fC}/100\,\text{pF} = 0.03\,\text{mV}$, but it's fine for the small fill factor ones: $v_{in} = Q/C_D \approx 3\,\text{fC}/3\,\text{pF} = 1\,\text{mV}$.

In the case of a resistor feedback, if the signal duration is longer than the discharge time ($\tau = R_S C_D$) of the detector the system works as current amplifier, as the signal is immediately transmitted to the amplifier; in the complementary case, if the signal duration longer than the discharge time, the system integrates the current on the C_D and operates as a voltage amplifier.

2.7 Readout logic

The readout logic includes the part of the circuit which takes the FE output signal, processes it and then transmit it out of pixel and/or out of chip; depending on the situation of usage different readout characteristics must be provided. To store the *analogical* information (i.e. charge collected, evolution of signal in time, ...) big buffers and a large bandwidth are needed; the problem that does not occur, or better occur only with really high rate, if one wants record only *digital* data (if one pixel is hit 1 is recorded, and if not 0 is recorded).

A compromise is to store the Time Over Threshold (ToT) of the pulse in clock cycle counts; this needs of relatively coarse requirement as the ToT can be trimmed down to use only a dozen bits but, being correlated with the deposited charge, it can provide a sufficient information. The ToT digitalization usually takes advantage of the distribution of a clock (namely BCID, bunch crossing identification) on the pixels' matrix. The typical required timing precision is better than ~25 ns, that corresponds to the period between bunch collisions at LHC; for such reason a reasonable BCID-clock frequency for pixels detector is 40 MHz.

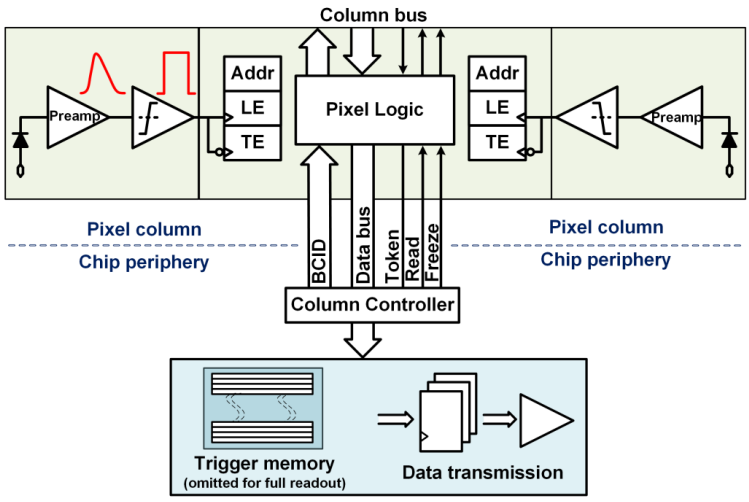


Figure 2.11: Column drain R/O scheme where ToT is saved

Moreover, the readout architecture can be *full*, if every hit is read, or *triggered*, if a trigger system decides if the hit should be stored or not. On one hand the triggered-readout needs buffers and storage memories, on the other the full readout needs an high enough bandwidth. A triggered readout is fundamental in accelerator experiments where the quantity of data to store is very large and some selection has to be applied by the

trigger: to give an order of magnitude, at LHC more than 100 TBit/s of data are produced, but the storage limit is about 100 MBit/s [11]. Typically, the trigger signal is processed in a few µs, so the pixel gets it only after a hundred clock cycles from the hit arrival time: the buffer depth must be able to handle such high trigger latency.

After having taken out the data from the pixel, it has to be transmitted to the End of Column (EoC) where one ore more serializers deliver it out of chip, typically to an FPGA. There are several ways of transmitting data from a pixel to the EoC, but two macro-categories can be distinguished: readout can be based on a *rolling shutter* logic or can be *sparsified*. In the former case, the matrix is divided in sections and in turn each part is selected, freezed and a rapid scan on the pixels draws the data out; since also pixels without hits are read, this readout results to be slower compared to the sparsified one. In this case the discrimation on pixel is used to select for the readout only the pixel which store hits. The sparsified readout need of a priority logic which establishes the pixels readout order across the matrix; among the possibilities a common priority logic is the column-drain (fig.2.11), developed for CMS and ATLAS experiments [12]. All the pixels in a double-column share a data bus and only one pixel at a time, according to a priority chain, can be read. The reading order circuit is implemented by shift register: when a hit arrives, the corresponding data, which might be made of timestamp and other information such as the ToT for example, is temporarily stored on a RAM until the shift register allows the access to memory by data bus. Even if many readout architectures are based on the column-drain one, it does not work for large size matrices. The problem is the increasing number of pixels on a column would also raise the number of pixels in the priority chain, which would result in a slowdown of the readout.

If there is not any *storage memory*, the double-column behaves as a single server queue and the probability for a pixel of waiting a time T greater than t, with an input hit rate on the column μ and an output bandwidth B_W is [13]:

$$P(T > t) = \frac{\mu}{B_W} e^{-(B_W - \mu)t} \tag{2.6}$$

To avoid hit loss (let's neglect the contribution to the inefficiency of the dead time τ due to the analog Front End), for example imposing $P_T > t \sim 0.001$, one obtains $(B_W - \mu)\, t_t \sim 6$, where t_t is the time needed to transfer the hit; since t_t is small, one must have $B_W \gg \mu$, that means a high bandwidth [13].

Eq.2.6 is actually an approximation, since each pixel sees a different bandwidth depending on the position on the queue: the first one sees the full bandwidth, while the next sees a smaller one because it can be occasionally blocked by the previous pixel. Then, the bandwidth seen by the pixel i is $B_i = B - \sum_j \mu_j$, where μ_j is the hit rate of the $j-$th pixel. The efficiency requirement on the bandwidth and the hit rate becomes: $B_{W,i} > \mu_i$, where the index i means that the constraint is for a single pixel; if all the N pixels on a column have the same rate $\mu = N\mu_i$, the condition reduces to $B_W > \mu$. The bandwidth must be chosen such that the mean time between hits of the last pixel in the readout chain is bigger than that. In order to reduce the bandwidth, a readout with zero suppression on pixel is typically employed; this means that only information from channels where the

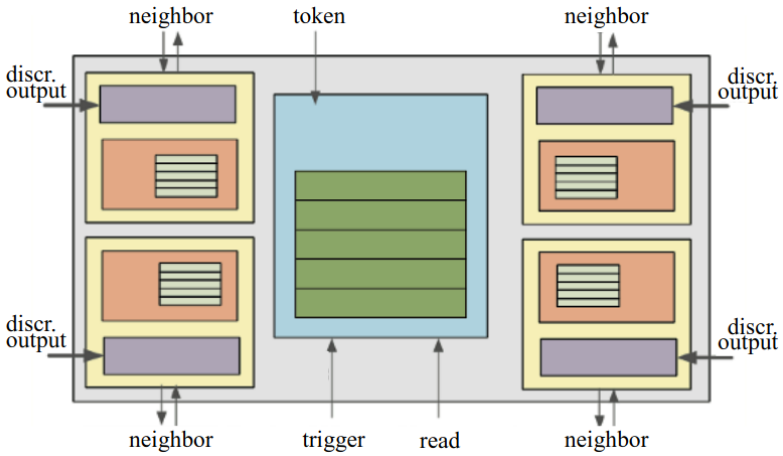


Figure 2.12: Block diagram of the FE-I4 R/O. Read and memory management section is highlighted in light blue; latency counters and trigger management section are highlighted in green; hit processing blocks are highlighted in purple; ToT counters and ToT management units are highlighted in orange

signal exceeds the discriminator threshold are stored.

If, instead, the signal is locally stored until a trigger signal arrives, the input rate to column bus μ' is reduced compared to the hit rate μ as: $\mu' = \mu \times r \times t$, where r is the trigger rate and t is the bunch crossing period. In this situation there is a more relaxed constraint on the bandwidth, but the limiting factor is the buffer depth: the amount of memory designed depends both on the expected rate μ and on the trigger latency t as $\propto \mu \times t$, which means that the higher the trigger latency the lower the hit rate to cope with.

In order to have an efficient usage of memory in the pixel area it is convenient to group pixels into regions with shared storage. Let's compare two different situations: in the first one a buffer is located on each pixel area, while in the second one a core of four pixels share a common buffer. This architecture is commonly called FE-I4 and in figure 2.12 it is shown the block diagram of a core with such architecture. Consider a 50 kHz single pixel hits rate and a trigger latency of 5 μs, the probability of losing hits is:

$$P(N > 1|\nu) = 1 - P(N = 0|\nu) - P(N = 1|\nu) = 1 - e^{-\nu}(1 + \nu) \approx 2.6\% \tag{2.7}$$

where I have assumed a Poissonian distribution with mean $\nu = 0.25$ to describe the counts N. To get an efficiency ϵ greater than 99.9 % a 3 hit depth buffer is needed:

$$P(N > 3|\nu) = 1 - \sum_{i=0}^{3} P(N = i|\nu) < 0.1\% \tag{2.8}$$

Consider the second situation: if the average single pixel rate is still 50 kHz, grouping four pixels the mean number of hits per trigger latency is $\nu = 0.25 \times 4 = 1$. To get an efficiency of 99.9% (eq. 2.8) a buffer depth of 5 hits in the four-pixels region, instead of 3 per pixels, is needed. Then to get the same efficiency for a group of 4 pixels, in the first case a buffer depth of 12 hits is needed, while in the second case only a 5-hits depth buffer is; this is one of the reasons why pixels are grouped in cores.

CHAPTER 3

MAPS devices description

This chapter presents a brief description of the two DMAPS devices that I have tested during the thesis work: the TJ-Monopix1 device, originating from developments for ATLAS upgrade, and the ARCADIA Main Demonstrator device, developed by INFN ARCADIA collaboration as a multipurpose chip for accelerator and space applications.

3.1 TJ-Monopix1

TJ-Monopix1 is a small electrode DMAPS with fast R/O capability, fabricated by TowerJazz foundry in 180 nm CMOS imaging process. It is part, together with prototypes from other series such as TJ-MALTA, of the ongoing R&D efforts aimed at developing DMAPS in commercial CMOS processes, that could cope with the requirements at accelerator experiments. Both TJ-Monopix and TJ-MALTA series [14], produced with the same technology by TowerJazz (the timeline of the foundry products is shown in figure 3.1), are small electrode demonstrators and principally differ in the readout design: while Monopix implements a column-drain R/O, an asynchronous R/O without any distribution of BCID has been used by TJ-Malta in order to reduce power consumption.

Figure 3.1: Timeline in TowerJazz productions in 180 nm CMOS imaging process

	LF-Monopix1	TJ-Monopix1
Resistivity	>2 kΩcm	>1 kΩcm
Pixel size	50 × 250µm^2	36 × 40 µm^2
Depth	100-750 µm	25 µm
Capacity	∼ 400 fF	∼ 3 fF
Preamplifier	charge	voltage
Threshold trimming	on pixel (4-bit DAC)	global threshold
ToT	8 bits	6 bits
Consumption	∼ 300mW/cm^2	∼ 120mW/cm^2
Threshold	1500 e^-	∼ 270 e^-
ENC	100 e^-	∼ 9 e^-

Table 3.1: Main characteristics of Monopix1 produced by TowerJazz and LFoundry [16][17]

Another Monopix series, but in 150 nm CMOS technology, has been produced by LFoundry [15]. The main differences between the LF-Monopix1 and the TJ-Monopix1 (summarized in table 3.2), lay in the sensor rather than in the readout architecture, as both chips implements a fast column drain R/O with Time Over Threshold (ToT) capability [16][17]. Concerning the sensors, either are based on a p-type substrate, but with slightly different resistivity; in addition LFoundry pixels are larger, thicker and have a large fill factor (the very deep n-well covers ∼55% of the pixel area). The primary consequence is that LF-Monopix1 pixels have a higher capacitance resulting in higher consumption and noise. As I discussed in section 2.5.1, the fact that LF-Monopix has a large fill factor electrode is expected to improve its radiation hardness. Indeed, a comparison of the perfomance of the two chips showed that TJ-Monopix suffers a comparatively larger degradation of efficiency after irradiation, due to the low electric field in the pixel corner; on the other hand, a drawback of the large fill factor in LF-Monopix is a significant cross-talk.

The TJ-Monopix1 chip contains, apart from the pixels matrix, all the required support blocks used for configuration and testing:

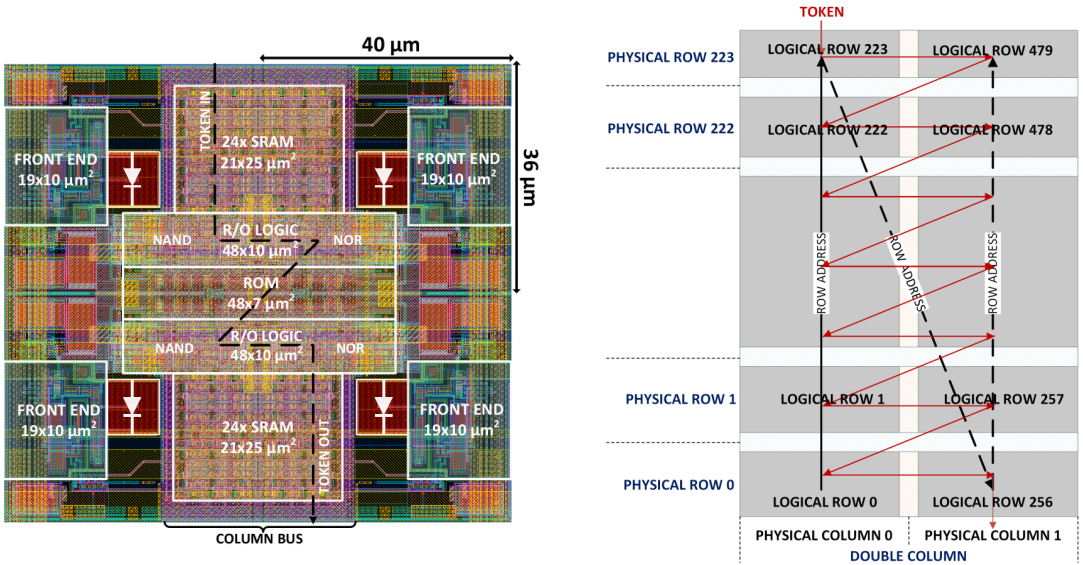
- the whole matrix contains 224 × 448 pixels, yielding a total active area approximately equal to 145 mm^2 over a total area of 1×2cm^2;
- at the chip periphery are placed some 7-bit Digital to Analog Converter (DAC), used to generate the analog bias voltage and current levels and to confiugure the FE;
- at the EoC is placed a serializer to transferred data immediately, indeed no trigger memory is implemented in these prototypes;
- the matrix power pads are distributed at the sides
- four pixels which have analog output and which can be monitored with an oscilloscope, and therefore used for testing

Parameter	Value
Matrix size	$1\times2\text{cm}^2$
Pixel size	$36 \times 40\ \mu\text{m}^2$
Depth	$25\,\mu\text{m}$
Electrode size	$2\,\mu\text{m}$
BCID	40 MHz
ToT-bit	6
Power consumption	$\sim 120\ \text{mW/cm}^2$

Table 3.2: Charateristics of TJ-Monopix1 chip

Pixels are grouped in 2×2 cores (fig.3.2a): this layout allows to separate the analog and the digital electronics area in order to reduce the possible interference between the two parts. In addition it simplifies the routing of data as pixels on double column share the same column-bus to EoC. Therefore pixels can be addressed through the physical column/row or through the logical column/row, as shown in figure 3.2b.

Concerning the integration of the chip in a readout path, TJ-Monopix1 chips have been wire-bonded on a dedicated carrier board , the Printed Circuit Board (PCB). Two other board between the DAQ and the chip: the General Purpose Analog Card (GPAC), which provides power supply channels, current/voltage bias sources and I/O buffer, and the MIO3 FPGA, which strictly interacts with the DAQ (fig.3.8).



(a) (b)

Figure 3.2: (a) Layout of a core containing 4 pixels: the analog FE and the digital part are separated in order to reduce crosstalk. (b) Propagation logic of signals along the double column buses.

3.1.1 The sensor

As already anticipated, TJ-Monopix1 has a p-type epitaxial layer and a n doped small collection electrode (2 µm in diameter); to prevent the n-wells housing the PMOS transistors from competing for the charge collection, a deep p-well substrate, common to all the pixel FE area, is used. TJ-Monopix1 adopts the modification described in section 2.5.2 that allows to achieve a planar depletion region near the electrode applying a relatively small reverse bias voltage. This modification improves the efficiency of the detector, especially after irradiation, however a simulation of the electric field in the sensor, made with the software TCAD (Technology Computer Aided Design), shows that a nonuniform field is still produced in the lateral regions of the pixel compromising the efficiency at the corner. Two variations to the process have been proposed in order to further enhance the transversal component of electric field at the pixel borders: on a sample of chip, which includes the one in Pisa, a portion of low dose implant has been removed, creating a step discontinuity in the deep p-well corner (fig.3.3); the second solution proposed consists in adding an extra deep p-well near the pixel edge. A side effect of the alteration in the low dose implant is that the separation between the deep p-well and the p-substrate becomes weak to the point that they must be biased together to avoid punch-through.

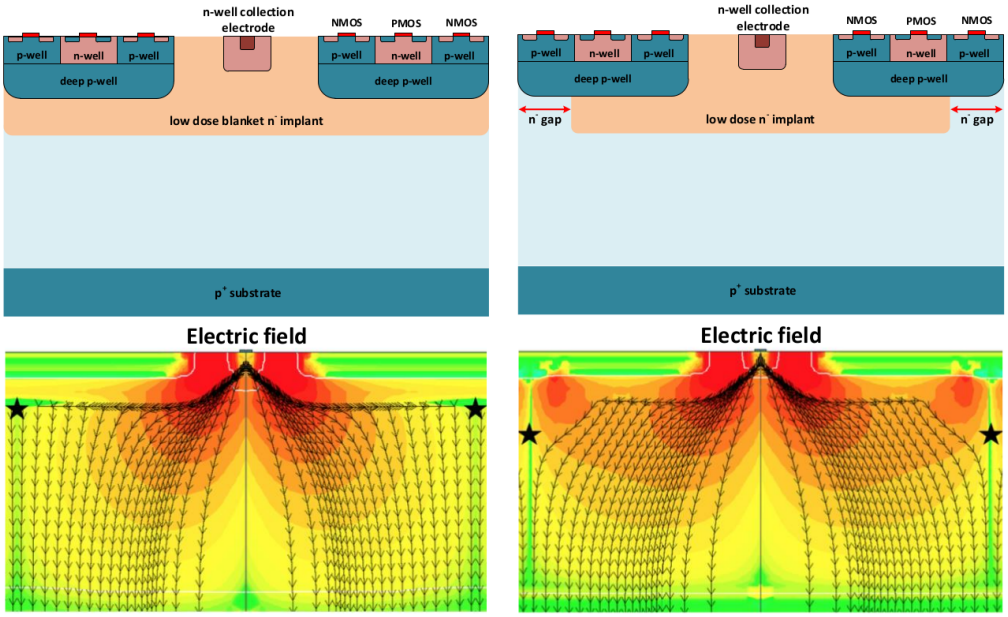


Figure 3.3: On the left the cross-section of a monolithic pixel in the TJ-Monopix with modified process; in the scheme on the right an additional gap in the low dose implant is created to improve the collection of charge due to a bigger lateral component of the electric field. The point indicated by a star at the pixel corner corresponds to the region where the transversal component of the electric field drops; introducing the gap then, a transversal component is restored and the collection becomes more efficient also in that region.

Moreover, to investigate the charge collection properties, pixels within the matrix are split between bottom top half and bottom half and feature a variation in the coverage of

the deep p-well: the electronics area can be fully covered or not. In particular the pixels belonging to rows from 0 to 111 are fully covered (FDPW) and pixels belonging to rows from 112 to 223 have a reduced p-well (RDPW), resulting in a enhancement of the lateral component of the electric field.

3.1.2 Front end

One of the main advantage of this chip is the small collection electrode, which results in a small capacitance (C_{in}=3 fF) allowing for high input signal amplitude and single stage of amplification, which obviously improves the signal to noise ratio performance of the FE. Assuming a fully depleted epitaxial layer of 25 µm, which corresponds approximately to a 20 µm of deep sensing volume, a MIP should produce ∼1600 e−, then:

$$V_{in} = \frac{1600\, e^- \times 1.6\, 10^{-19}\,\mathrm{C}}{3\,\mathrm{fF}} = 85\,\mathrm{mV} \tag{3.1}$$

Secondly, a reset mechanism which slowly discharges the detector capacitance must be included in the circuit:

$$V_{in} = \frac{Q_s}{C_{in}} e^{-t/R_b C_{in}} \tag{3.2}$$

where R_b is the equivalent reset element. The general constraint which must be satisfied is that the discharge time $\tau = R_b C_{in}$ must be slower than the characteristic time of the amplifier, otherwise a signal loss could occurs. Traditionally the reset can be implemented in two different way: with a forward biased diode, that might be implemented by a simple p^+ diffusion inside the well of the collection electrode n, or with a PMOS transistor. Despite of the simplicity of the diode reset, since it is a non-linear element, the discharge would depend on the quantity of charge Q generated on the n electrode, prejudicing the linearity of the analog output (Q-ToT). To ensure the linearity, a PMOS reset is the method preferred in design such TJ-Monopix1 with analog output; the PMOS transistor, indeed, can acts as a constant current source and then used to discharge the sensor. Although the PMOS reset is capable of providing a constant current, it has to be manually re-tuned every time in order to restore the input DC baseline voltage; to do that a low-frequency feedback is implemented in the FE circuit.

The matrix is split in four sections, each one corresponding to a slightly different flavor of the FE, implemented in order to test more options. All the flavors implement a source-follower readout with common buses among double-column: the standard and reference flavor is the flavor B, that features a PMOS input reset (referred as "PMOS reset"). Flavor A is identical to flavor B except for the realization of the source follower (it is a gated one): in the circuit of the gated version there is one transistor more that operates on the baseline and on the feedback mechanism; this aim to reduce the power consumption and results in a higher signal baseline, and then in a lower effective threshold. C instead implements a novel leakage compensation circuit, with a PMOS reset configuration. Moreover the collection electrode can be either DC-coupled to the readout electronics, as in flavors A, B, C, or AC-coupled through a metal-oxide-metal capacitance, as in D is AC-coupled. The latter one allows applying a high bias voltage to the electrode n and for this reason

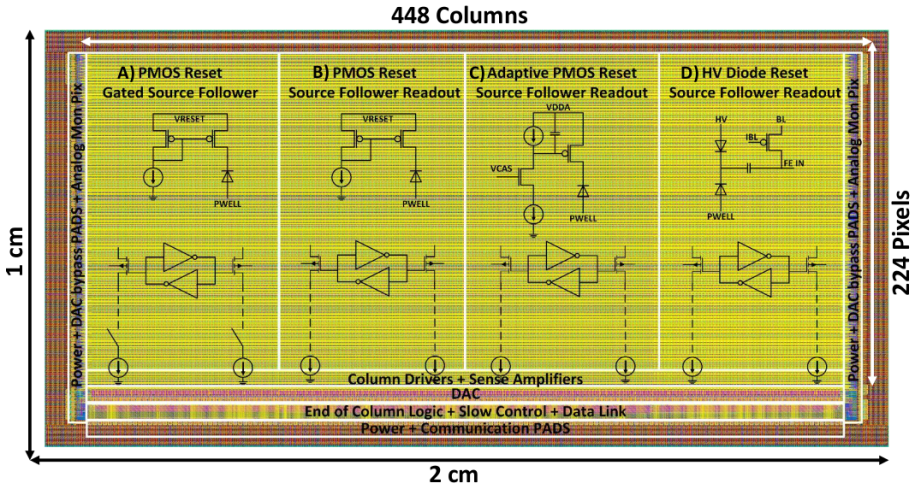
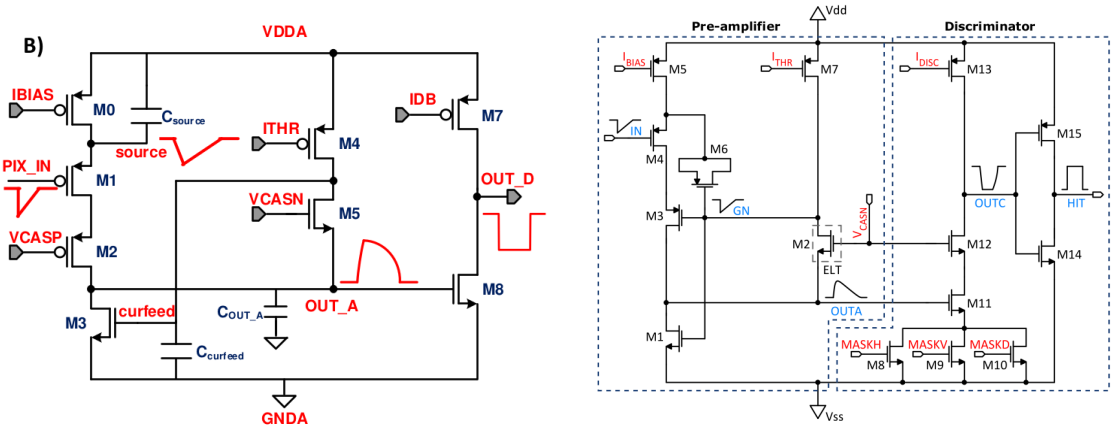


Figure 3.4: TJ-Monopix1 has been developed in four different flavor. The flavor PMOS reset (B) is considered as the reference one.



(a) (b)

Figure 3.5: In (a) basic ALPIDE FE circuit; in (b) the real circuit implemented in TJ-Monopix1

the flavor D is also called "HV flavor". Unfortunately the "HV" suffer from a signal loss, which can achieve even the 50%, due to the additional parasitic capacity introduced at the input node. The nominal HV voltage above which the breakdown begins is ∼50 V; however at values bigger than 20 V, the gain is not expected increasing anymore, since the depletion zone is already at its greatest value.

ALPIDE chips, developed by the ALICE collaboration, implemented a standard FE to the point that many CMOS MAPS detectors used a similar FE and are called "ALIPDE-like". Considering that both TJ-Monopix1 and ARCADIA-MD1 have an ALPIDE-like FE, I am going to explain the broad principles of ALPIDE early FE stage. The general idea of the amplification is to transfer the charge from a bigger capacitance[18], C_{source}, to a smaller one, C_{out}: the input transistor M1 with current source IBIAS acts as a source follower and this forces the source of M1 to be equal to the gate input $\Delta V_{PIX\_IN} =$

Q_{IN}/C_{IN}.

$$Q_{source} = C_{source}\Delta V_{PIX\_IN} \tag{3.3}$$

The current in M2 and the charge accumulates on C_{out} is fixed by the one on C_{source}:

$$\Delta V_{OUT\_A} = \frac{Q_{source}}{C_{OUT\_A}} = \frac{C_{source}\Delta V_{PIX\_IN}}{C_{OUT\_A}} = \frac{C_{Source}}{C_{OUT\_A}}\frac{Q_{IN}}{C_{IN}} \tag{3.4}$$

A second branch (M4, M5) is used to generate a low frequency feedback, where VCASN and ITHR set the baseline value of the signal on C_{OUT_A} and the velocity to goes down to the baseline. The CURFEED transistor is then used to control the current in the input branch and also determines the current that flows at the discriminator input. Finally IDB defines the charge threshold with which the signal OUT_A must be compared: depending on if the signal is higher than the threshold or not, the OUT_D is high or low respectively. In table 3.3 are listed the FE parameters settable by the DAQ, while in figure 3.7 are shown some photos of the oscilloscope of the analog output available on the TJ-Monopix1 board; in the acquisitions we have used the injection mode to inject pulses of a fixed voltage amplitude and look at the response at different values of the parameters of the FE.

Despite TJ-Monopix1 has an ALPIDE-like FE, it implements a circuit slightly different from the one in figure 3.5a, but the actual circuit implemented is shown in figure 3.5b: the principal difference lays in the addition of disabling pixels' readout. This possibility is uttermost important in order to reduce the hit rate and to avoid saturating the bandwidth due to the noisy pixels, which typically are those with manufacturing defects. In the circuit transistors M8, M9 and M10 have the function of disabling registers with coordinates MASKH, MASKV and MASKD (respectively vertical, horizontal and diagonal) from readout: if all three transistors-signals are low, the pixel's discriminator is disabled. Compared with a configurable masking register which would allow disabling pixels individually, to use a triple redundancy reduces the sensitivity to SEU but also gives amount of intentionally masked ("ghost") pixels. This approach is suitable only for extremely small number N of pixel has to be masked: if two coordinate projection scheme had been implemented, the number of ghost pixels would have scale with N^2, if instead three coordinates are used, the N's exponential is lower than 2 (fig.3.6)

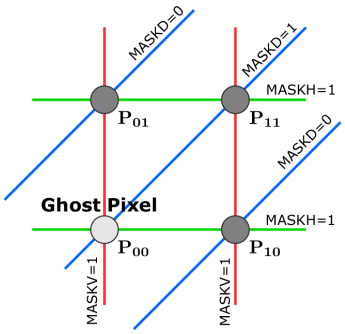


Figure 3.6

Parameter	Meaning	
IBIAS	mainly controls the rise time	yes
IDB	sets the discriminator threshold	yes
ITHR	sets the velocity of the return to the baseline	yes
VCASN	sets the baseline of the signal	yes
VRESET	sets the return to the baseline	yes
IRESET	sets the return to the baseline	no

Table 3.3: FE parameters which must be set through the DAQ. "Function" means that higher parameter implies higher value

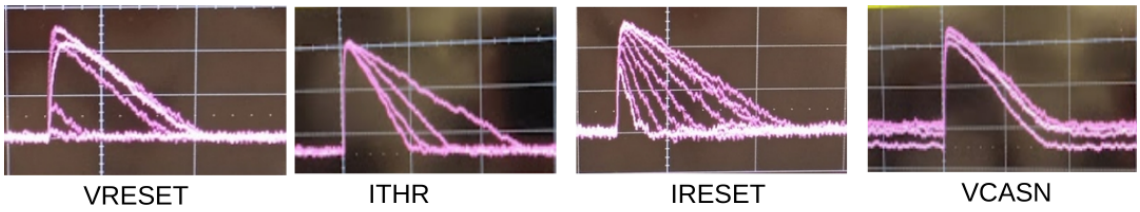


Figure 3.7: Some photos of the analog output available on the TJ-Monopix1 board. The parameters have been sweep in the following range: VRESET in 10-60 DAC, ITHR in 5-15 DAC, IRESET in 0-127 DAC, VCASN in 0-15 DAC.

3.1.3 Readout logic

TJ-Monopix1 has a triggerless, fast and with ToT capability R/O which is based on a column-drain architecture. On the pixel are located two Random Access Memory (RAM) cells to store the 6-bit LE and 6-bit TE of the pulse, and a Read-Only Memory (ROM) containing the 9-bit pixel address. Excluded these memories, TJ-Monopix1 has not any other buffer: if a hit arrives while the pixel is already storing a previous one, the new data get lost. After being read, the data packet is sent to the EoC periphery of the matrix, where a serializer transfers it off-chip to an FPGA (3.8). There a FIFO is used to temporarily stored the data, which is transmitted to a computer through an ethernet cable.

The access to the pixels' memory and the transmission of the data to the EoC, following a priority chain, is managed by control signals and is based on a finite state machine composed by four state: no-operation (NOP), freeze (FRZ), read (RD) and data transfer (DTA). The readout sequence (fig.3.9) starts with the TE of a pulse: the pixel immediately tries to grab the column-bus turning up a hit flag signal called *token*. The token is used to control the priority chain and propagates across the column indicating what pixel that must be read. To start the readout and avoid that the arrival of new hit disrupt the priority logic, a *freeze* signal is activated, and then a *read* signal controls the readout and the access to memory. During the freeze, the state of the token for all pixels on the matrix remains set: this does not forbid new hits on other pixels from being recorded, but forbids pixel hits from turning on the token until the freeze is ended. The freeze stays on until

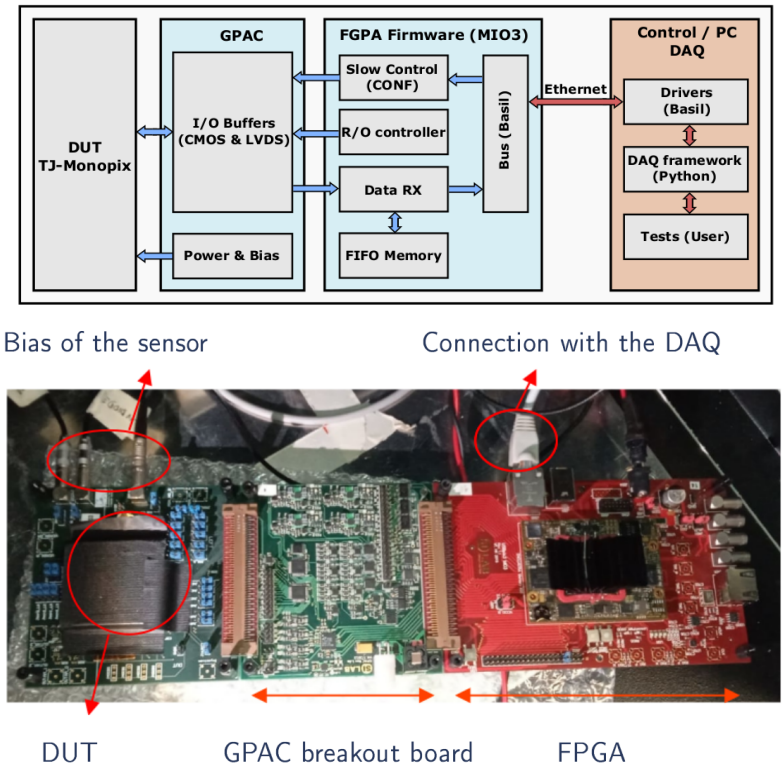


Figure 3.8: TJ-Monopix1 is controlled by the PC via the DAQ. Between the PC and the DUT commands and data pass through an FPGA and a breackout board.

the token covers the whole priority chain and gets the EoC: during that time new tokens cannot be turned on, and all hits arrived during a freeze will turn on their tokens at the end of the previous freeze. Since the start of the token is used to assign a timestamp to the hit, the token time has a direct impact on the resolution of the time measurement. This could be a problem especially with high hit rate, when the number of hits arriving during the freeze may be large.

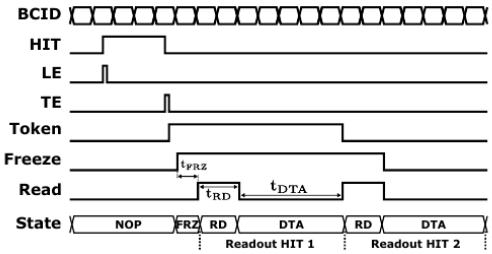


Figure 3.9: Readout timing diagram: in this example two hits are being processed

The analog FE circuit and the pixel control logic are connected by an edge detector which is used to determine the LE and the TE of the hit pulse(fig.3.10): when the TE is stored in the first latch the edge detector is disabled and, if the FREEZE signal is not set yet, the readout starts. At this point the HIT flag is set in a second latch and a token signal is produced and depending on the value of Token in the pixel can be read or must wait until the Token in is off. In figure an OR is used to manage the token propagation, but

since a native OR logic port cannot be implemented with CMOS logic, a sum of a NOR and of an inverter is actually used; this construct significantly increases the propagation delay (the timing dispersion along a column of 0.1-0.2 ns) of the token and to speed up the circuit optimized solution are often implemented. When the pixel becomes the next to be read in the queue, and at the rising edge of the READ signal, the state of the pixel is stored in a D-latch and the pixel is allowed to use the data bus; the TE and the HIT flag latches are reset and a READINT signal that enable access of the RAM and ROM cells is produced.

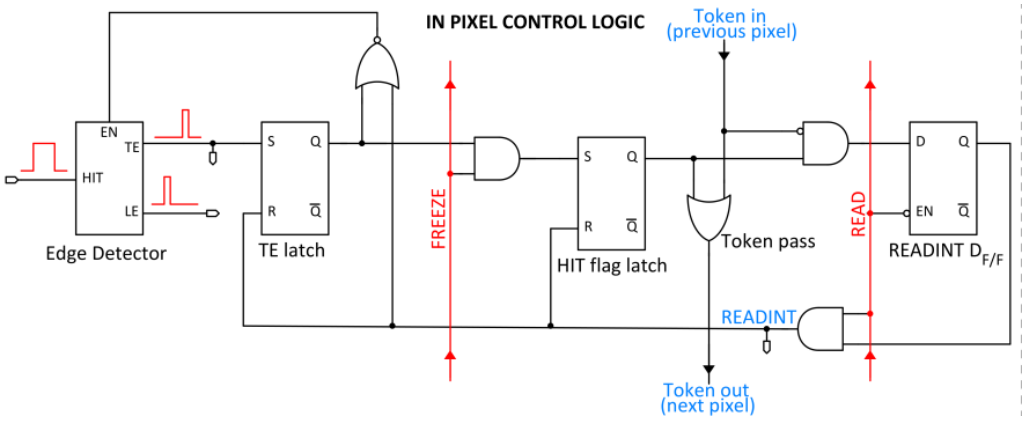


Figure 3.10

The final data must provide all the hits information: the pixel address, the ToT and the timestamp. All those parts are assigned and appended at different time during the R/O chain:

- **Pixel address:** while the double column address (6-bit) is appended by the EoC circuit, the row address (8-bits for each flavor) and the physical column in the doublet (1-bit) are assigned by the in-pixel logic

- **ToT:** is obtained offline from the difference of 6-bits TE and 6-bits LE, stored by the edge detector in-pixel; since a 40 MHz BCID is distributed across the matrix, the ToT value is range 0-64 clock cycle which corresponds to 0-1.6 µs

- **Timestamp:** the timestamp of the hit correspond to the time when the pixel set up the token; it is assigned by the FPGA, that uses the LE, TE and a 640 MHz clock to derive it. For all those hits which arrived while the matrix is frozen, the timestamp is no more correlated with the time of arrival of the particle

When the bits are joined up together the complete hit data packet is 27-bit. The fact that the ToT is represented with a 6-bits variable implies that, for a very high charge, when the ToT is very long and exceeds the maximum value (64 clock counts), it comes back to zero (fig.3.11). While the minimum and maximum values of the ToT are equal to all the pixels in the matrix (0 and 64 respectively), the corresponding charge when rollover happens is different from pixel to pixel and, as I will clarify in section 4.1.3.

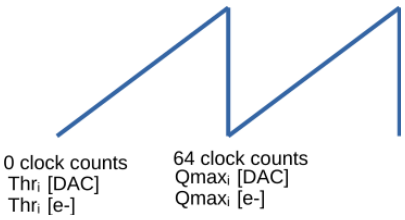


Figure 3.11: ToT rollover scheme

3.2 ARCADIA MD1

ARCADIA (Advanced Readout CMOS Architectures with Depleted Integrated sensor Arrays)[19][20] and SEED (Sensor with Embedded Electronic Development) are collaborations involved in the development of MAPS sensors based on the CMOS technology and both having LFoundry as industrial partner. Many concept and performances studies have been carried out with simulations and small-scale test structure by SEED, before ARCADIA, applying the experience developed with SEED to a full chip prototype, the MD1. An example of small-scale prototypes produced for testing is MATISSE: it is made by 24×24 pixels organized in 4 columns; each pixel has an analog output, which allows for energy loss measurements, and a shutter snapshot readout with a speed that can reach 5 MHz.

The ARCADIA-MD series target is the development of a novel CMOS sensor platform allowing for fully depleted active sensors with thickness in the range 50 µm to 500 µm. A small charge collecting electrode to achieve a good signal to noise ratio, a high time resolution (the lower bound is set at $\mathcal{O}$(µs) but also more advanced solutions are being investigated for a $\mathcal{O}$(10 ns)) and a scalable readout architecture with low power consumption are the main requirement imposed by ARCADIA; the Main Demonstrator 1, has been submitted in 2020, and its characteristics are shown in table 3.4. A second main demonstrator, ARCADIA-MD2, has been submitted in Summer 2021; it features a similar design of MD1, but it is expected to be faster and to have a lower power consumption thanks to a logic and buffering optimization.

Parameter	Value
Matrix size	1.28×1.28 cm^2
Pixel size	25×25 µm^2
Depth	48/100/200µm
Electrode size	9×9µm^2
Power consumption	$\sim$ 10mW/cm^2
Output signal	digital

Table 3.4: ARCADIA MD1 characteristics

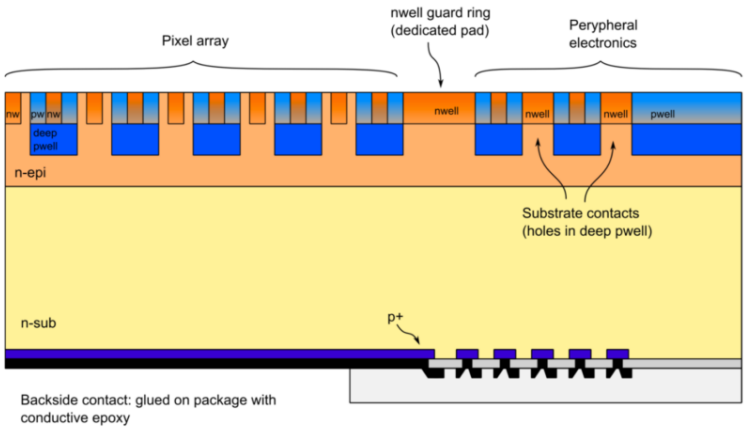


Figure 3.12: Cross section of the ARCADIA MD1 sensor

3.2.1 The sensor and the front end

ARCADIA-MD1 is an LFoundry chip fabricated in 110 nm CMOS technology.The sensor (fig.3.12) is made by a p substrate and an n doped diode within a n epitaxial layer; a custom patterned backside has been developed in collaboration with LFoundry to introduce junction at the bottom surface which allows a full depletion. A deep p-well enclosure has been used to shield the n-well contained in the electronics circuit and deny competing in charge collection; considering the isolation the resulting area available for the analog circuit is 223 µm^2.

Up to now the sensor has been implemented in three different variant: 48 µm, 100 µm and 200 µm thick, each with the same FE and readout logic but requiring a different biasing (always higher than 10 V). In figure 3.13 is shown a TCAD simulation, which includes two pixels and a guard ring, of the electric-field line within the sensor; being part of DMAPS and being operated in fully depletion, the charge is fastly collected by drift along the electric field lines.

There are three types of configuration registers which are used to configure the matrix:

- the Pixel Configuration Register (PCR), which is a 2-bits word used for enabling respectively the masking and injection functionalities. Each bit is made by a latch which occupy 14.6 µm^2 out of the per-pixel area available, 223 µm^2 then it is clear that there is not much extra space for any more configuration bits. The on-pixel PCR circuit is shown in figure 3.14.

- the Internal Configuration Register (ICR), which are used for the comunication with the FPGA, for example to send a pulse, reset or configure the whole matrix.

- the Global Configuration Registers (GCR), which are used to set the configuration of the FE parameters are similar to the one of the TJ-Monopix1 circuit, and they are (partially) listed in table 3.5.

All the bias (for the FE and for the sensor) are supplied by padframes (a top, a bottom and a side one) placed aside the matrix, which also provide the clock, the reset, the test pulse

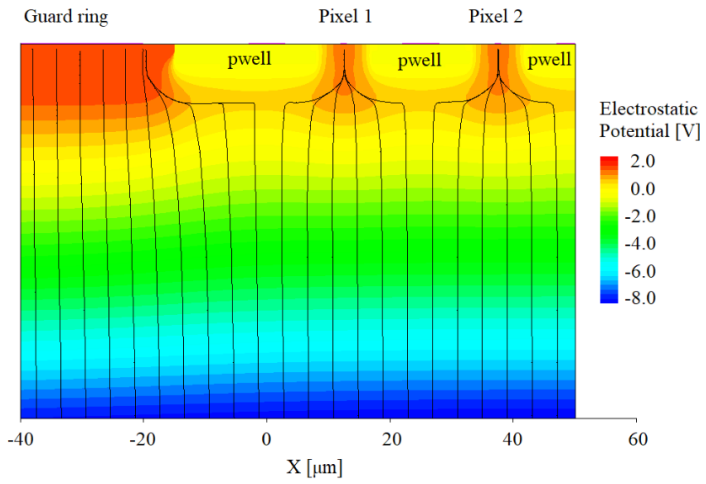


Figure 3.13: TCAD simulation of the electric field in the surface region of a portion of the sensor including two pixels.

for the injection circuit and the comunication signals. The timestamp clock, which defines the timestamp granularity, is not internally generated but it is obtained from an external clock of 320 MHz with a clock divider; a 4-bit GCR is used set the base-2 logarithm of the dividing ratio, such as the timestamp clock frequency is:

$$f_{timestamp} = \frac{320\,\text{MHz}}{2^{GCR}} \tag{3.5}$$

and then varies in range 320 MHz and 20 MHz.

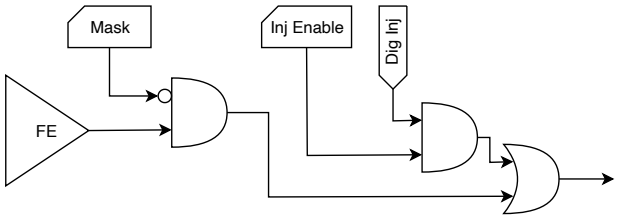


Figure 3.14: Logic used for each pixel to implement the injection and the masking.

MD1 chips have been submitted in two different front end options: they are commonly called ALPIDE-like and bulk-driven. The differences between them are in the FE circuit and in the biasing current of the registers, while the underlying readout is the same. The main difference is in the amplification stage, while in the ALPIDE-like flavor the amplification is implemented as explained in section 3.1.2, in the bulk-driven flavor the gain is adjusted by the ratio of two transconduttances. Consequently, some of the biasing registers, whose current is settable externally by the DAQ, have different default values and they might not be available at all in one of the flavor. An example is the ICLIP register, which is available only in the bulk driven flavor despite the transistor to which refers is implemented in both the flavor; its function is similar to the *curfeed* capacitor in figure 3.5(a), which controls the current in the input branch of the FE and also influences the value of the baseline at the discriminator input.

Parameter	Meaning
CLK_DIVIDER	$\log_2$ number to divide the input clock
VINREF	provides the current to restore the input node i
VCASN	sets the threshold
IBIAS	sets the baseline
IFB	current in the feedback branch
ID	discriminator current
ICLIP	baseline

Table 3.5: FE MD1 parameters which must be setted through the DAQ.

3.2.2 Readout logic and data structure

One of the main ambition of the MD1 is to achieve the lowest possible power consumption, hopefully less than $20\,\mathrm{mW/cm^2}$; this is important for applications in the field of space experiment, where the power consumption and the cooling are a major issue. In order to fulfill that requirement, the matrix is clockless and the readout is triggerless; moreover the chip can be operated both in the high rate mode and low rate by enabling on if only one or all serializers, placed at the periphery of the matrix. In addition, to save as much area as possible, buffers have not been included on the matrix, at the expense of the maximum hit rate sustainable. The readout then is completely data push and when a hit is received immediately starts the readout mechanism to transmit it off chip. The board hosting the chip is connected with a breakout board, which is connected to the FPGA; a data packet sent to the EoS, is then encoded and transmitted to the FPGA using a 320 MHz DDR serializers and then transmitted by ethernet to the PC. A photo of the experimental setup is shown in figure 3.15.

The chip structure is meant to optimize the power consumption and the scalability for future up-scaling retaining high rate operation; in particular it is divided into a physical and logical hierarchy, which also reflects in the way the data packets are built (tab.3.6). First of all, the 512 columns are split in 16 sections each one containing 512×32 pixels and having its own biasing lines and serializers at the matrix periphery. Each section is is divided 512×2 double-column mirrored, which just as in TJ-Monopix1, share the same readout buses placed between them and having analog logic on the sides. The rows, then, are divided in group of 32, resulting in core with 32×2 pixels. Finally each core is sub-divided in regions, each one containing 4×2 pixels.

The readout has been designed with the constraints of being capable of handling a rate of $100\,\mathrm{MHz/cm^2}$, and it has been optimized to minimize the amount of logic and to have a high bandwidth of transmission of the data to the periphery. For this reason not all pixels have been provided of the readout logic. In particular, each pixel region can either be Master or Slave, depending on if has or has not the readout capability. The Master's data packets are therefore composed of two parts: the hitmap of the Master itself and the one of Slave. Moreover, the pioneer idea of ARCADIA-MD1, which has as

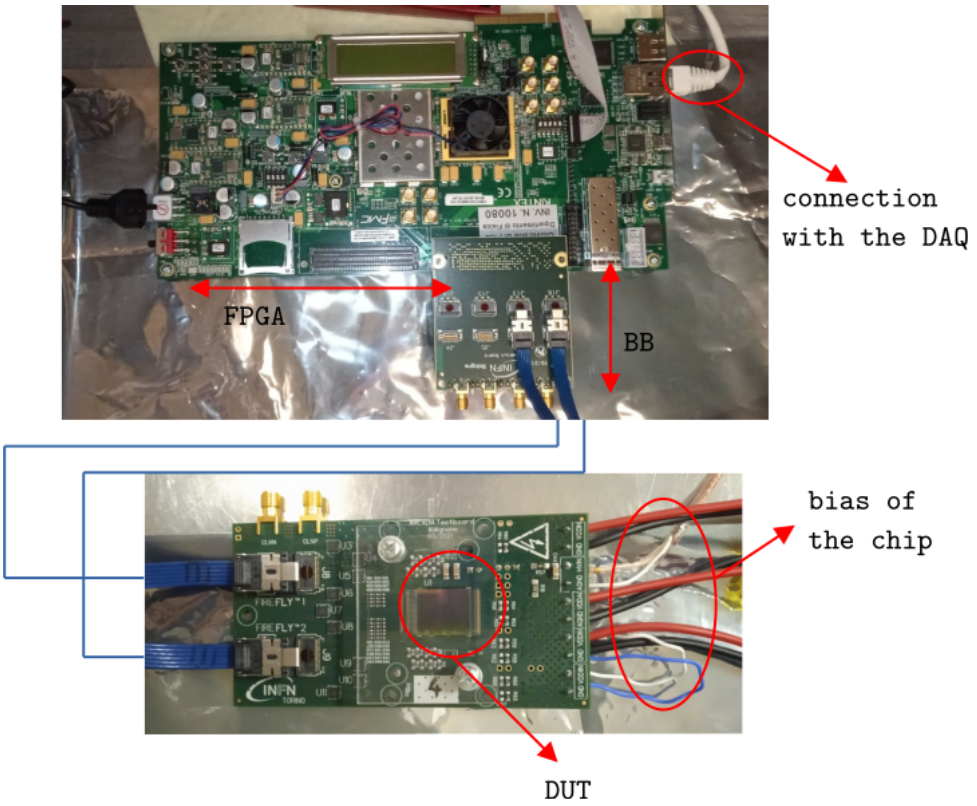


Figure 3.15: Board hosting the MD1 chip, FPGA and breakout board. The chip and breakout boards must be connected with the blue cables

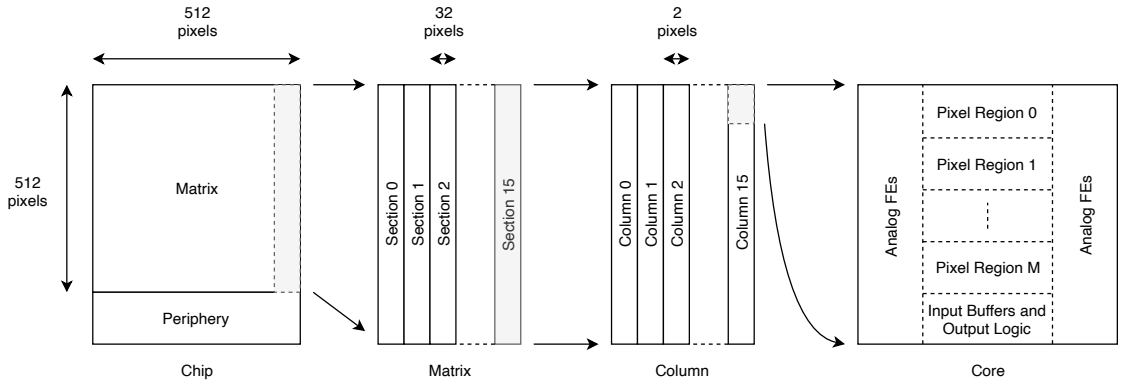


Figure 3.16: Hierarchy of the matrix division

finally goal the test of a readout capable of transmit cluster data in as few data packets as possible, is the possibility of the Master to decide what Slave (top or bottom) to read; the information of what Slave has been selected is represented by a bit, often called *hot bit*, in the data-packet. Every pixel has an associated status register, that essentially is a flip flop (FF), which is set to 1 when the pixel stores a hit; an OR of the FF within the Master or the Slave region generates an active flag which is used to require a readout by the EoS. In figure 3.17 is shown the circuit with the logic of assignment of the Slave to the Master. Depending on the active flags of the neighbours Masters, the Slave hitmap is assigned to the one at the top or bottom: if both the Masters have an active flag, the

Bits	Meaning
31:24	timestamp
23:20	section index
19:16	column index
15:9	pixel region
8:0	hitmap

Table 3.6: Data packet structure implemented by the MD1 readout logic.

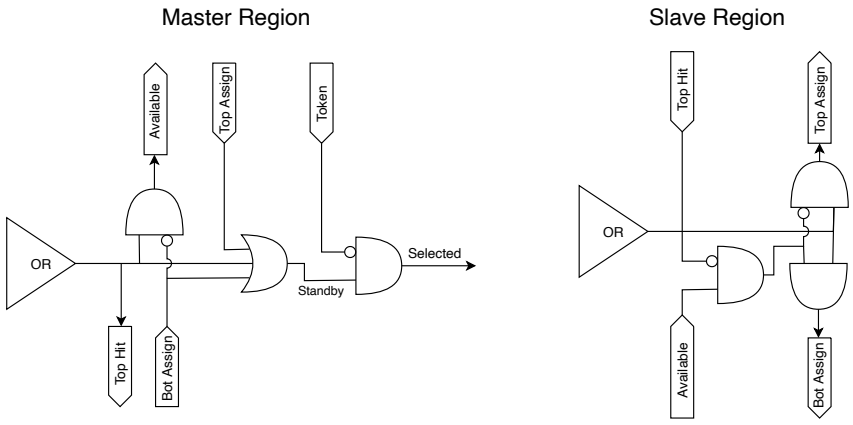


Figure 3.17: Logic of the circuit to implement the online clustering and deciding if to assign the Slave to the top or bottom Master.

Slave is assigned to the top one. In the example in figure 3.18 two Master-Slave regions are considered: the hitmaps of the Master (red colored in the example) and Slave (blue colored) are joined together within a unique data packet; a bit (green colored) is used to specify what Slave the hitmap corresponds.

The data packets are then transmitted to the End of Section (EoS) with a priority chain similar to what happens in TJ-Monopix1. If at least one Master set a high flag, a Token signal is generated and is assigned to the high priority Master in the column, together with a Full flag which is distributed to the active Masters in the whole column in order to deny more region to be accessed at the same time. The readout then propagates down the column from Master to Master, skipping the empty cores; the Master selected for the readout is the one with the flag high and with an input (from top) Token equal to 0. In the example in figure 3.19 the Token is propagated from the Pixel Region (PR) 10 to the PR 7. In the three readout steps the red Masters are the ones selected for the readout, while the yellow are the ones which an active flag high; gray color is used for empty regions. When a specific Master has been selected, a Read signal is generated both to transmit the data to the EoS and also to generate a reset for the just read pixels. Once the pixels are reset, the Master's Full and Token flags fall, and the following region which satisfies the two readout conditions explained above, becomes selected.

The performances of the readout has been studied with simulations by the designer of the chip. Random hits events with cluster size of 4 pixels on average, with a Pois-

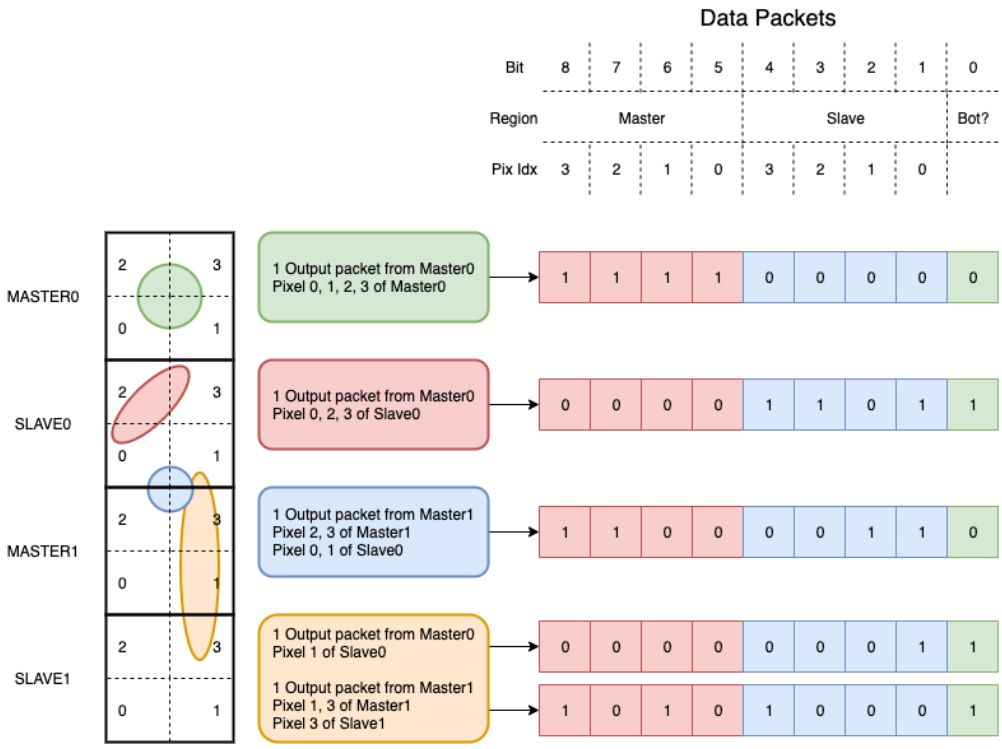


Figure 3.18: Different cluster structures and the data packet produced by them are shown in the example.

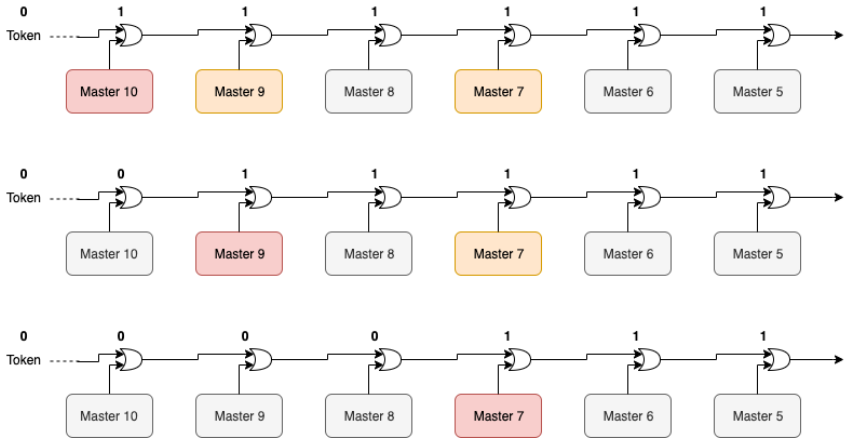


Figure 3.19: Three steps in the readout sequence on the region-column: the Token is propagated from the Master 10, to Master 9 and then to Master 7, according to the priority chain readout.

sonian distribution in time and uniformly distributed on the matrix has been generated. They state that with particle hit rate of 100 MHz/cm^2, considering a portion of matrix of three section (512×96), the efficiency results to be 98.7%, while reducing the hit rate to 80 MHz/cm^2 it is even higher achieving the 99.95%.

CHAPTER 4

Characterization

In this chapter I will present the tests performed on TJ-Monopix1 and ARCADIA-MD1; firstly I will describe the characterization of TJ-Monopix1, in terms of threshold, noise, dead time, and also provide a calibration of the output signal, and then I will show some preliminary results obtained with ARCADIA-MD1.

4.1 TJ-Monopix1 characterization

4.1.1 Threshold and noise: figure of merit for pixel detectors

A characterization of threshold and noise is typically necessary since these values have an impact on the operating conditions and on the performance of the chips, so much that the signal to threshold ratio may be considered as the figure of merit for pixel detectors rather than the signal to noise ratio. The mean minimum stable threshold evolved through different generation of chips: in the 1st generation it was around $2500\,e^-$ while in the 3rd (corresponding to nowadays chips) is less than $500\,e^-$; the threshold of TJ-Monopix1 is around $500\,e^-$ and is in agreement with this trend. Lower threshold allows for smaller signals produced in thinner sensor: from $16\,000\,e^-$ produced in 200 µm, the signal expected moved down to $2000\,e^-$ produced in 25 µm.

The threshold should be located between the noise peak around the baseline and the signal distribution, in particular it has to be low enough to maintain a high signal efficiency, but also high enough to cut the noise: for a low threshold many pixels can fire at the same time and a positive crosstalk feedback can set off a chain reaction eventually, causing all the other pixels to fire. Thus, the noise sets a lower bound to the threshold: if an occupancy $\leqslant 10^{-4}$ is required, for example, a probability of firing lower than that value is needed and this, assuming a gaussian noise, requires that the threshold is set at least $3.7\,\sigma$, with σ the RMS of the noise. In this case, if the noise is $100\,e^-$ (reasonable), the threshold must be higher than $3.7 \times 100\,e^-$.[13] Typically this argument sets only a minimal bound to the threshold since the variation with time and from pixel to pixel have to be taken into

account: the temperature, the shift in MOSFET threshold voltage caused by radiation damage in the oxide layer and the process parameters variation across the wafer (as for example process mismatch between transistors).

The noise is parameterized as Equivalent Noise Charge (ENC), which is defined as the ratio between the noise N at the output expressed in Volt and the out voltage signal S produced by $1\,e^-$ entering in the preamplifier:

$$ENC = \frac{N_{out}[V]}{S_{out}[V/e-]} = \frac{V_{noise}^{RMS}}{G} \tag{4.1}$$

with G expressed in V/e−. Considering the threshold dispersion a requirement for the ENC is:

$$T > n\,ENC \oplus T_{RMS}(x) \oplus T_{RMS}(t) \tag{4.2}$$

where T_{RMS} is the threshold variation during time (t) and across the matrix (x), and n corresponds to the number of σ of noise from which the threshold must set in order to obtain a firing probability lower than a certain value; taking up the example above, assuming a gaussian distribution, if n equal to 3.7, the probability that the pixel fires is less than 10^{-4}.

Because of the changing of the "real" threshold, the possibility of changing and adapting the setting parameters of the FE, both in time and in space is desirable: these parameters are usually set by Digital to Analog Converter (DAC) with a number of bits in a typical range of 3-7. Unfortunately DAC elements require a lot of space that may be not available on the pixel area; therefore, the FE parameters are typically global, which means that they are assigned for the whole chip, or they can be assigned for regions the matrix is divided into. The former case corresponds to TJ-Monopix1's design in which 7 bits are used for a total 127-DAC possible values, while the latter corresponds to the ARCADIA-MD1's one, where the number of bits depends on the particular parameter (for the threshold 6-bits DAC are used). Another possibility, for example implemented in TJ-Monopix2, is to allocate the space on each pixel for a subset of bits, then combining the global threshold with a fine tuning. If so, the threshold dispersion after tuning is expected to decrease depending on the number of bits available for tuning:

$$\sigma_{THR,tuned} = \frac{\sigma_{THR}}{2^{nbit}} \tag{4.3}$$

where σ_{thr} is the RMS of the threshold spread before tuning.

To measure the threshold and noise of pixels a possible way is to make a scan varying a known charge injected through a dedicated injection capacitor: the threshold corresponds to the value where the efficiency of the signal exceeds the 50%, and the ENC is determined from the inverse of the slope at this point. Assuming a gaussian noise, e.g. a noise whose transfer function turns a voltage δ pulse in a gaussian distribution, the efficiency of detecting the signal and the noise can be described with the function below:

$$f(x, \mu, \sigma) = \frac{1}{2}\left(1 + erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right) \tag{4.4}$$

where erf is the error function. Referring to 4.4 the threshold and the ENC corresponds to the μ and σ.

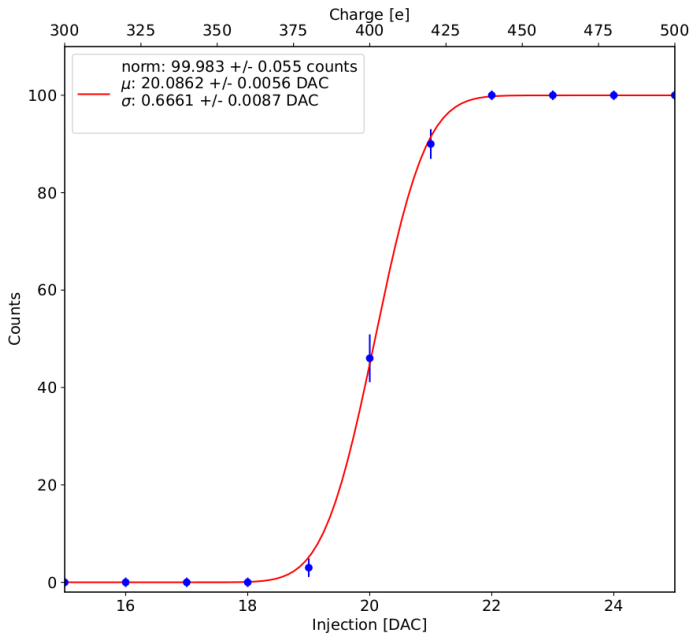


Figure 4.1: S-curve for pixel (10, 10) of the PMOS flavor (flavor B) with IDB fixed at 40 DAC. The conversion of charge injected from DAC to electrons has been performed using a nominal conversion factor of 20.3 e−/DAC

I used the injection circuit available on the chip to inject 100 pulses for each input charge for a fixed threshold. The charge injection is realized sending a voltage step V_{inj}(DAC) on an injection capacitance implemented at the input of the FE circuit (Q_{inj}=$C_{inj}V_{inj}$). Knowing the nominal value of the C_{inj} (230 aF for the PMOS flavor) one can calculate the conversion factor for the signal injected from DAC to e^-. Since the DAC are biased at 1.8 V, the Least significant Bit (LSB) corresponds to a voltage of 14.06 mV from which (for the PMOS flavor) the charge for LSB 1.4375 e/mV and the conversion factor therefore is 20.2 e/DAC. While this value is equivalent for all the PMOS flavor, the HV flavor is expected to have a different conversion factor, $\sim$ 33 e/DAC, because of the different injection capacitance. Besides the charge, also the duration and the period of the injection pulse can be set; it is important to make the duration short enough to have the falling edge during the dead time of the pixel (in particular during the FREEZE signal) in order to avoid the undershoot, coming at high input charge, triggering the readout and creating spurious hits. Since the injection circuit is coupled in AC to the FE, if the falling edge of the pulse is sharp enough to produce an undershoot, this can be seen as a signal.

Therefore I fitted the counts detected using the function in equation 4.4. Figure 4.1 shows an example of such fit for a pixel belonging to the flavor B with the register IDB, which sets the discriminator threshold in voltage, fixed at 40 DAC; in figure 4.2 are shown the 1D and 2D distributions of the parameters μ and σ of the fit found for the PMOS B flavor. Then I fitted the 1D-histograms with a gaussian function to found the average and RMS of the noise and the threshold across the matrix. The results for each flavor are reported in table 4.1; no relevant differences among the flavors have been observed regarding the noise, which results to be $\lesssim$15 e^-, while the threshold has been found to be

$\sim$400 e^- except for the PMOS C flavor, where, with the same FE settings, it is $\sim$540 e^-.

	PMOS A	PMOS B	PMOS C	HV
Threshold [e^-]	401.7±0.2	400.8±0.2	539.7±0.6	403.9±0.2
Threshold dispersion [e^-]	32.9±0.1	33.0±0.2	55.5±0.4	44.7±0.2
Noise [e^-]	13.01±0.06	12.26±0.07	13.9±0.1	11.7±0.1
Noise dispersion [e^-]	1.61±0.04	1.50±0.05	1.91±0.07	1.58±0.07

Table 4.1: Mean threshold and noise parameters for all flavor and their dispersion on the matrix.

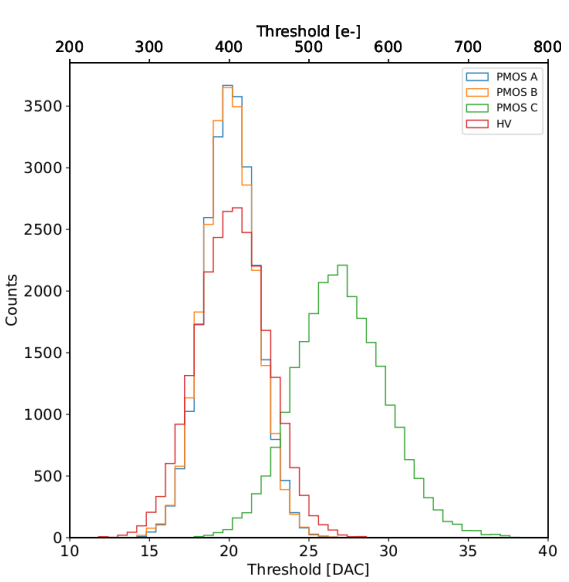
Although a slightly lower threshold is visible in the first biasing section (columns from 0 to 14) in the map in figure 4.2b, the threshold and noise are rather uniform across the matrix but a small systematic variation appears more evidently when using different IDB values. The systematic threshold variation seems connect with the column-section and has not a well established explanation, if not a relation with the biasing group. An interpretation could certainly be the transistor mismatch of the biasing DAC registers IDB and ICASN, which both adjust the effective threshold (ICASN regulates the baseline and in the presented measurement has been set at the minimum value, that is 0 DAC).

To verify the trend of the threshold as a function of the front end parameter IDB and find its dynamic range, I have performed different scans changing the FE register IDB. For each value I have injected the whole matrix and searched for the mean and the standard deviation of the threshold and noise distributions. The results are shown in figure 4.3: the blue points are the mean threshold found within the matrix, while in green is shown the width (threshold ± trhreshold dispersion) of the threshold distribution, i.e. the threshold dispersion. While the threshold increases at higher IDB, the ENC decreases of $\sim$4 e^-,which is $\sim$1/3 of the noise at IDB=40 DAC. Then, to evaluate the operation and the occupancy of the chip at different threshold I have checked how the number of pixel masked changes with the threshold (fig.4.4). The masking algorithm I have used search for pixels with rate >10 Hz and mask them. In our standard condition a very low noise hit rate of $\sim$3 Hz is intentionally achieved masking a dozen of pixels on the whole flavor.

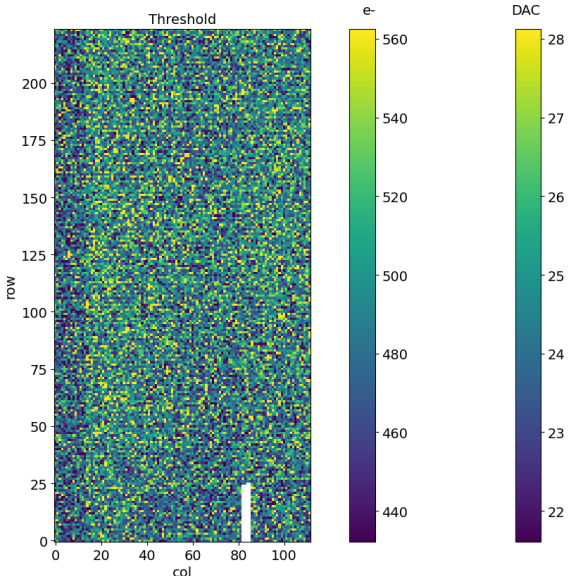
4.1.2 Linearity of the ToT

I have already stated in chapter 3 that TJ-Monopix1 returns an output signal proportional to the charge released by a particle in the epitaxial layer, which is the Time over Threshold; the ToT is saved as a 6-bit variable and therefor its dynamic range is 0-64, which corresponds to 0–1.6µs assuming a clock frequency of 40 MHz. When a pulse is longer than 1.6 µs the counter rolls back to zero and there is no way to distinguish that charge from a lower one with the same ToT: that is the rollover of the ToT (fig.4.5a).

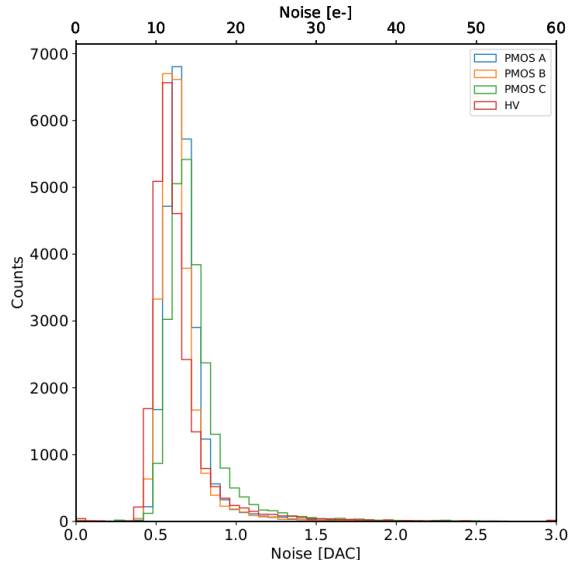
In order to associate the ToT (in range 0-64) to the charge, a calibration of the signal is necessary. The output of TJ-Monopix1 is approximately a triangular pulse, resulting in



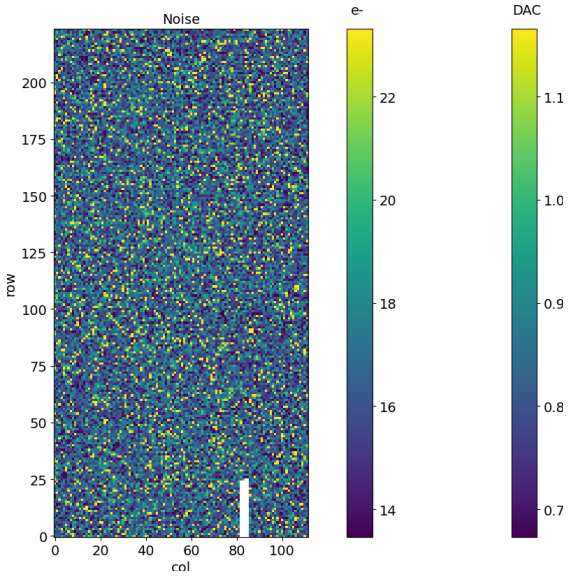
(a) Histograms of the threshold



(b) Map of the threshold



(c) Histogram of the noise



(d) Map of the noise

Figure 4.2: The threshold and the noise have been found fitting the s-curve of all flavor with IDB fixed at 40 DAC. The white pixels have the injection circuit broken

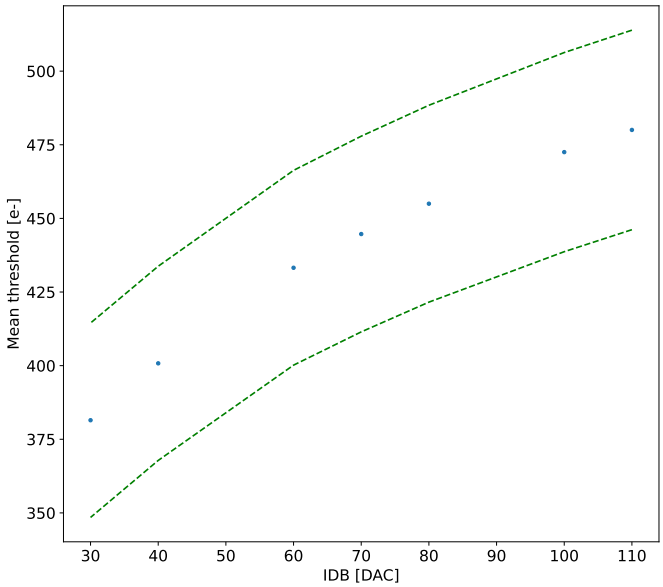


Figure 4.3: Flavor PMOS (B) with Psub-Pwell biased at -6 V. Threshold converted in electrons (using the nominal conversion factor 20.2 e/DAC) vs the register which sets the threshold, IDB.

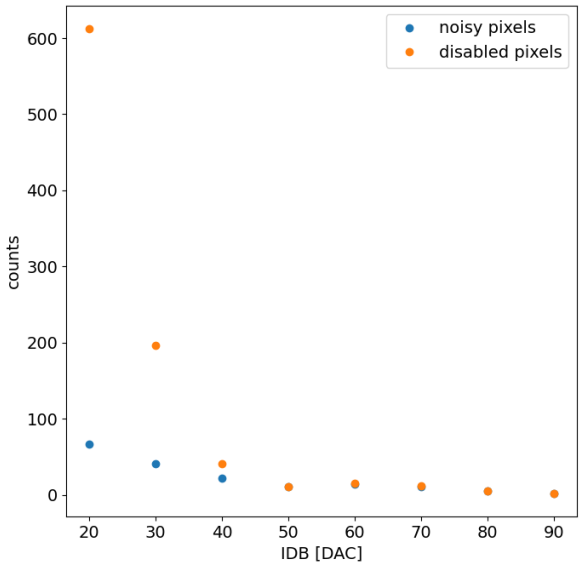
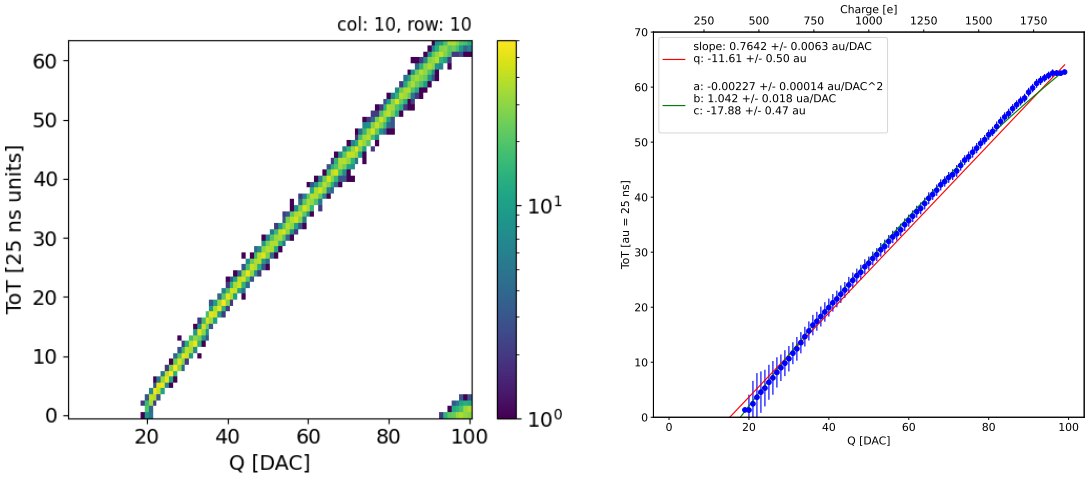


Figure 4.4: Number of pixels masked at different IDB: beacuse of the algorithm for masking not all the pixels masked are noisy. In blue the number of noisy pixels, while in orange the number of pixels disabled

a linear relationship between ToT and charge:

$$Q\,[DAC] = \frac{(ToT\,[au]\ -\ offset\,[au])}{slope\,[au/DAC]} \tag{4.5}$$



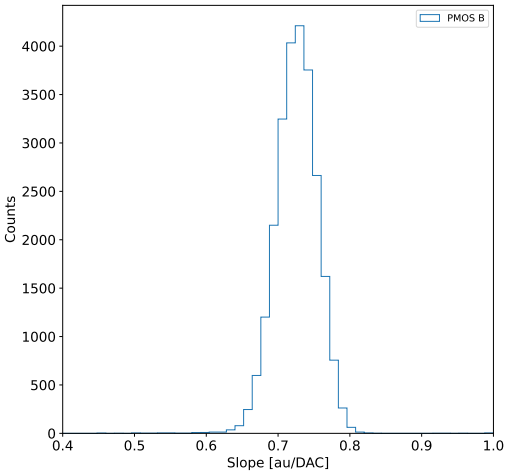
(a) (b)

Figure 4.5: The figures refer to pixel (10, 10) of the PMOS-reset flavor B with IDB fixed at 40 DAC. (a) Histogram of the injection pulses: the ToT is in range 0-64 since it is represented by 6 bit, so when achieving the 64 it rolls over back to the zero. (b) Mean ToT vs the the charge: the mean has been calculated removing the rollover hits.

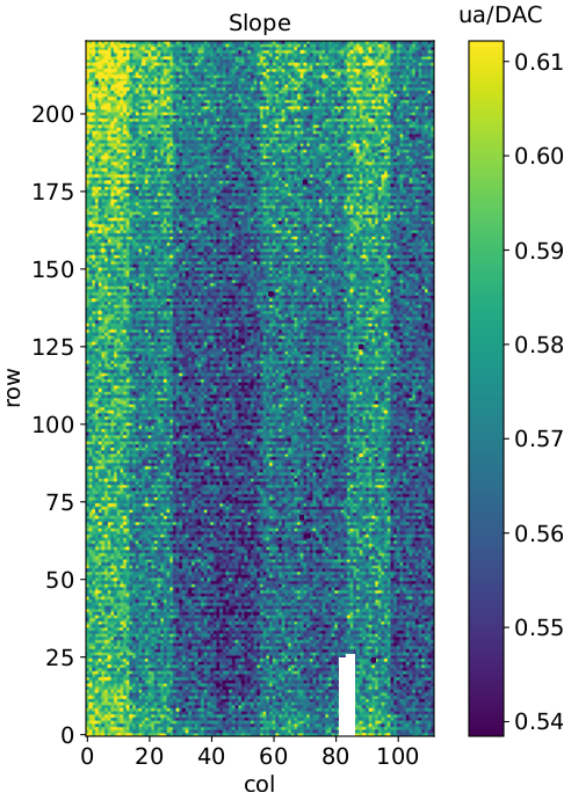
where $slope$ and $offset$ are the fitted parameters of the calibration. It is important to keep in mind that the main application target of TJ-Monopix1 is in the inner tracker detector of HEP experiments, then the main feature is the efficiency, then a rough calibration of the signal to charge is fine. The ToT information can be used both to better reconstruct the charge deposition in cluster in order to improve the track resolution, and for particle identification, through $\frac{dE}{dx}$, especially for low momentum particles which do not reach the dedicated detectors.

The study of the output signal has been possible via the injection: I fitted the ToT versus the pulse amplitude injected for all the pixels within the matrix. In figure 4.5b there is an example of fit for a pixel belonging to the flavor B, while in figure 4.6 there are the histograms and the maps of the parameters of the line-fit for all flavors with IDB fixed at 40 DAC. Here a difference among the biasing section appears: since the slope of the ToT is related to the gain of the preamplifier (increasing the gain also increases the ToT), the mismatch is probably due to the transistor contributing to the amplification stage.

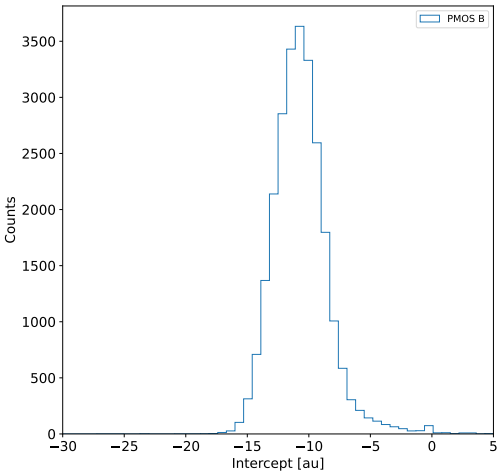
I fitted the average ToT of all the pulses recorded as a function of the pulse amplitude; data affected by rollover have been removed in order to avoid introducing a bias in the mean values. In figure 4.5b are shown both the fits with a line (red) and with a second order polynomial (green): at the bounds of the ToT range values deviate from the line model. Since the deviation is lower than 1% and it only interests the region near the 0 and the 64, in first approximation it is negligible.



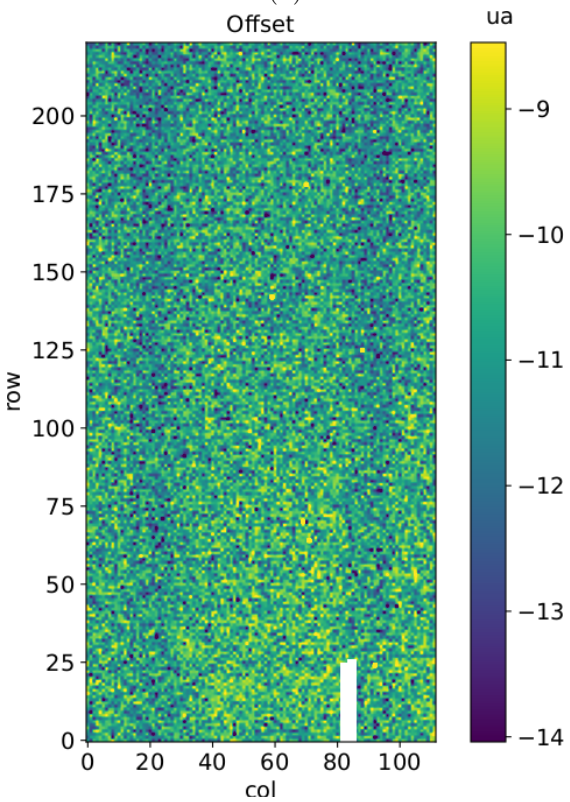
(a)



(b)



(c)



(d)

Figure 4.6: Histograms of the calibration parameters, slope (a) and offset (b), found fitting the ToT with a line, for the flavor B and with IDB fixed at 40 DAC. Maps of the calibration parameters, slope (a) and offset (b), found fitting the ToT with a line, with IDB fixed at 40 DAC.

4.1.3 Calibration of the ToT

For the calibration of the Time over Threshold signal I have used both the injection and a radioactive source in order to fix an absolute scale of the charge detected. The calibration is based on the following assumptions:

- 0 DAC corresponds to 0 e^-
- a ToT of 1 clock count corresponds to a charge produced in the bulk equal to the threshold

Considering that the charge injected in the FE depends on the value of the C_{inj} (Q_{inj}=$C_{inj}V_{inj}$) which is different from pixel to pixel, the true charge injected does not correspond to the nominal value expected assuming C_{inj}=230 aF. Accordingly to that, a measurement of the injection capacitance provides both an absolute calibration of C_{inj} and a conversion factor K to have a correspondence of the DAC signal in electrons. K and C_{inj} are defined respectively as:

$$K\,[e^-/\mathrm{DAC}] = \frac{1616\,[e^-]}{Q\,[\mathrm{DAC}]} \tag{4.6}$$

$$C_{inj}\,[\mathrm{F}] = K\,[e^-/\mathrm{DAC}]\,\frac{1.6\,10^{-19}\,[\mathrm{C}]}{14.06\,[\mathrm{mV}]} \tag{4.7}$$

where 1616 e^- is the number of electrons produced in the detector by the calibration source (Fe55) and 14.06 mV is the voltage value of a DAC (LSB). K is expected to be 20.2 e/DAC, assuming the nominal value of C_{inj} equal to 230 aF, and where 1616 is the expected number of electrons produced by the calibration source used, Fe55. Fe55 is en extremely important radionucleotide in the calibration of X-ray spectrometers, proportional counter and scintillator detector since it emits two two X-photons during the electron capture decay: the first one (K_α) at 5.9 keV with an emission probability of 24.4% and the second one (K_β) at 6.5 keV with a probability of 2.86%. The K_α photon, the one with the higher emission probability, which does photoelectric effect in silicon, has an absorption length $\lambda\sim$29 µm, then the probability of being assorbed in the 25 µm thick epitaxyal layer is $\sim$0.58%. The photo-electron emitted has an energy equal to the photon, so recalling that the mean energy needed to produce a couple electron-vacuum is 3.65 eV, the signal produced by the Fe55 source is expected to be 1616 e^-.

In figure 4.7 are shown two histograms of the ToT spectrum of the Fe55 source for two different pixels. The peak on the right corresponds to the events with complete absorption of the charge in the depleted region, while the long tail on the left to all the events with partial absorption due to charge sharing among neighbors pixels. In order to reduce the consistent charge sharing, the pixel dimension in TJ-Monopix2 has been reduced down to 30×30 µm^2. The events on the right side of the higher peak, instead, corresponds to the charge released by K_β photo-electrons. Looking at the histograms for pixel (30, 185) and (30, 68) respectively at top and bottom of figure 4.7, a significant difference in the peak to tail ratio leaps out, which can be related with the position of the pixel in the matrix. In particular, because of a different charge collection property, pixels in the upper part of the matrix (rows 112-224) have a more prominent peak, while in pixels in the lower part (rows

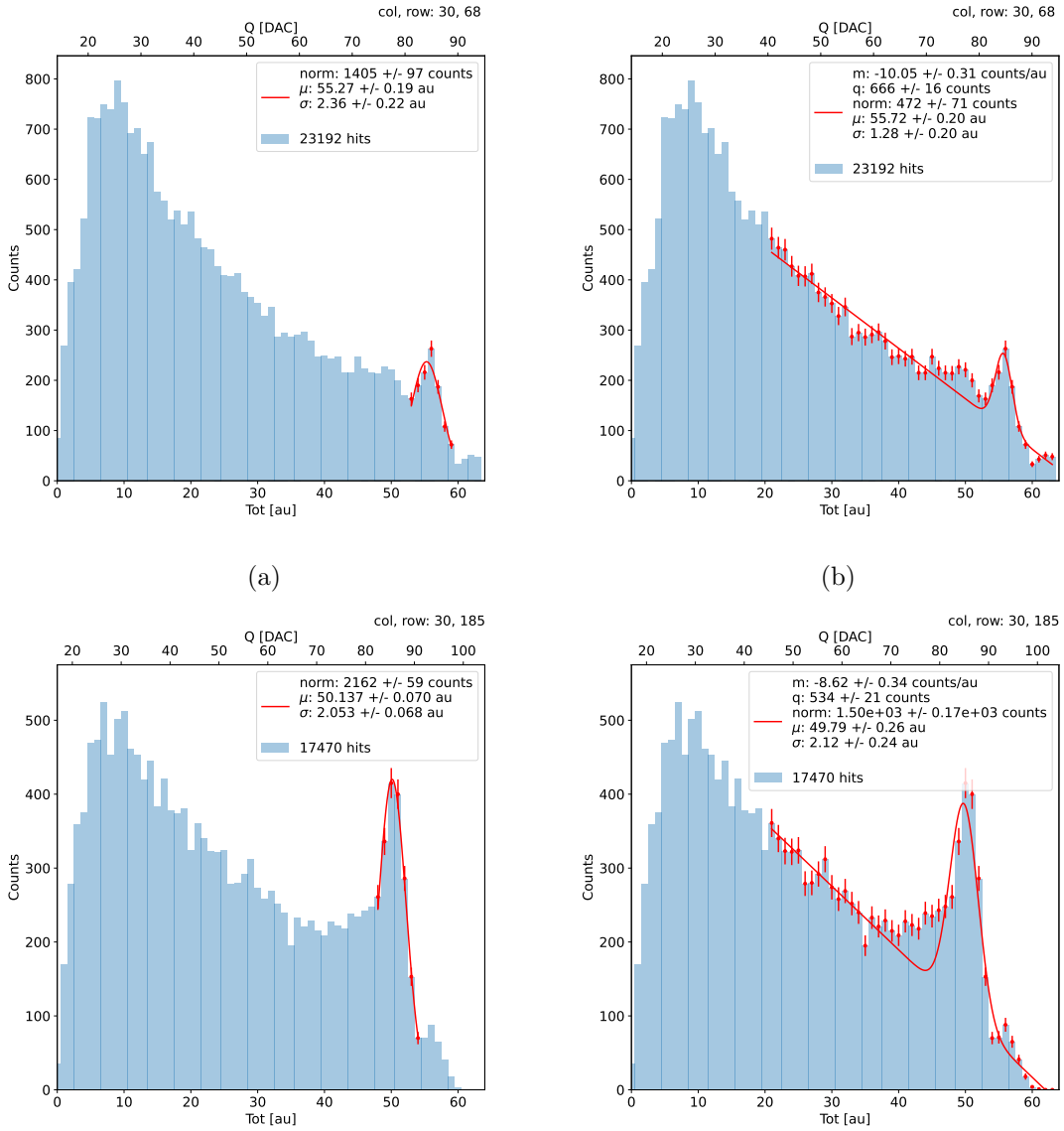
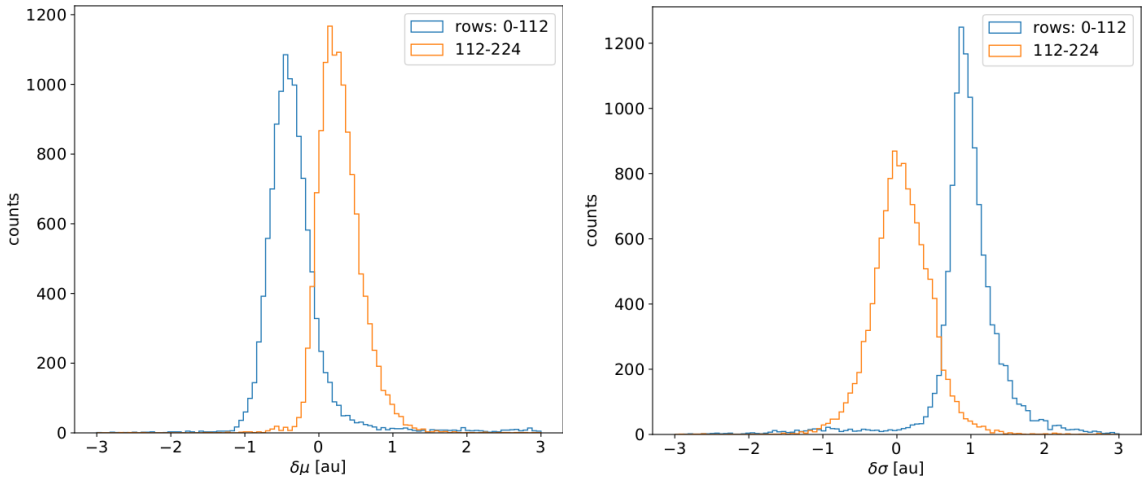


Figure 4.7: Both strategies 4.8 and 4.9 of fitting the Fe55 peak are shown for two pixels on the matrix: the (a) and (b) refers to pixel (30, 68) which has a FDPW, while the (c) and (d) refers to pixel (30, 185) which has a RDPW. The fit has been performed using the bins colored by red.



(a) (b)

Figure 4.8: Difference between the parameters μ and σ obtained with the gaussian fit and those obtained with a gaussian plus a line. When μ <0 the fit with function 4.8 is generally worse (the peak is shifted to the left); when $\sigma < 0$, the fit with 4.9 is worse (larger sigma).

0-111) there is a higher partial absorption. Indeed, as discussed in section 3.1.1, there is a distinction in the structure of the low dose-epi layer among the rows, in particular pixels in rows 112-224, which have a Reduced Deep P-Well (RDPW), are supposed to have a higher efficiency in the pixel corner; this seems to verified considering that a more prominent peak means a higher probability of complete charge absorption.

For the calibration I needed to establish the peak position, then to do that I fitted the ToT histogram of each pixel: I tested two different fit functions which I report below

$$f(x, N, \mu, \sigma) = \frac{N}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{(x-\mu)}{\sigma})^2} \tag{4.8}$$

$$f(x, m, q, N, \mu, \sigma) = m\,x + q + \frac{N}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{(x-\mu)}{\sigma})^2} \tag{4.9}$$

The additional linear term in equation 4.9 is meant to model the tail due to incomplete charge collection and prevent it from introducing a bias in the fitted peak position. For this reason, when I fitted with eq.4.9, I selected a larger region of the spectrum compared to the fit with eq.4.8, for which I used only a small reagion around the peak. The optimal fit range was chosen in both cases through an iterative routine: for the fit with eq.4.9 it starts from an interval including all the ToT above 20 clock counts and progressively reduces it by increasing the left boundary; for the fit with eq.4.8, it starts from an interval of 5 bins around the expected peak position and reduces the interval of 1 bin at each iteration.

Even if the difference in the peak position between the two fit strategies is not really relevant for the purpose of the calibration, being of the order of 0.8-1.5% (4.8), it still introduces a systematic bias towards lower values due to the contribution of the tail. Indeed, we know that the sharp edge on the right must correspond to the case of complete

absorption of the photon, so that, in general, the closest to the edge the fitted peak position is, the better the fit is. Besides the peak position, a poor fit tends also to overestimate the peak width. Even looking at the χ^2, the fit function 4.8 seems to be the better choice, except for a set of pixels in the lower part of the matrix, the ones with a FDPW and lower efficiency in pixel corner.

The resolution of the detector, which is expected to be determined by the statistical fluctuations in the number of charge carries generated in the detector as well as by the ENC, can be compared to the observed Fe55 peak width. Ideally:

$$\sigma_{Fe} = \sqrt{ENC^2 + F \times N} \tag{4.10}$$

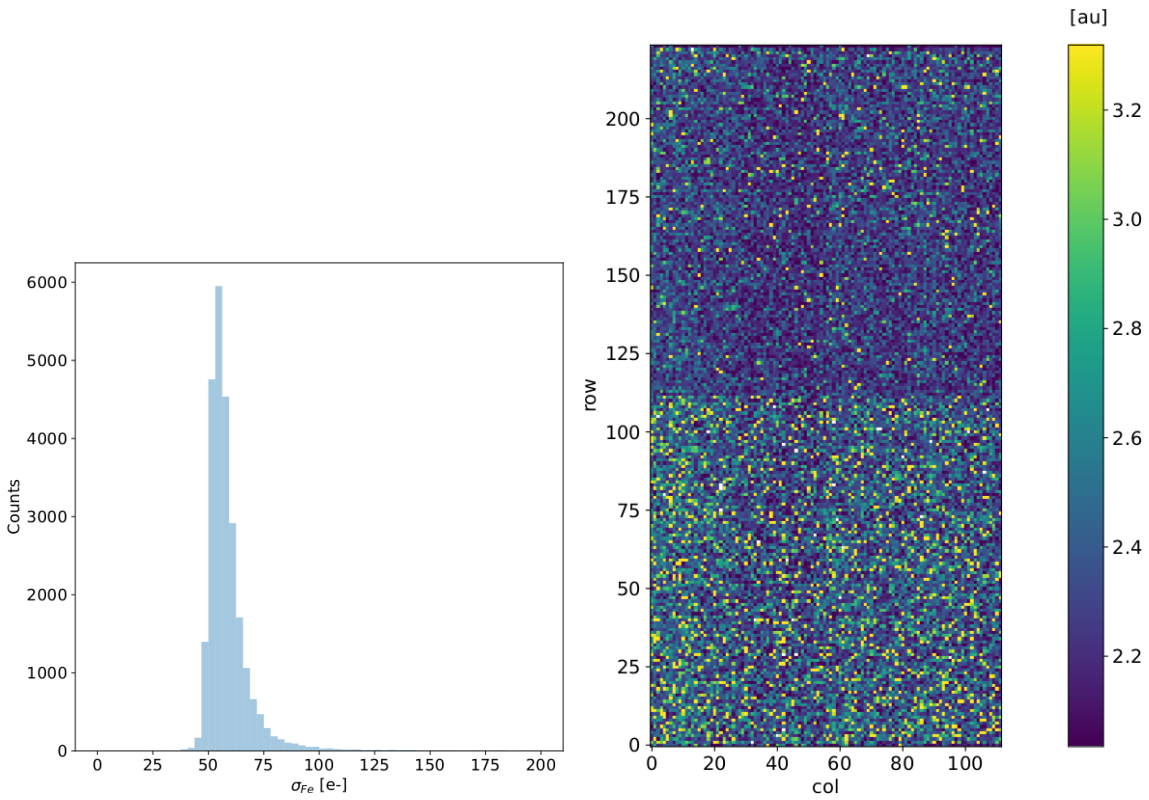
Since the number of e/h pairs produced in the sensor is 1616, recalling that the Fano factor F for a silicon detector is 0.115 and that the ENC measured with the injection is $12\,e^-$, the σ_{Fe} is expected to be $\sim 18\,e^-$. Looking at figure 4.9 the resolution achieved with the Fe55 source seems to be much worse. A contribution we have not taken into account but is certainly relevant is the systematic overestimation of the standard deviation of the Fe55 peak: this, as I already explained, is principally due to the high background of incomplete charge collection, which broadens the fitted peak. Although, this effect is not sufficient to justify a such high peak width value. 2D maps of the value of the capacity and of the conversion factor found are shown in 4.10. The evident stripe-structure in the matrix shows an evident correlation among the same row; the same structure, which is also visible in the slope map of the calibration of the ToT (fig.4.6b), may be related with the structure of the bias lines.

An attempt of calibrating the HV flavor, which is the most different from the PMOS B flavor, has been performed; however, because of a loss of signal of $\sim$50% caused by the higher capacitance, we have been unable to identify the Fe55 peak, and then the calibration of the ToT in electrons has been impossible. Moreover the HV flavor did not seem to work properly, as we have observed that all pixels sometimes fire one time simultaneously. For these reasons unfortunately a complete characterization of the HV flavor has not been possible: in fact, since it has the most particular FE compared to the other PMOS flavor, a comparing the results would be particularly interesting. An example of Fe55 spectrum collected with the HV flavor is shown in figure 4.11.

4.1.4 Changing the bias

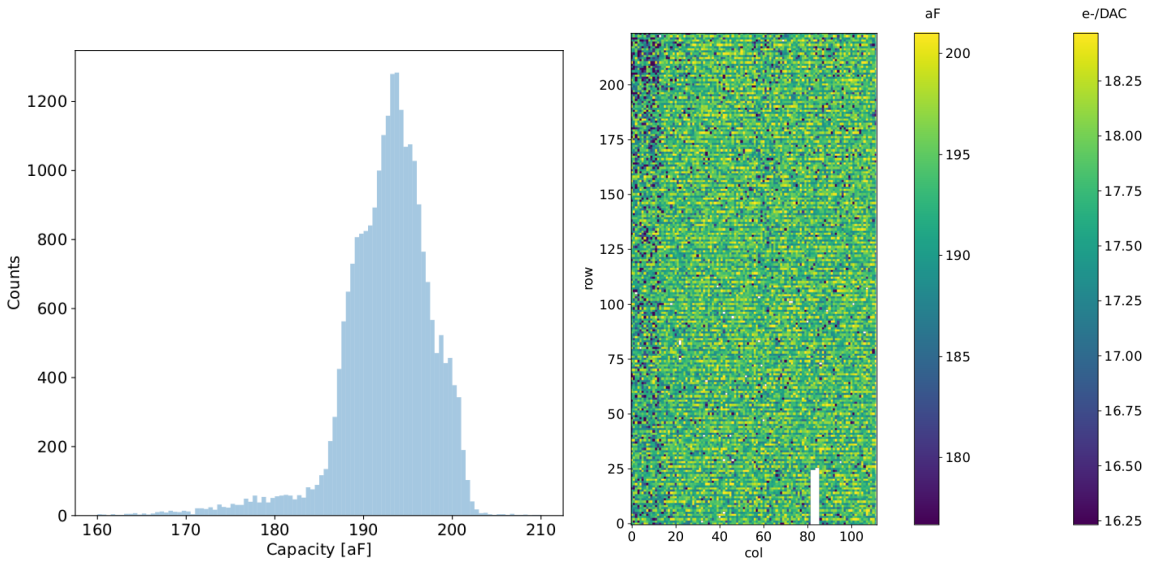
In order to study the behavior of the sensor as a function of the bias, I performed several injection scans in different sensor bias conditions. The thickness of the depletion region has to be considered an important parameter affecting the signal efficiency, and in particular it affects the charge released by a particle which crosses the sensor, since the signal is proportional to the thickness of the epitaxial layer.

Another important benefit of operating the FE with higher bias is the reduction of the capacitance of the collection diode C_{in}, and the corresponding increase of the gain of the first stage of the FE, which goes as $\sim 1/C_{in}$ (as explained in sec.3.1.2).



(a) (b)

Figure 4.9: (a) Histogram and (b) map of the Fe55 width found by the fit with function 4.8 converted in electrons using the calibration. In the map a clear difference between the two parts of the matrix can be be distinguished: in particular, as already stated, the rows with RDPW have a better resolution. It worth noting that the pixels which in the above maps appear disable, here do not show any problem. This prove that they have a problem in the injection circuit but not in the sensor and in the FE.



(a) (b)

Figure 4.10: Histogram (a) and map (b) of the calibrated capacity of the injection circuit.

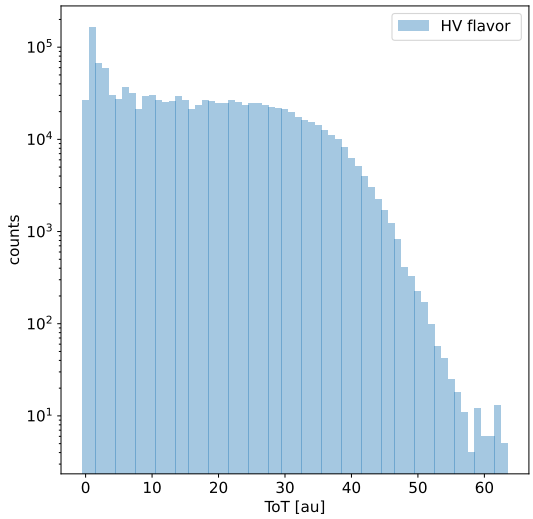


Figure 4.11: Fe55 spectrum with the HV flavor. No peak has been identifyied because of the loss of the signal due to the higher input capacity.

An example of expected change of gain with increased bias is shown in fig.4.12, that reports the output voltage amplitude and gain for the PMOS and HV flavours for chips characterized by other groups. Given that the chip under examination has a gap in the low dose epi-layer (3.3), we were not able to change independently the bias of the substrate (PSUB) and of the p-well (PWELL), but they must be kept at the same value, differently from other chips of the same submission. Lowering the bias, the depletion region is expected to narrow and the efficiency to reduce, especially in the pixel corner, thus raising the threshold and the noise and decreasing the slope as a consequence of the reduction in the gain.

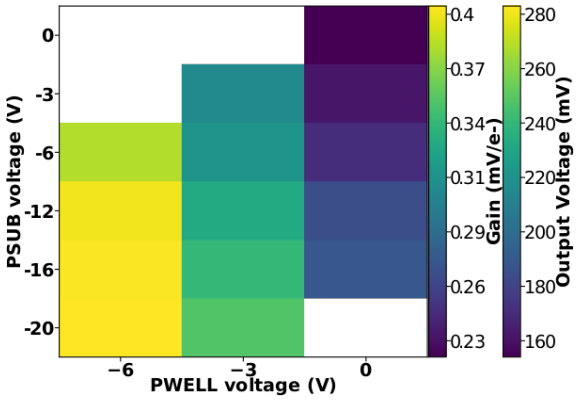


Figure 4.12: Output voltage amplitude and gain with respect to the p-well and p-substrate voltage in the case of the PMOS reset front-end.

In order to test the behavior of the chip when not completely depleted, I have performed an injection scan with PSUB/PWELL bias at 0 V, -3 V and -6 V (results in tab.4.2); passing from -6 V to a smaller depletion at 0 V, the slope of the output signal reduces of

	-6 V	-3 V	0 V
Threshold [DAC]	20 ± 2	21 ± 2	24 ± 2
Noise [DAC]	0.61 ± 0.08	0.62 ± 0.08	0.82 ± 0.1
Slope [au/DAC]	0.73 ± 0.03	0.71 ± 0.03	0.57 ± 0.02
Offset [au]	-11 ± 2	-11 ± 2	-11 ± 2

Table 4.2: The errors of the values are the standard deviations of the corresponding distributions. To convert DAC values to electrons can be used the conversion factor ∼20 e−/DAC (nominal) or ∼18 e−/DAC (measured).

∼1/4, which is smaller than the reduction in gain reported in figure 4.12, that is ∼1/3. Morover the increase in the threshold and noise at smaller bias is due to the fact that diffusion becomes a competing collection mechaninsm, with all the consequences described in section 2.5. Figure 4.13 shows the values of the K_α peak position, the normalization of the events above the peak that is the normalization coming from the gaussian fit of the peak, and the rate as a function of the PSUB/PWELL biases. These quantities have been normalized to their value at -6 V, which is then defined as the reference condition. As expected with reduced bias two effects occur: firstly the position of the Fe55 peak moves to lower values due to a lower gain (the reduction of ∼1/3 is in agreement with the value measured in 4.12), secondly the number of events in the Fe55 peak and rate both become smaller since the depletion region is reduced. So, what happens is the decrease of both the number of events with full collection (of the 1616 e^- from the Fe55 photon) in a single pixel, which contributes to the normalization of the peak, but also the reduction of the events with charge sharing among neighbours pixels.

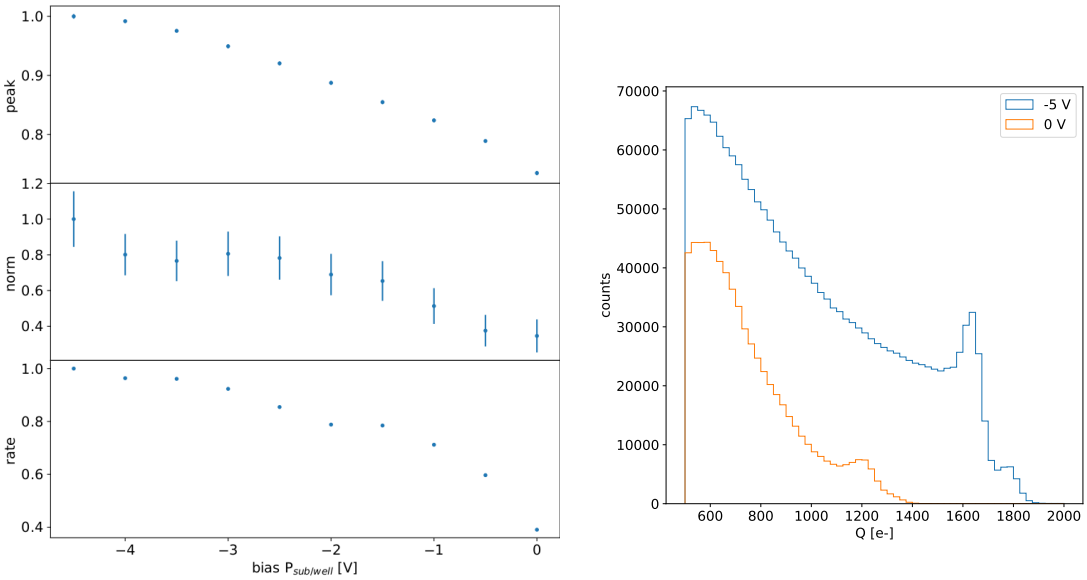
4.1.5 Measurements with radioactive sources

In order to completely validate the operation of the whole sensor[1], I have performed several acquisitions with radioactive sources, specifically Fe55 and Sr90Y, which is a β^- emettitor with electron endpoint at 2.3 MeV, and cosmic rays. In particular I used the data collected with Sr90 and cosmic rays, to study charge sharing and events with more than one hit. I define *cluster* the ensemble of all the hits with the same timestamp. This is obviously a coarse requirement, but it gave me the opportunity of using a simple and fast clustering algorithm, which is fine when the random coincidence probability is neglibile. Defining R_1 and R_2 as two any events rate (can be both signal of the same source or one can be a source signal rate and the other the noise rate), and τ as the dead time of the detector, the random coincidence rate can be found:

$$R_{coinc} = R_1 \times R_2 \times \tau \tag{4.11}$$

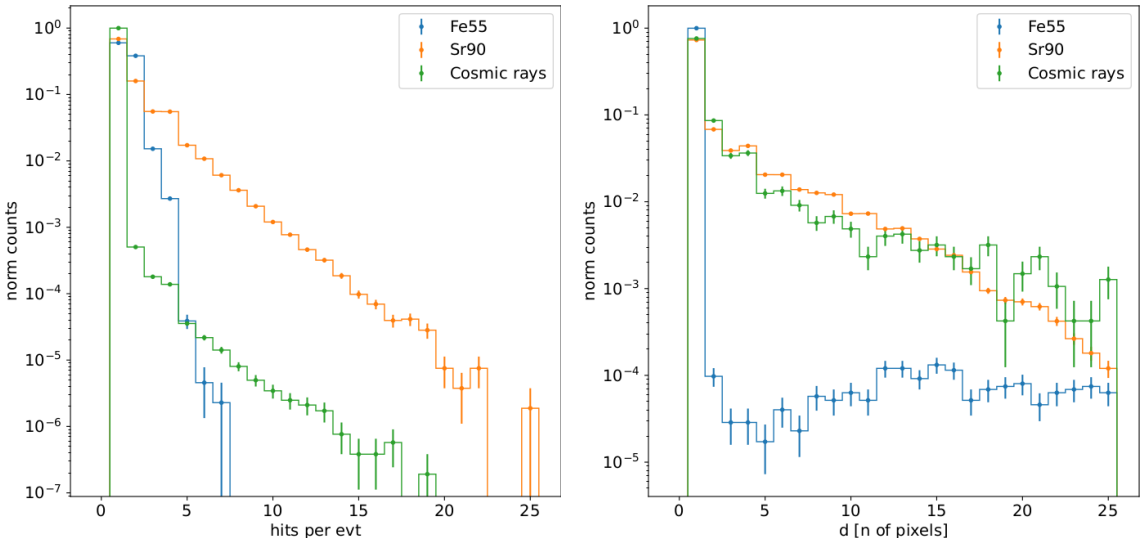
As I am going to prove in the next section, the dead time strictly depends on the occupancy of the matrix, even though we can assume a dead time of ∼1 µs, which corresponds to

[1] As I will discuss in chapter 5.2 these measurements serves also as a reference for the spectrum observed at the test beam



(a) (b)

Figure 4.13: (a) Peak position, peak amplitude and rate as a function of the bias. Since during the collection of the whole data the source has been moved, it is not guaranteed that it has always had a repositioning in the same exactly place, then small the fluctuation of the rate along the decreasing trend are accettable. The peak position and amplitude are estimated by fitting the spectrum with a gaussian in the region around the peak. (b) Fe55 spectrum at different $P_{sub/well}$ bias. The ToT values have been calibrated as explained in section. 4.1.3.



(a) (b)

Figure 4.14: (a) Distribution of the number of hits per event with different sources. (b) Dimension of cluster defined as eq.4.1.5. Compared with the Sr90 and the cosmic rays, the Fe55 d distribution is characterized by a clear discontinuity around d=5. The very thin peak at 0 corresponds to the effective clusters, while the long tail at bigger d is principally made of random coincidences distant on the matrix.

the mean dead time per pixel. However, if in an event a particle hits two different pixels producing a cluster, the total dead time simply doubles. Since the measured rate on the whole matrix of noise, Fe55, Sr90 and cosmic rays are ∼Hz, 3.3 kHz, 40 Hz and ∼10 mHz[2], the random coincidence probability are neglibile except the one of two Fe55 events, which is 11 Hz.

In figure 4.14 I report the histograms of the number of pixels in the cluster and of the dimension of clusters, defined in terms of the max and min coordinates on the matrix as:

$$d = \sqrt{(y_{max} - y_{min})^2 + (x_{max} - x_{min})^2} \tag{4.12}$$

Looking at the shape of the histogram of the dimension, generally the Sr90 and the cosmic rays produce bigger clusters and hit a higher number of pixels, a trend that can be explained considering that the Fe55 photoelectron is less energetic than the Sr90 electron and cosmic rays. A sample of hitmap of events produced by the three different sources is shown in figures 4.15, 4.16 and 4.17. In figures 4.19, 4.20, 4.21 are shown the distributions per different cluster dimension events, of the charge collected by a single pixel (figures on the left) and the charge collected by summing the charge collected by the pixels within the cluster (figures on the right). Since the noise rate is comparable with the cosmic rays and Sr90 ones, I have removed the single pixel events which are separately shown in figure 4.18; although we cannot identify and select only the noise events, these distributions, and especially the cosmic rays one, are expected to be mostly populated by noise events. The distributions have a peak around the threshold, which is compatible with the fact that the

[2]The cosmic rays rate at the sea level is expected to be $\sim 1/\mathrm{cm}^2/\mathrm{s}$

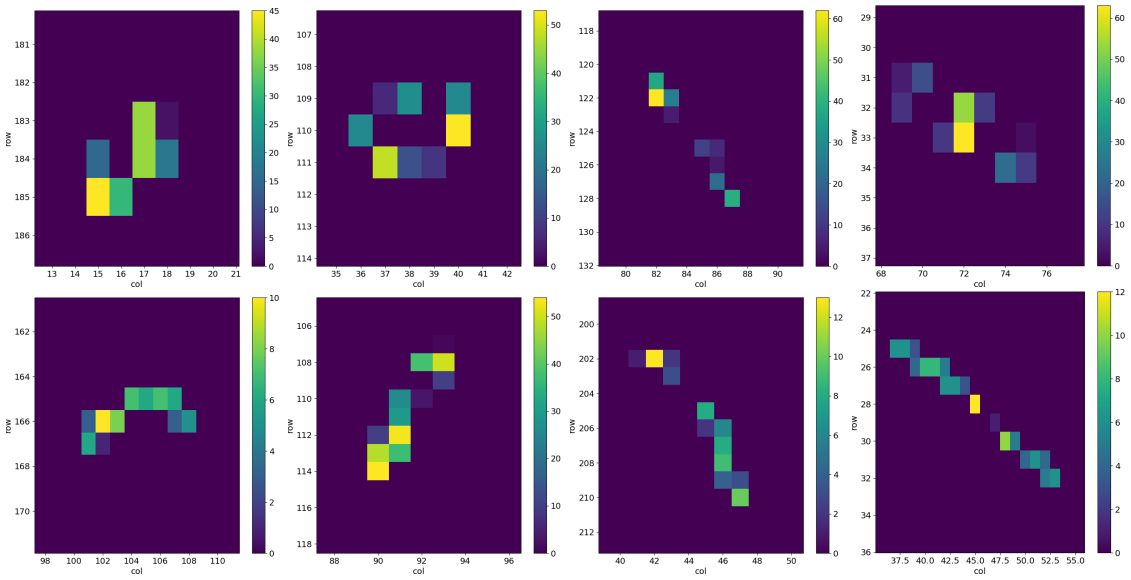


Figure 4.15: 2D histograms of the ToT in different events in an aquisition of cosmic rays.

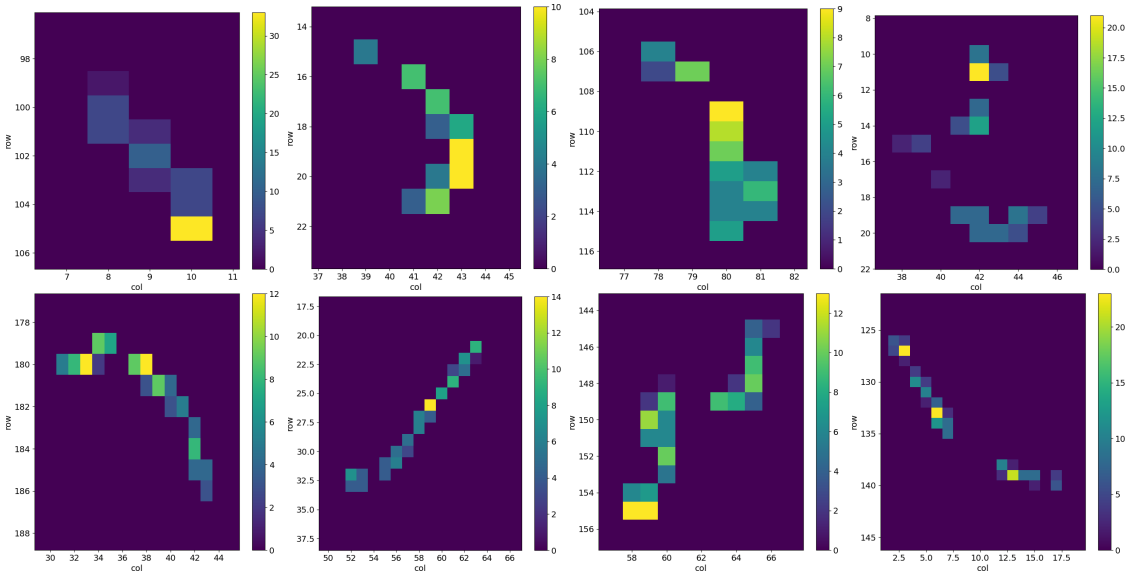


Figure 4.16: 2D histograms of the ToT in different events in an aquisition of Sr90.

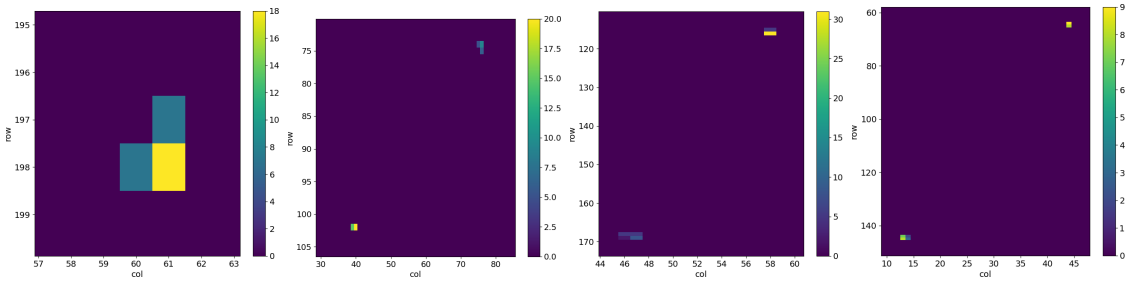


Figure 4.17: 2D histograms of the ToT in different events in an aquisition of Fe55

noise events typically have a low ToT.

Looking at the spectra of Sr90 instead (fig:4.20), the maximum of the distribution of the cluster charge seems to follow a linear dependence on the number of pixels hit

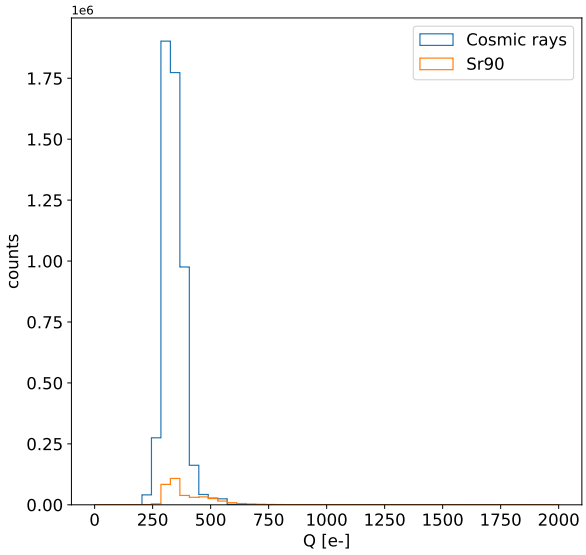


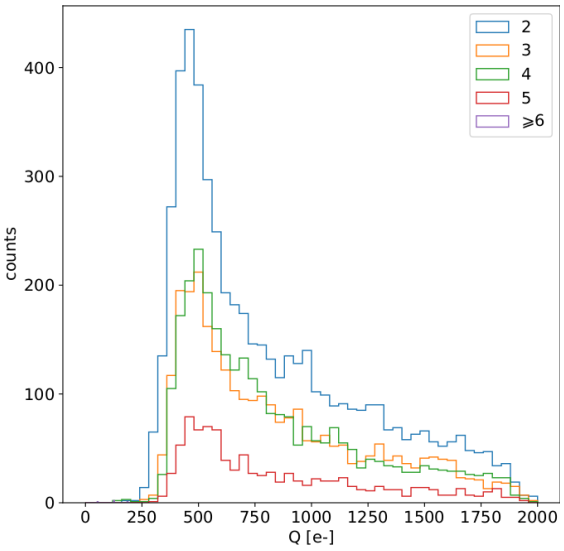
Figure 4.18: Histograms of the charge released in the pixels in events in which only a single pixel turns on.

Pixel per evt	Measured [e-]
2	950 ± 30
3	1450 ± 30
4	2050 ± 30
5	2450 ± 30

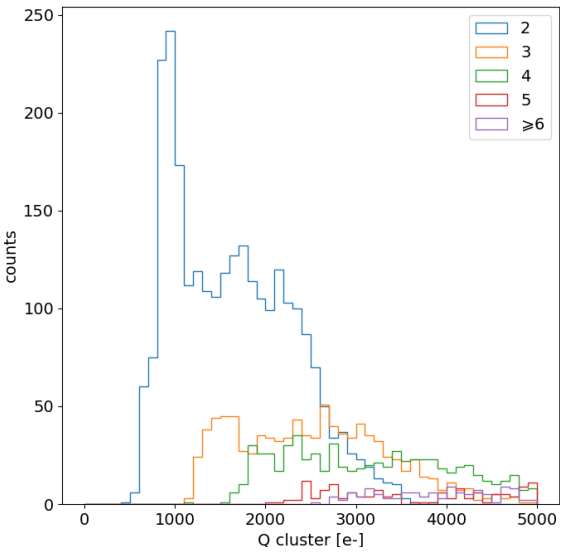
Table 4.3: Position of the maximum of the distributions in figure 4.20b of the summed charge released in the clusters depending on the number of pixel in the cluster.

(tab.4.3); this can be accepted as a first approximation considering that the pitch (36 µm and 40 µm) depends on the direction, and the epitaxial layer thickness (25-30µm) are comparable. However a more accurate model which takes into account the impact angle of the particle and charge sharing among neighbours pixels should be developed for a more precise comparison.

Regarding the Fe55, the bump in the cluster spectrum at $\sim$1616 e^- corresponds to photons which had converted at the boundary of nearby pixels thus sharing their charge among them. Starting from 4-pixels clusters the peak moves to the right: this is due to the fact that the cluster with more than 3 pixels are principally random coincidence events Fe55-Fe55 or Fe55-noise. Recalling that the noise typically just exceeds the threshold and then has low ToT, the peak position in the spectrum 4.21b of 4-pixel cluster can be explained admitting that one of the four pixel is a noise signal. The shoulder on the right, instead, which have an edge at about 3200 e^- corresponds to the events with coincidence of two photons. Looking at the charge on the single pixel spectrum (fig.4.21a), instead,

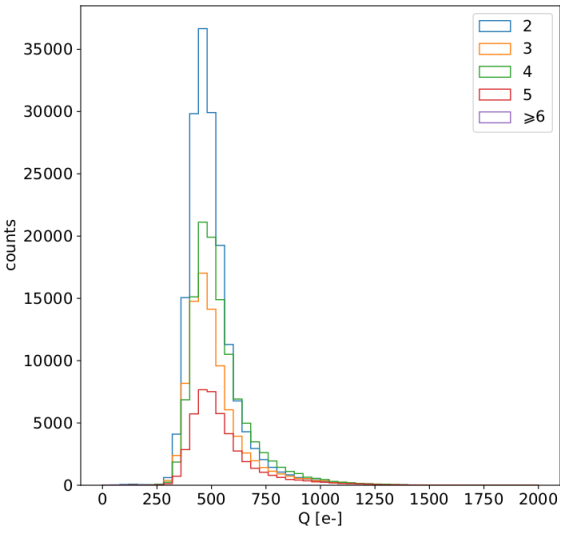


(a) Distribution of the charge collected on individual pixels

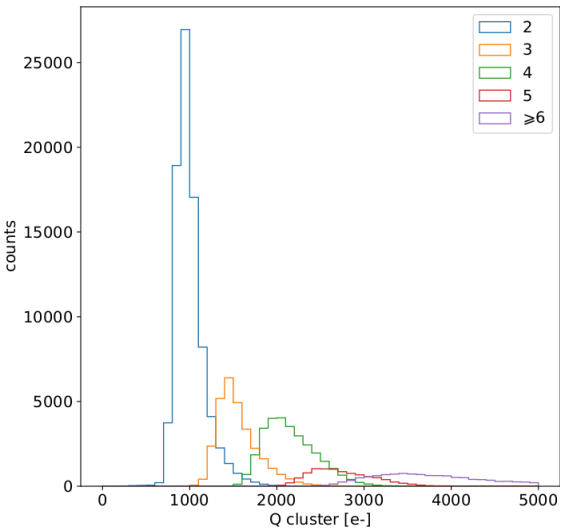


(b) and of the charge collected on clusters

Figure 4.19: Acquisition of cosmic rays with the PMOS B flavor with the same FE setting of the calibration (in particular IDB=40 DAC)

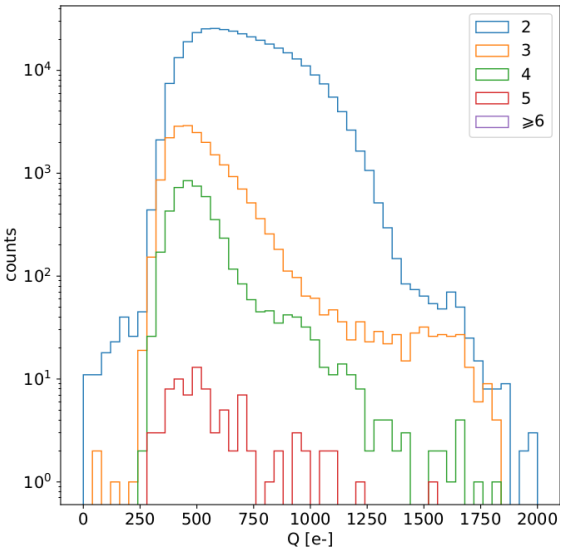


(a) Distribution of the charge collected on individual pixels

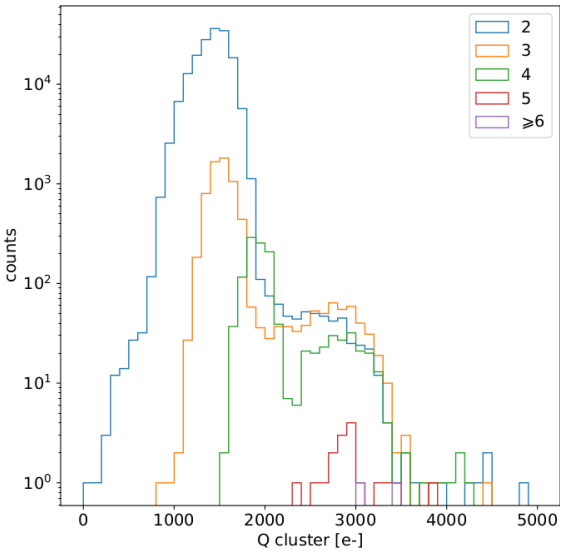


(b) and of the charge collected on clusters

Figure 4.20: Acquisition of the Sr90 with the PMOS B flavor with the same FE setting of the calibration (in particular IDB=40 DAC)



(a) Distribution of the charge collected on individual pixels



(b) and of the charge collected on clusters

Figure 4.21: Acquisition of the Fe55 with the PMOS B flavor with the same FE setting of the calibration (in particular IDB=40 DAC)

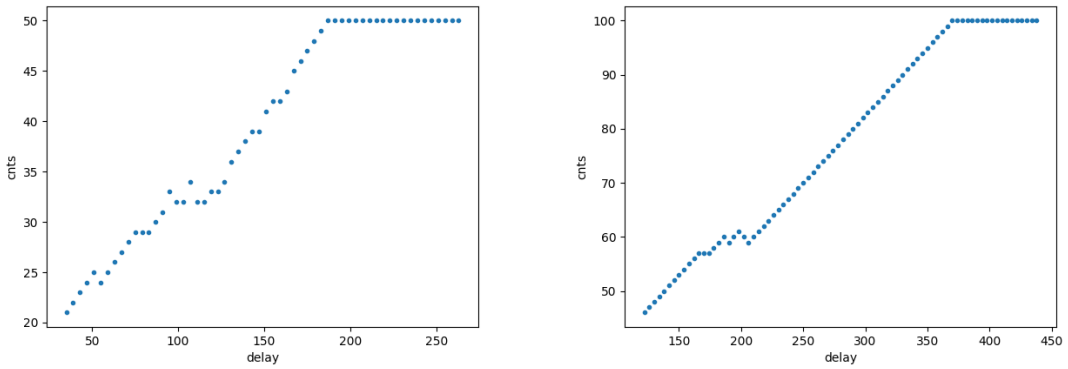
a small bump can be seen around 1616 e^-: these events correspond to photons which released almost all the charge on one pixel.

4.1.6 Dead time measurements

The hit loss is due to analog and digital pile up: the first one occurs when a new hit arrives during the pre-amplifier response to a previous event, the second instead when the hit arrives while the information of the previous hit has not yet been transferred to the periphery. Since the pre-amplifier response has a characteristic time ∼ToT, the dead time τ_a introduced by it will be at most 1.6 µs; using the IRESET and VRESET FE parameters the reset time can be lowered down, but as explained in section 3.1.2 it must be longer than the preamplifier charateristics time in order to not cut the signal. Regarding the latter contribution instead, since only one hit at a time can be stored on the pixel's RAM, until the data have completed the path to get out, the pixel is paralyzed. Moreover since there is no storage memory included on TJ-Monopix1 prototypes, the digital dead time τ_d almost corresponds to the time needed to trasmit the data-packets off-chip.

The exportation of data from pixel to the EoC occurs via a 21-bits data bus, therefore only one clock cycle is needed and the dead time bottleneck is rather given by the bandwidth of the serializer which trasmits data off-chip from the EoC. In our setup the serializer operates at 40 MHz, thus to transmit a data packet (27-bit considering the addition of 6 bits to identify the double-column at the EoC) at least 675 ns are needed. For what we have said so far, the R/O is completely sequential and therefore is expected a linear dependence of the reading time on the number of pixels to read:

$$\tau = 25\,\text{ns} \times (\alpha\, N + \beta) \tag{4.13}$$



(a) Distribution of the charge released on pixels (b) and of the charge released on clusters

Figure 4.22: Efficiency vs the DELAY parameters. (a) I made a scan injecting 5 pixels with 50 pulses for each DELAY configuration and (b) 10 pixels with 100 pulses for each DELAY

where α and β are parameters dependent on the readout chain setting.

To test the linearity of the reading time with the number of pixels firing and to measure it, I have used the injection circuit which allows me choosing a specific hit rate: I made a scan injecting a fixed number of pulses and each time changing the number of pixels injected. Indeed the injection mode allows fixing not only the amplitude of the pulse, which corresponds to the charge in DAC units, but also the time between to consecutive pulses (DELAY). The hit rate then corresponds to 25 ns/DELAY.

Unfortunately a high random hit rate on the matrix cannot be simulated by the injection because of the long time (∼ms) needed to set the pixel registers of the injection; then I was forced to specify at the start of the acquisition the pixels to inject on, and for convenience I chose those on a same column. In figure 4.22 is shown the dependence of the efficiency on the DELAY parameter in two different cases. For the 5 pixels example the efficiency goes down the 90% at a DELAY of ∼185 clock counts, which corresponds to 4.625 µs and to a rate of 216 kHz, while in the 10 pixels example, the efficiency goes under the 100% at ∼380 clock counts, which corresponds to 9.5 µs and to a rate of 105 kHz.

From the efficiency curves I have then looked for the time when the efficency decreases. In figure 4.23(a) is shown the dead time per pixels as a function of N with different R/O parameters configuration, the meaning of which is explained in chapter 3.1.3. The default value suggested by the designer of the chip are reported in table 4.4; moving too much the readout parameters from the default ones, the readout does not work properly, and no hits can be read at all. The problem probably comes from the firmware setting of the readout which are specially fixed for our chip. The single pixel readout time is indipendent of its position in the matrix, and it is equal to 37.5±1 clock counts. However if many pixels are fired, the dead time τ_d depends on the position because the reading sequence goes from row 224 to row 0, and from column 0 to column 112, making the pixel on the bottom right corner the one with the longest dead time.

Furthermore to test that there is no dependence of the digital readout time from the

Parameter	Value [DAC]	Value [µs]
START_FREEZE	64	1.6
STOP_FREEZE	100	2.5
START_READ	66	1.65
STOP_READ	68	1.7

Table 4.4: Default configuation of the R/O: START and STOP refer to the begin and the end of the respective signals starting from the TE of the hit.

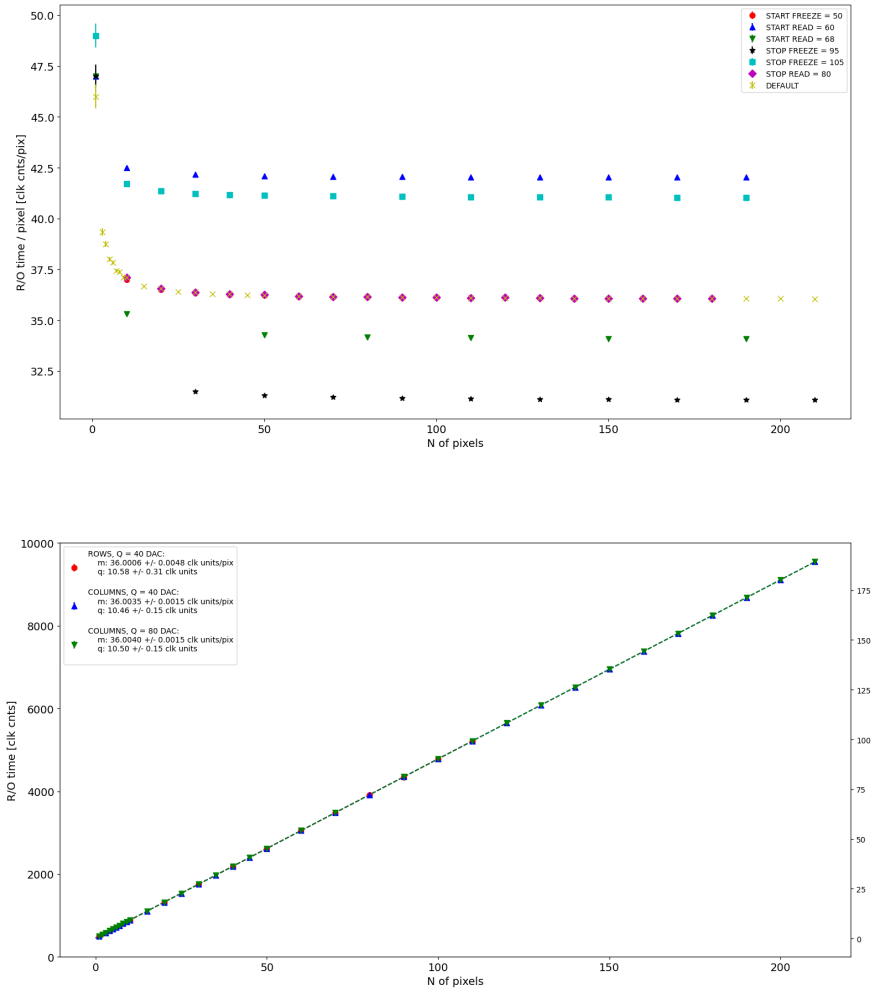


Figure 4.23: (a) Readout time per pixel as a function of the number of pixel injected obtained with different FE setup. (b) Readout time as a function of the number of pixels injected obtained injecting pulses with amplitude of 80 DAC (green), of 40 DAC on the same row (red) and on the same column (blue).

charge of the pulse, I have tried to change the amplitude of the pulse injected, but the parameters found were consistent with the default configuration ones. No difference in the α and β coefficients has been observed between the two cases. Referring to eq.4.13, the

factor α is proportional to the difference (STOP_FREEZE - START_READ), while the offset β lies between 5 and 15 clock counts.

The readout time found by this test is so long because in the prototypes no parallelization of the informations (with the instroduction of more serializer for example) and no storage memory are included; this feature are typically added in the final prototypes. An example closely linked to TJ-Monopix1 is OBELIX: it will include on the chip a storage buffer to optimize the dead time and to keep a low occupancy even at high fluence.

4.2 ARCADIA-MD1 characterization

Unfortunately the characterization of MD1 has not yet been completed because the first chip we received was not fully functional, so that we have only been able to perform a few electrical and communication tests, in order to assess the operations of the FPGA and the breakout board (BB). At the moment this document is being written, a fully operational chip has been available only for a week, due to delays in the extraction and the bonding of the wafer; an initial characterization and testing of the new chip is currently undergoing in the clean room of INFN, and here I will show some preliminary results of that work.

The problem with the damaged chip manifests itself when the chip is biased; in particular when the HV voltage is lowered down to 0 V, the sensor requires too much power and a too high current draw sets. We have discussed the problem with the designers of the chip, who helped us indentifying the motivation of the malfunction: the chip has been glued using too much conductive tape and hence has a short-circuit between the sides and the back, which makes the biasing impossible. Unfortunately, since both the sensor and the FE require at least -10 V to work properly, no measurement was possible except the acquisition of the noise in the FE circuit.

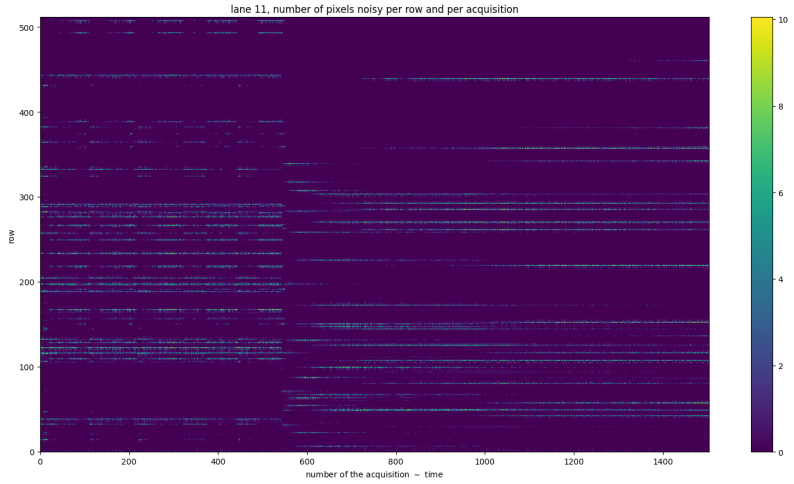


Figure 4.24: Noise in the front end circuit depending on the bias road across the matrix was recorded.

The second chip we received is a minid2, that is a "mini demonstrator" from the second

submission. The two share the same characteristics, but the minid2 is smaller than the MD1, in particular it only have 32×512 pixels, instead of 512×512.

Up to now we used the injection circuit ($C_{inj} = 2.325\,fF$) in order to make a threshold scan on a few pixels: differently from the TJ-Monopix1's characterization, where we performed a scan changing the injection charge of the pulse, with the minid2we have changed the threshold (whose register is VCASN) instead, keeping the charge of the pulse fixed. For each threshold we injected 100 pulses of amplitude 10 µs. The dependece of the efficiency on the threshold for two pixels is shown in figure 4.25. Even if the behavior is reasonable, as the efficiency becomes higher when the threshold is reduced, it is possible that the bias (-50 V) is not enough to full deplete the sensor, since the counts does not reach the 100% steadily.

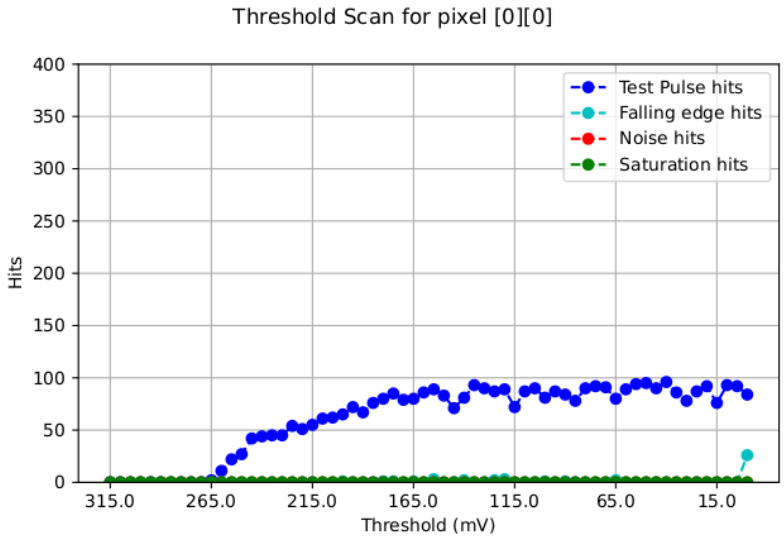
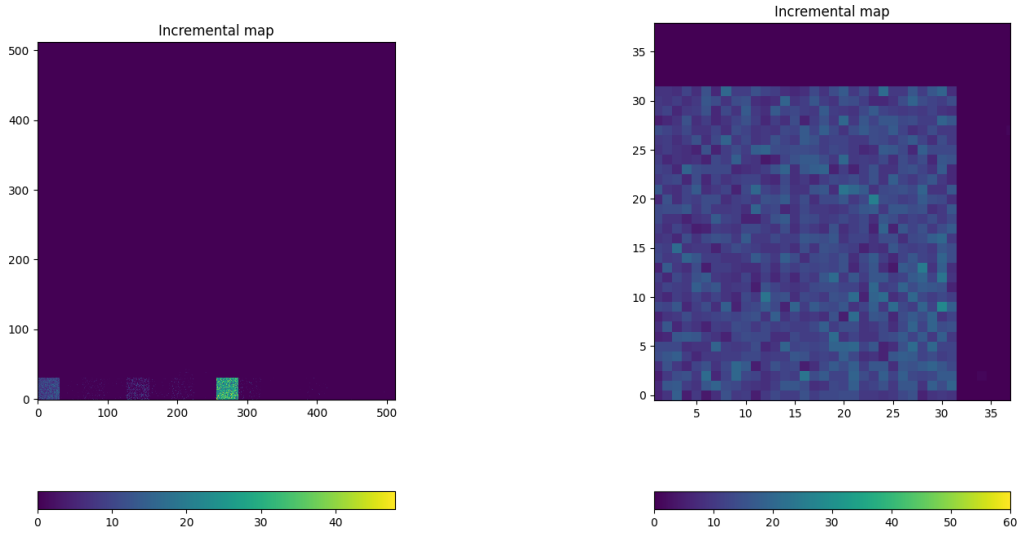


Figure 4.25: Threshold scan on the pixel (0,0). The sensors is polarized with ΔV=-50 V.

The SNR, the ENC and the threshold dispersion on the matrix are expected to be respectively ∼90, $3\,e^-$ and ∼$35\,e^-$ dor a detector with an expected capacitance of about 7 fF. The injection capacity is expected to be ∼2.325 fF, and in this condition the the minimum and maximum signals generated are respectively 0.08 fC and 2.6 fC.

Substantial differences have been observed with VCASN=40 DAC in both the efficiency and the threshold among the sections; this suggests that with this particular FE configuration there is a big threshold dispersion on the matrix. The hitmap of an acquisition with the Fe55 source is shown in figure 4.26: the whole MD1 matrix with only the bottom region (32 rows) working is represented in (a), while in (b) there is a zoomed hitmap. The rate seen within the region 8 (green region in the figure (a)) is compatible with the rate of the same radioactive source measured with TJ-Monopix1, that it ∼3.3 kHz.

Looking to the Sr90 acquisitions (fig.4.27) many clusters and tracks can be immidiately distiguished, confirming what observed with TJ-Monopix1. More tests will be performed in the future to fully characterize ARCADIA-minid2.



(a) (b)

Figure 4.26: Fe55 acquisitions with VCASN=40 DAC. (a) All the matrix 512×512 is plotted even if the minid2 has only the rows in range 0-32. (b) A zoom on the first section (col 0-32).

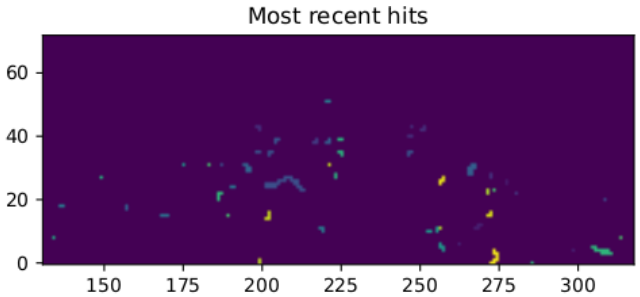


Figure 4.27: Sr90 acquisition with VCASN=40 DAC. The different colours are related with the time of arrival of the hits: in yellow the most recent hits, while in blue the old ones.

CHAPTER 5

Test beam measurements

During August 2022 I participated to a beamtest at the Santa Chiara hospital in Pisa, where a new accelerator designed both for medical research and R&D on FLASH-RT, called "ElectronFlash", was installed a few months ago. The test-beam was meant to test TJ-Mopopix1 at high dose rate with a focus on investigating the possibility of the application in radiotherapy. Despite this particular device does not seem to match the requirements required for that application, especially regarding the readout time, the measurements have helped us charaterizing the setup for future activities, and also have given us the possibility of a complete charaterization of the chip. In this chapter I will describe the setup used and some preliminary results.

Given that, in medical physics, the *dose* is the standard metric used to characterize the beam, because of its obvious relation with the damage caused in the patient, I am going to explain the meaning of it from the point of view of the instrumentation. In fact, when interacting with measuring systems, a more common and useful metric is the *rate* or the *fluence* of particles. The conversion between the two quantities can be find starting from the definition of dose: it is defined as energy per unit area deposited in a material as a result of an exposure to ionizing radiation. Assuming total absorption of electrons in water, defined by law as the reference medium, the dose can be expressed as:

$$D[\mathrm{Gy}] = \frac{NE[\mathrm{eV}]}{\rho[\mathrm{g/cm^3}]A[\mathrm{cm^2}]x[\mathrm{cm}]} \tag{5.1}$$

where N is the number of incoming particles, E is their energy, x is their range, A is the section of the beam and finally ρ is the density of the absorbing medium.

After having applied the conversion of the energy from eV to J and noticed that $\mathrm{E}/\rho\mathrm{x} \approx \mathrm{dE}/\rho\mathrm{dx}$ for MIP electrons and roughly corresponds to the stopping power S of electrons of energy E in water, and defining N_A as the fluence of particle on an area A (beam section), a simple estimation of the dose released is:

$$D[\mathrm{Gy}] = 1.602\ 10^{-10}\, N_A[\mathrm{cm^2}]\, S[\mathrm{MeVcm^2/g}] \tag{5.2}$$

$\bar{D}$	Dose rate (mean dose rate for a multi-pulse delivery)	0.005-10000 Gy/s
$\dot{\bar{D}}$	Intra pulse dose rate (dose rate in a single pulse)	0.01-1 10^6 Gy/s
DPP	Dose in a single pulse	0.04-40 Gy
PRF	Pulse repetition frequency	1-350 Hz
t_p	Pulse width	0.2-4 µs
n	Number of pulses	single/pulse train

Table 5.1: The parameters that can actually be set by the control unit are the PRF, DPP, t_p and n (in particular the modality of single irradiation or pulse train), while the other changes consequently.

Then, for 9 MeV electrons, whose stopping power in water[1] is 2.17 MeVcm^2/g, a dose of 1 Gy corresponds to a fluence of 2.9 10^9cm^2; if we assume a beam section of 10 cm, then the number of particle expected at the exit of the accelerator is 9.1 10^{11}.

5.1 Apparatus description

In order to shield the environment from ionizing radiation, the accelerator is placed in a bunker inside the hospital. The bunker has very thick walls of concrete and both the control units of the accelerator and of the detector are placed outside in a neighboring room.

5.1.1 Accelerator

The ElectronFlash accelerator, fabricated by S.I.T. - Sordina IORT Technologies S.p.A, is an electron Linear Accelerator (LINAC) with two energy configurations, at 7 MeV and 9 MeV, and it can reach ultra high intensity (over 5000 Gy/s) while keeping the possibility of accessing many different beam parameters and changing them independently from each other, a characteristic that makes it almost unique worldwide and which is fundamental for research in FLASH-RT, both for the medical aspects and for the studies on detectors. The accelerator implements the standard beam structure used in RT with electrons (fig. 1.7), that is a macro pulse divided in many micropulses; the parameters used to set the dose and their range of values settable by the control unit is reported in table 5.1.

The accelerator is also equipped with a set of plexiglass applicators with diameters in range from 1 cm to 12 cm and a collimator that can be used as is needed shaper to produce a squircle (between square and circle) shape. The plexiglass applicators must be fixed to the gun during the irradiation and are needed for producing, via the scattering of electrons with it, an uniform dose profile (fig.5.1) which is desired for medical purposes.

[1] Water is the reference medium for dose measurements because of it is equivalent-tissue

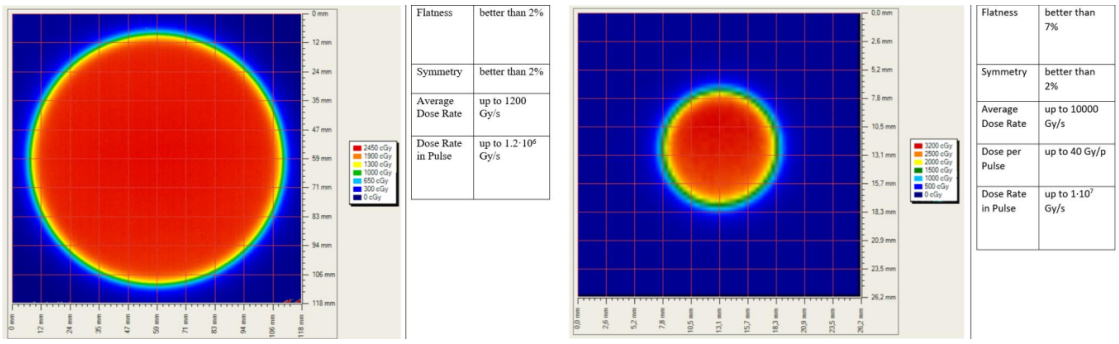


Figure 5.1: Two example of x-y isodose curves for two different plexiglass applicators, 10 cm and 1 cm respectively, reported by the producer in the manual with the specific of the accelerator (S.I.T. - Sordina IORT Technologies S.p.A.). With the smaller applicator the dose rate in pulse is comparatively higher.

DPP [Gy]	$N_{acc.\_exit} \times 10^9$ [cm^2]	N_{on_collB} 10^5 [cm^2]
1	2.88	11.52
0.6	1.73	6.92
0.07	0.20	0.8
0.04	0.12	0.48

Table 5.2: To obtain $N_{acc.\_exit}$ I have used the equation 5, while to obtain N_{on_DUT} I have taken into account the attenuation factor due to the collimator A.

5.1.2 Mechanical carriers

The tested detector consists in one chip, the Device Under Test (DUT), mounted on a board and connected to the FPGA with the same arrangement of figure 3.8. These boards have been positioned vertically in front of the plexiglass gun on a table specifically built for the testbeam. The three boards have been enclosed in a box of alluminium with a window on the DUT and with the required holes at the side to enable the biasing via cables and the connection with the DAQ provided via ethernet cable. A trigger signal coming from the control unity and synchronized with the pulses emitted from the beam was also sent to the FPGA. This digital signal cannot be considered a real trigger, since the TJ-Monopix1 prototype has been designed to be triggerless, but its Time of Arrival (ToA) has allowed the reconstruction of the correct timing during the analysis.

In order to reduce the particle flux on the sensor, two alluminium collimators have been fabricated: one has been positioned at the plexiglass gun exit while the other in front of the DUT. The collimators are t=32 mm thick and have a diameter d equal to 1 mm: assuming a beam divergence bigger than d/t=1/32 = 1.8°, which is the case, the collimator at the plexiglass gun output was supposed to work as a point source and to reduce the rate on the DUT of a factor at least 4 10^{-4}. In table 5.2 are reported, as a function of the Dose Per Pulse (DPP) setted by the control unit of the accelerator,the number of electrons which exit from the gun, the number of electrons which are expected arrive on the DUT if the

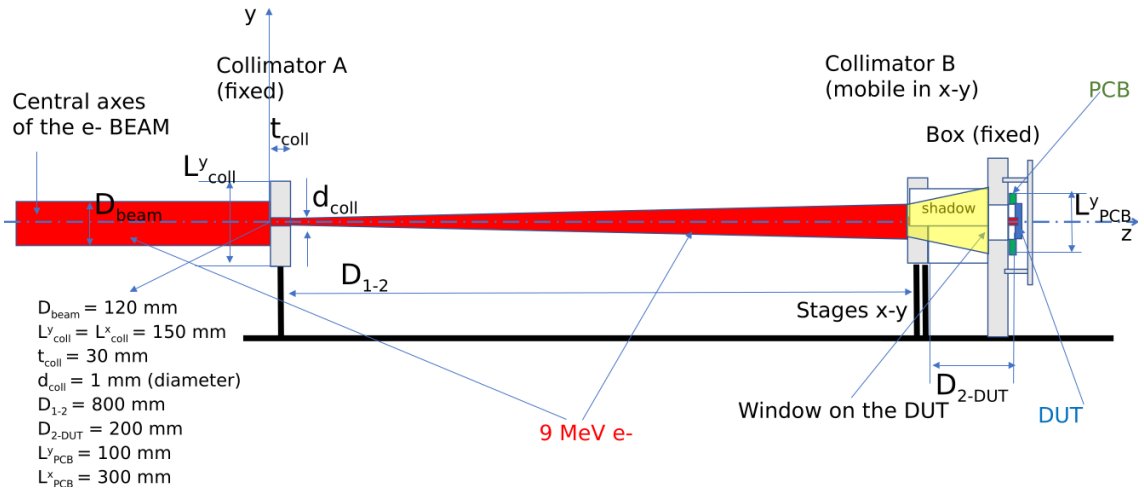


Figure 5.2: Scheme of the setup at the beamtest.

collimator B is not mounted. To obtain the rate on pixel, N_{on_DUT} must be divided by 4 µs.

The second one, located near the DUT, was instead supposed to shield the sensor from the electrons which passed the first one, except for a region of 1 mm^2 configurable using micrometers screws. It must be said that this arrangement of the collimators was not optimized. Simulations performed after the beam test indicate that multiple scattering in air plays an important role and the back of the envelope calculation of the flux was not correct.

5.2 Measurements

Because of the dead time of TJ-Monopix1, it is not possible to resolve the bunch substructure and almost no pixel can read more than a hit per bunch. This is unfortunately a major limitation that prevents operating the sensor as dosimeter, since the dead time per pixel depends on the location on the readout priority chain and for each pixel $\lesssim$1 µs are needed. Assuming a pulse duration of 4 µs, only a few pixels at the top of the priority chain (placed at the upper left on the matrix) can fire a second time, as they can be read a first time before the end of the pulse and then can be hit again.

Since resolving the single electron track is impossible, a way this sensor could be used in such context is reducing its efficiency and taking advantage of the analog pile up and of the linearity of the analog output (ToT), in order to see a signal produced not by the single particle but by more electrons. Reducing the efficiency and the sensibility of the sensor is essential in order to decrease the high charge signal produced in the epitaxial layer and mitigating the saturation limit: the smaller the output signal produced by a particle, the higher the fluence the detector can cope with. There is an obvious limit in this context that is the ToT rollover; indeed, the signal stops giving information when this value has been overridden and is no more bijective. With the standard configuration of the FE parameters and the epitaxial layer completely depleted, a MIP produces a charge at the limit of representation with a 6-bit ToT; to obtain smaller output signals one can

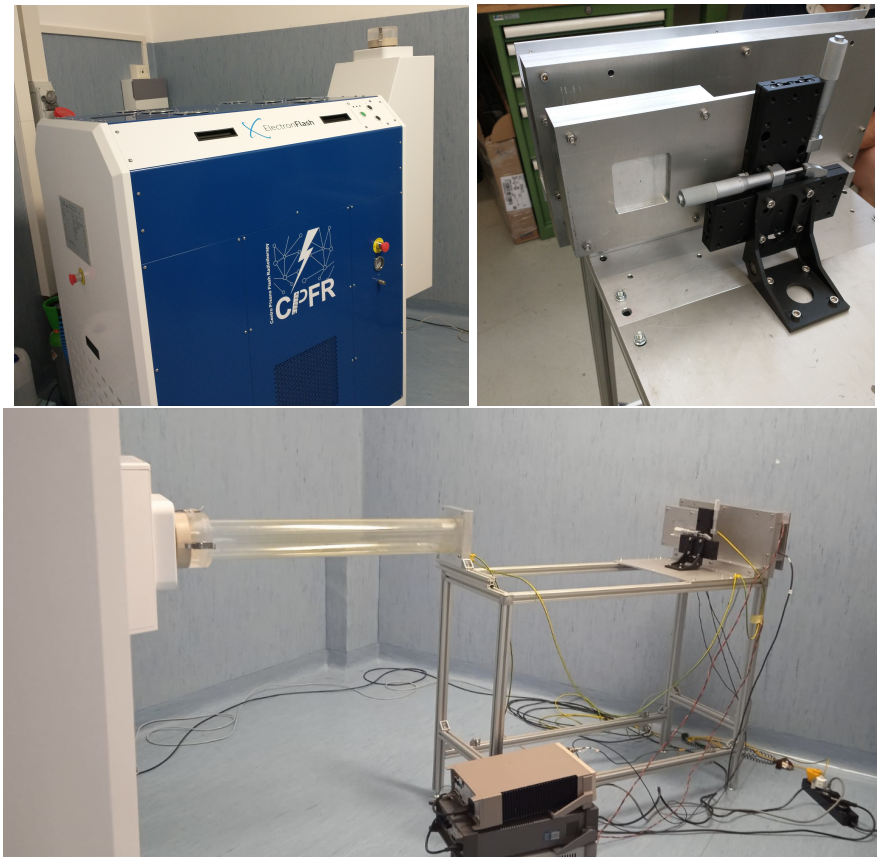


Figure 5.3: Experimental set up. Top left: ElectronFlash accelerator; a rotating gantry allows the gun orientation from 0° to 90° (horizontal /vertical). Top right: collimator B and DUT box. Bottom: whole structure mounted: we used the 10 cm diameter and 1.2 m long plexiglass tube; the DUT which is in the box behind the two collimators is connected to the power supply units.

operate on the reduction of the gain.

Recalling the results in section 4.1.4, I have shown that concerning the PMOS flavor B, decreasing the bias from -6 V to 0 V brings a reduction of efficiency down to 40 %, and in the gain of a factor ∼1/3.

In order to take advantage of the analog pile up and integrate the charge, two consecutive electrons must hit the pixel in a relatively small time. In fact, as already explained in section 3.1.3, the pixel completely paralyzes when its pulse goes under the threshold (TE); then the rate of arrival of electrons must be high enough to prevent that the second electron arrives before the TE. Since the typical ToT of a particle depends on the FE settings, this condition requires careful consideration.

During the testbeam many runs have been performed, spanning the energy, the dose per pulse and the four possible configurations with/without the collimators. We have collected data with the PMOS flavor B in the standard configuration: with the PWELL and PSUB biased at -6 V and we have used the default configuration of the FE parameters (the same used for the calibration and for the acquisition of spectrum in section 4.1.5). Meanwhile, we have selected pulses with t_p of 4 µs and with the smallest settable Pulse Repetition Fre-

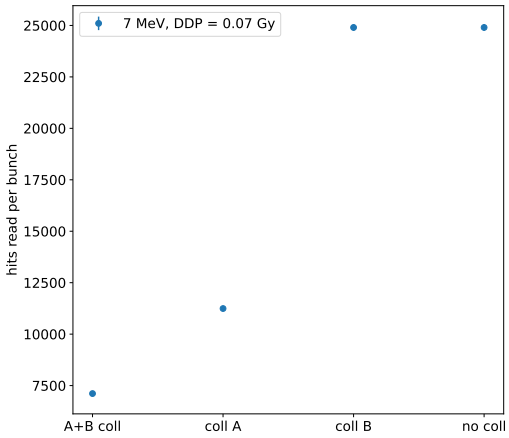


Figure 5.4: Mean number of hits read per bunch at DDP=0.07 Gy, with all the possible setup condition: with both the collimator, with only the collimator far from the chip (A), with only the collimator near the chip (B), and without any collimator. With the configuration B and without any collimator all pixels in the matrix fire.

quency, which is 1 Hz, in order to start in the most conservative working point, excluding the digital pile up of events from different bunches. Under these conditions, even if the whole matrix turns on, the total readout time corresponds to 25000×1 µs=25 ms and is still lower than the time between two consecutive pulses. In figure 5.4 is shown the mean number of hits read during one accelerator pulse in different setup conditions.

I will briefly discuss a few details of how the readout of the chip works (for a complete description see section 3.1.3), since it has a direct consequence on how the data were collected.

The readout starts with the trailing edge (TE) of the first pulse going below the threshold: about 50 clk=1.25 µs after this moment the FREEZE signal is sent to the whole matrix, and the transmittion of the data to the EoC begins. The hits read during the FREEZE signal are the ones whose TE occurred before the start of the FREEZE; instead, the ones whose TE occur during the FREEZE are stored in the pixel memory until the end of the first FREEZE signal. At this point, after ∼50 clk, a second readout starts and a second FREEZE is sent to the matrix. A time scheme of the sequence of the signal is shown in figure 5.5.

An example of the two sub-pulses corresponding to an electron bunch is shown in figure 5.6. In the acquisition we injected 5 pulses with both the collimators mounted on the table. Looking at the spectrum we can see that the second sub-pulse has a populated tail on the right; this is due to the fact that the hits which arrive before the start of the first FREEZE but have a long ToT that falls during the FREEZE, are read at the second sub-pulse.

No effect of the collimator can be seen (fig.5.6) and the distribution is uniform, indicating that the collimators do not shield particles as expected. It is possible that this

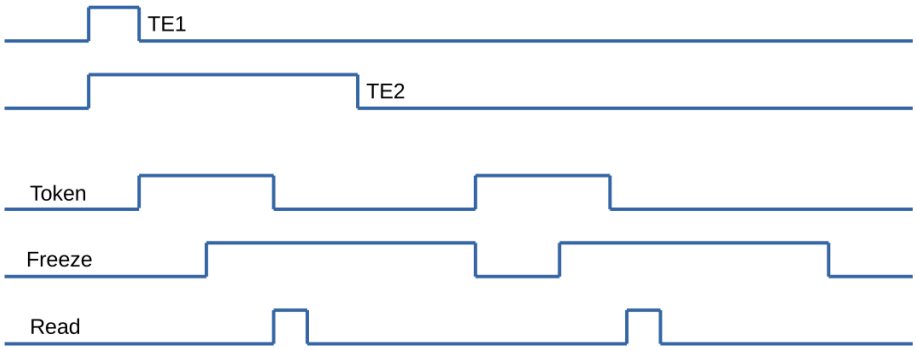


Figure 5.5: Readout sequence for two consecutive pulses which arrives at the same time but are read in two different FREEZE.

is due to a background of Bremsstrahlung photons higher than expected, but a full verification of that and the analysis of the data is still going on. In figure 5.7, instead, the histograms with a higher Dose Per Pulse value is shown; in the example the matrix turns on completely, but again this happens in two different consecutive read out steps.

Without the collimators, instead, the fluence greatly increased and the two-pulses substructure is no longer visible (fig.5.8). However, because of the high activity of the matrix, after each readout new hits with a fixed ToT were induced due to crosstalk. This problem had already been observed on other prototypes of TJ-Monopix1, and thanks to a simulation it has been observed that the main source of crosstalk is the voltage drop of the pre-amplifier ground as a result of the accumulated current that is drawn from the discriminator.

Unfortunately the available beam time was limited and we could not perform futher tests. Clearly TJ-Monopix1 is not well suited for dosimetry at high rates. Possible directions of improvements are: 1) significant reduction of the dead time with a fast readout allowing separeting the pulse substructure and 2) a biasing sheme that reduces the response od the sensor (the opposite of what is done for MIP detection) to reduce the saturation effect at high dose rate.

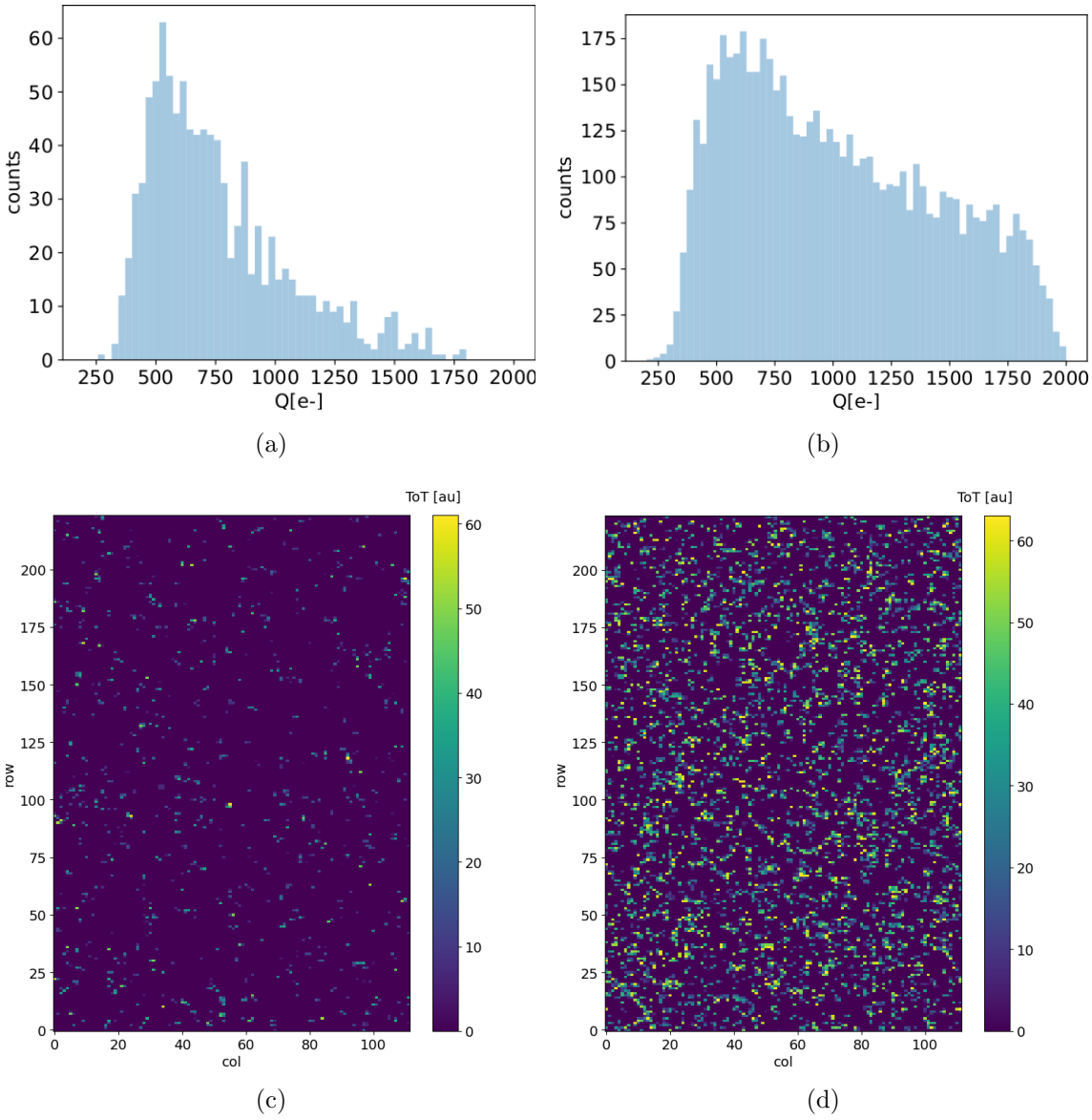
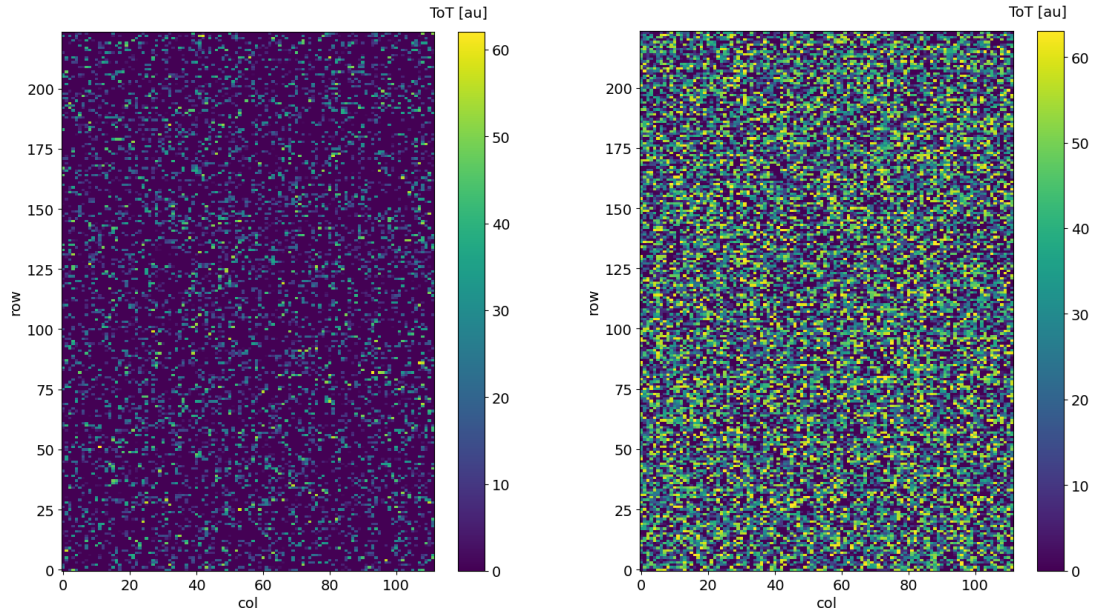
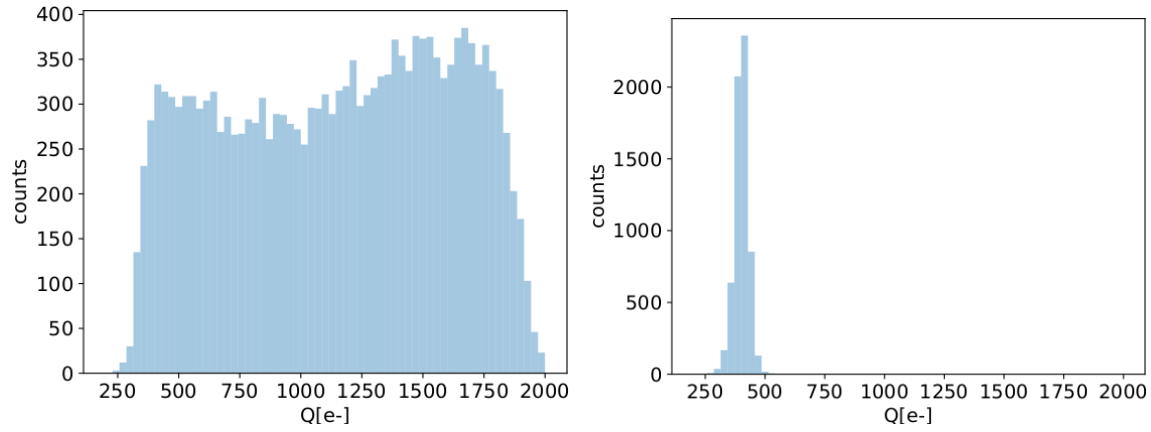


Figure 5.6: Acquisition with both the collimators: 5 pulses at DDP=0.07 Gy. In (a) and (b) the spectra of the charge released in the sensor and read during the first and the second FREEZE respectively. In (c) and (d) the 2D histogram of the ToT of the hits arrived in the sub-pulses.



(a) (b)

Figure 5.7: Acquisition with both the collimators: 5 pulses at DDP=0.6 Gy. In (a) and (b) the 2D histogram of the ToT of the hits read in the two FREEZE. Compared with the previous maps, more pixels turn on since the DDP is much higher.



(a) (b)

Figure 5.8: Acquisition without any collimator: 5 pulses at DDP=0.04 Gy.

CHAPTER 6

Conclusions

In this thesis I have presented the characterization of two Monolithic Active Pixel Devices, the TJ-Monopix1 and the ARCADIA main demonstrator. They are prototypes still in an initial development phase and the purpose of this characterization is to help understand the detailed operation of the devices, not to reach clear conclusions about their suitability for specific applications. For both devices I contributed to the setup of the test environment in the INFN clean laboratories and carried out the measurements personally.

Concerning TJ-Monopix1 the values found are reasonably in agreement with the simulations, although the threshold and noise ($\sim$400 e^- and $\sim$15 e^-) have been found to be higher than the expected values ($\sim$270 e^- and $\sim$9 e^-). This difference is not too surprising, and can be justified considering that the simulations were performed with the front end in a optimized status, while in our measurements the front end working point optimization was limited by the need to keep under control the number of noisy pixels.

The threshold dispersion was measured to be $\sim$30 e^-, in agreement with the simulation; the dispersion across the matrix can be reduced and can be make comparable with the ENC by adding a bit for trimming on each pixel. TJ-Monopix1 was tested with Fe55 and Sr90 sources, and with cosmic rays, allowing the absolute calibration of the injection circuit and a first characterization of the chip response to radiation. An initial test of the device response to a high rate FLASH beam was also performed, although the limitations of the chip prototype prevented reaching conclusions on the suitability of the device for this application.

Regarding the ARCADIA-MD1 prototype the very preliminary results have shown that its behavior is in agreement with what expected and, after a complete characterization of the front end, the test of its pioneering readout mode and of its coarse hardware clustering algorithm depending on the operating range will be certainly a main target.

The R&D of monolithic active devices is and important and active sector since they represent a low-cost and versatile technology, with possible future applications in many field and, as stated several times, they will possible open new scenarios particle detectors

both in accelerator experiments and in medical physics. The future perspective is now the development of bigger and faster devices for what concerns the HEP experiments; a prosecution of the work activities on the Monopix series, for example, is going to lead to the characterization of the following prototype, TJ-Monopix2, within the Belle2 collaboration.

For what concerns the application in radiotherapy, FLASH-RT does not only requires new dosimeters, but also beam monitor and imaging detector to used in the diagnostic. In a such wide landscape a more detailed study of the sensor itself, studying the fabrication parameters and focusing on the operating limits at high dose rate, would better clarify the applicability of Monolithic Active Pixel Devices in this field.

Bibliography

[1] Chris Damerell. *Tracking the rise of pixel detectors*. https://cerncourier.com/a/tracking-the-rise-of-pixel-detectors. July 2021. URL: https://cerncourier.com/a/tracking-the-rise-of-pixel-detectors.

[2] J. Baudot et al. "First test results Of MIMOSA-26, a fast CMOS sensor with integrated zero suppression and digitized output". In: *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*. 2009, pp. 1169–1173. DOI: 10.1109/NSSMIC.2009.5402399.

[3] A. Dorokhov et al. "High resistivity CMOS pixel sensors and their application to the STAR PXL detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 650.1 (2011). International Workshop on Semiconductor Pixel Detectors for Particles and Imaging 2010, pp. 174–177. DOI: https://doi.org/10.1016/j.nima.2010.12.112. URL: https://www.sciencedirect.com/science/article/pii/S0168900210028925.

[4] Giacomo Contin et al. "The STAR MAPS-based PiXeL detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 907 (2018). Advances in Instrumentation and Experimental Methods (Special Issue in Honour of Kai Siegbahn), pp. 60–80. DOI: https://doi.org/10.1016/j.nima.2018.03.003. URL: https://www.sciencedirect.com/science/article/pii/S0168900218303206.

[5] J.-C. Brient, R. Rusack, and F. Sefkow. "Silicon Calorimeters". In: *Annual Review of Nuclear and Particle Science* 68.1 (2018), pp. 271–290. DOI: 10.1146/annurev-nucl-101917-021053. eprint: https://doi.org/10.1146/annurev-nucl-101917-021053. URL: https://doi.org/10.1146/annurev-nucl-101917-021053.

[6] C. Marinas and M. Vos. "The Belle-II DEPFET pixel detector: A step forward in vertexing in the superKEKB flavour factory". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 650.1 (2011). International Workshop on Semiconductor Pixel Detectors

for Particles and Imaging 2010, pp. 59–63. DOI: https://doi.org/10.1016/j.nima.2010.12.116. URL: https://www.sciencedirect.com/science/article/pii/S0168900210028962.

[7] Pasi Kostamo et al. "GaAs Medipix2 hybrid pixel detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 591.1 (2008). Radiation Imaging Detectors 2007, pp. 174–177. DOI: https://doi.org/10.1016/j.nima.2008.03.050. URL: https://www.sciencedirect.com/science/article/pii/S0168900208004294.

[8] Nolan Esplen, Marc S Mendonca, and Magdalena Bazalova-Carter. "Physics and biology of ultrahigh dose-rate (FLASH) radiotherapy: a topical review". In: *Physics in Medicine & Biology* 65.23 (Dec. 2020), 23TR03. DOI: 10.1088/1361-6560/abaa28. URL: https://doi.org/10.1088/1361-6560/abaa28.

[9] Fabio Di Martino et al. "FLASH Radiotherapy With Electrons: Issues Related to the Production, Monitoring, and Dosimetric Characterization of the Beam". In: *Frontiers in Physics* 8 (2020). DOI: 10.3389/fphy.2020.570697. URL: https://www.frontiersin.org/articles/10.3389/fphy.2020.570697.

[10] W. Snoeys et al. "A process modification for CMOS monolithic active pixel sensors for enhanced depletion, timing performance and radiation tolerance". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 871 (2017), pp. 90–96. DOI: https://doi.org/10.1016/j.nima.2017.07.046. URL: https://www.sciencedirect.com/science/article/pii/S016890021730791X.

[11] H. Kolanoski and N.Wermes. *Particle Detectors: Fundamentals and Applications.* OXFORD University Press, 2020.

[12] E. Mandelli et al. "Digital column readout architecture for the ATLAS pixel 0.25 /spl mu/m front end IC". In: *IEEE Transactions on Nuclear Science* 49.4 (2002), pp. 1774–1777. DOI: 10.1109/TNS.2002.801528.

[13] Maurice Garcia-Sciveres and Norbert Wermes. "A review of advances in pixel detectors for experiments with high rate and radiation". In: *Reports on Progress in Physics* 81.6 (May 2018), p. 066101. DOI: 10.1088/1361-6633/aab064. URL: https://doi.org/10.1088/1361-6633/aab064.

[14] M. Dyndal et al. "Mini-MALTA: radiation hard pixel designs for small-electrode monolithic CMOS sensors for the High Luminosity LHC". In: *Journal of Instrumentation* 15.02 (Feb. 2020), P02005–P02005. DOI: 10.1088/1748-0221/15/02/p02005. URL: https://doi.org/10.1088/1748-0221/15/02/p02005.

[15] M. Barbero et al. "Radiation hard DMAPS pixel sensors in 150 nm CMOS technology for operation at LHC". In: *Journal of Instrumentation* 15.05 (May 2020), P05013–P05013. DOI: 10.1088/1748-0221/15/05/p05013. URL: https://doi.org/10.1088/1748-0221/15/05/p05013.

[16] K. Moustakas et al. "CMOS monolithic pixel sensors based on the column-drain architecture for the HL-LHC upgrade". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 936 (2019). Frontier Detectors for Frontier Physics: 14th Pisa Meeting on Advanced Detectors, pp. 604–607. DOI: https://doi.org/10.1016/j.nima.2018.09.100. URL: https://www.sciencedirect.com/science/article/pii/S0168900218312531.

[17] I. Caicedo et al. "The Monopix chips: depleted monolithic active pixel sensors with a column-drain read-out architecture for the ATLAS Inner Tracker upgrade". In: *Journal of Instrumentation* 14.06 (June 2019), pp. C06006–C06006. DOI: 10.1088/1748-0221/14/06/c06006. URL: https://doi.org/10.1088/1748-0221/14/06/c06006.

[18] D. Kim et al. "Front end optimization for the monolithic active pixel sensor of the ALICE Inner Tracking System upgrade". In: *Journal of Instrumentation* 11.02 (Feb. 2016), pp. C02042–C02042. DOI: 10.1088/1748-0221/11/02/c02042. URL: https://doi.org/10.1088/1748-0221/11/02/c02042.

[19] L. Pancheri et al. "A 110 nm CMOS process for fully-depleted pixel sensors". In: *Journal of Instrumentation* 14.06 (June 2019), pp. C06016–C06016. DOI: 10.1088/1748-0221/14/06/c06016. URL: https://doi.org/10.1088/1748-0221/14/06/c06016.

[20] Lucio Pancheri et al. "Fully Depleted MAPS in 110-nm CMOS Process With 100–300-μm Active Substrate". In: *IEEE Transactions on Electron Devices* 67.6 (2020), pp. 2393–2399. DOI: 10.1109/TED.2020.2985639.

Acknowledgements

Vorrei ringraziare chi negli anni e nei mesi passati mi ha sostenuto, aiutandomi a portare a termine questo percorso impegnativo, stimolante e anche tanto divertente.

Per gli insegnamenti, la tolleranza e la fiducia (forse anche troppa) ringrazio moltissimo Francesco Forti. Grazie per avermi aiutato in una corsa contro il tempo garantendomi disponibilità e sostegno, tanto dal punto di vista lavorativo (grazie di aver rimandato gli impegni del fine settimana per correggere la tesi) quanto per il sostegno dal punto di vista emotivo (grazie di avermi trasmesso quell'incoraggiamento e quella determinazione per me necessari). Cercherò di ricordarmi che la determinazione è un ingrediente fondamentale nella vita, sia, come in questo caso, nella fase finale di realizzazione di un lavoro, sia nella fase iniziale in cui essere determinati a rispondere alle proprie domande è imprescindibile (prometto a lei e a me stessa che la prossima volta non aspetterò l'ultimo momento per chiedere chiarimenti). Un altro insegnamento che ho appreso nel confrontarmi con lei in questi mesi è come la determinazione debba essere sempre accompagnata dall'umiltà: determinazione e umiltà, non semplice, ma proverò a costruire il mio lavoro, secondo il suo esempio, su questi principi.

Il secondo ringraziamento va alla mia attuale e futura relatrice, la professoressa Maria Giuseppina Bisogni, con cui spero di poter applicare quanto imparato finora e che spero di non deludere dandole ripensamenti sulla scelta di lavorare assieme.

Un grande ringraziamento va a Giuliana, per avermi guidato e aiutato a risolvere moltissimi dei miei dubbi. Per me sei stata un punto di riferimento costante in questi mesi; mi auguro in futuro di poter lavorare ancora insieme in camera pulita e di poter condividere con te l'entusiasmo di una nuova misura o di un nuovo test su qualcosa che si è rivelato non funzionare perfettamente (cosa c'è di più entusiasmante di un rivelatore di cui bisogna testare e indagare il comportamento perchè non funziona esattamente come ci si aspetta?!).

Voglio poi ringraziare chi praticamente ha reso possibile il mio lavoro aiutandomi all'occorrenza a risolvere i problemi tecnici e sperimentali che si sono presentati: mi riferisco in particolare a Stefano, Ludovico, gli ingegneri Massimo, Maurizio e Andrea.

Non posso non menzionare Giulia, Francesco, Suryanarayan, e Luigi, persone dedite al lavoro, che saranno sicuramente fonte d'ispirazione per me.

Un enorme ringraziamento va in generale a tutto il gruppo di Belle II di Pisa, un gruppo in cui sono cresciuta, sono stata bene e in cui ho consolidato la mia passione per questo lavoro: voi tutti siete per me un esempio di come io spero di diventare da grande.

Un grazie va poi alle mie numerose famiglie: a quelle dei miei genitori e i miei amici più cari, per avermi supportato economicamente e moralmente. Ognuno di voi ha ovviamente contribuito in modo diverso alla realizzazione di questo percorso, ma tutti mi avete appoggiato dandomi la fiducia e la sicurezza di poter fare e di poter diventare ciò che volevo.

Un grazie va a chi ha condiviso questa esperienza con me dallo stesso mio punto di vista rendendola piacevole e divertente: ai laureandi, laureati, compagni di ufficio e amici dell'auletta, con cui in questi mesi ci siamo posti gli stessi dubbi, ci siamo confrontati e abbiamo fatto squadra.

Infine un grazie particolare ad Alberto, che mi ha sostenuto in *tutto*, senza eccezioni. Grazie per avermi trasmesso e insegnato quante più cose potessi (e qui per la mia testardaggine probabilmente si potrebbero trovare numerose eccezioni), per avermi incoraggiato prima degli esami, e poi per aver festeggiato con me o per avermi consolato a seconda dell'esito. Credo che il nostro lavoro che appassiona tanto entrambi non mi avrebbe dato tanta soddisfazione se non avessi avuto la possibilità di condividerlo a tavola nei momenti di vita quotidiana con te.