Università degli Studi di Bari
XXXV Ciclo del Dottorato di Ricerca in Fisica

Silicon pixel sensor characterization for ultra-light tracking detector with truly-cylindrical geometry.

Relatori:
Prof. Giuseppe Bruno
Dott. Domenico Colella

Dottoranda:
Arianna Grisel Torres Ramos

Anno Accademico 2022/2023

Abstract

ALICE (A Large Ion Collider Experiment) is the heavy-ion experiment at the Large Hadron Collider (LHC) at CERN; it is designed to address the physics of strongly interacting matter and in particular the properties of quark-gluon plasma. The experiment took data from 2009 to 2018 (LHC Run1 and Run 2) and went under a major upgrade during the LHC Long Shutdown 2 (2018-2021). Starting from 2022 detector is back in operation and it is foreseen to keep performing measurements up to 2032. A pillar of the ALICE future upgrade program is the improvement of the Inner Tracking System (ITS2) performance by the replacement of its three innermost layers (ITS3) during the LHC Long Shutdown 3, taking place between 2026 and 2028. The proposal is based on a vertex detector consisting of three cylindrical concentric layers composed by curved wafer-scale silicon sensors, built on CMOS Monolithic Active Pixel Sensors (MAPS) with the sensor matrix and readout integrated in a single chip. The first detector layer will be positioned at a radial distance from the interaction point of 18 mm. The new detector will present a significant reduction of the material budget, from 0.36% X_0 (in ITS2) to 0.05% X_0 per layer, thus improving the vertexing and tracking performance, especially for particles with low transverse momentum ($p_T < 1$ GeV/c).

The main scope of this thesis is to report on the activities associated with the characterization of bent MAPS, fabricated with TowerJazz 180 nm CMOS Technology, as well as the development of tools and techniques dedicated to bend accurately such devices.

Extensive characterization of a single ALPIDE chip (used for the current ITS2), positioned flat and bent down to the intended radius for the innermost layer of the ITS3 (18 mm), was carried out in order to evaluate its performance under the mechanical stress involved in the bending process. The measurement of the main variables steering the chip functioning was performed by evaluating their behavior over the time and through the variation of different in-pixel circuit parameters. The results were compared for both geometries. These studies aimed to demonstrate that ALPIDE functionalities are preserved despite the curvature effects.

An important component in the investigation of a detector performance is the realization of a so-called test-beam experiment. The data acquired during the first ever in-beam characterization of a bent MAPS, performed in a test-beam campaign at the DESY II test-beam facility in June 2020, were analyzed

and systematically compared with data taken with a flat ALPIDE. The study had the main scope to evaluate the sensing performance of a bent chip through the variation of the working point defined by front-end parameters and contrast it with the results of an originally flat sensor.

The analysis of efficiency, cluster size and preliminary position resolution was also performed for an experiment realized in July 2021 at CERN SPS facility, where a set of six ALPIDE chips were bent at the foreseen ITS3 layer radii, forming the so-called micro-ITS3. The objective was the simultaneous in-beam characterization of several concentrically arranged bent ALPIDE chips, thus modeling the tracking layers of the future ITS3.

An additional study was done from measurements where a copper target was introduced in the center of the detector barrel of the micro-ITS3 in order to mimic particles emerging from a real collision. Through this activity the possibility to reconstruct tracks and vertices from hadronic interactions in a geometrical configuration very close to the next full ITS (including ITS3) has been explored and demonstrated.

The tests on small sensors have opened the way to the investigation of a large scale sensor. The "super-ALPIDE" project has the goal to assembly a detector with geometry similar to one half-layer of the ITS3, using a working sensor with large dimensions made of a matrix of not diced ALPIDE chips. The concept of the super-ALPIDE design, including the components and mechanical structures implemented for readout and support, is presented in this thesis. Techniques and tools employed to perform sensor bending, parts assembly and wired interconnections are described in detail.

Another subject, within the context of the ITS upgrade R&D, is the preliminary study of new sensor prototypes based on TPSCo 65 nm CMOS process, the first choice of the technology for the implementation of the wafer-scale sensor, which allows for more than four-fold increase of the number of transistors per pixel compared to the older 180 nm and higher spatial resolution resulting from the possibility to fabricate smaller pixels. A first submission in TPSCo 65 nm was done as a Multi Layer Reticle (MLR1). An overview of the three main prototype structures composing MLR1 (Digital Pixel Test Structure, CE65 and Analogue Pixel Test Structure) is given in the last chapter of this thesis. The APTS op-amp, modified with gap version with DC coupled in-pixel architecture and 10 μm pitch, was tested in laboratory in order to characterize this chip flavor. The response of the pixels front-end was monitored by the measurement at different chip bias settings to evaluate the influence of the main parameters of the in-pixel circuit on the signal processing. The study of the device charge collection properties was performed through measurements with X-rays emitted by an ^{55}Fe source.

Contents

Chapter 1

Introduction

ALICE (A large Ion Collider Experiment) at the Large Hadron Collider (LHC) at CERN [1], has been designed and optimized to address the physics of strongly interacting matter, in particular the properties of the quark-gluon plasma. It performs particle tracking and identification through several techniques and multiple detector technologies, allowing a comprehensive study of hadrons, electrons, muons and photons, produced in nucleus-nucleus collisions at high energies.

The innermost detector in the central barrel is the Inner Tracking System (ITS), which is placed around a beryllium beam pipe close to the Interaction Point (IP). The ITS has been replaced with an upgraded detector (ITS2) [2] during the second LHC Long Shutdown (LS2) in early 2021. The ITS2 consists of seven approximately-cylindrical detector layers based on CMOS Monolithic Active Pixel Sensors (MAPS) with the sensor matrix and readout integrated in a single chip, named ALPIDE (ALice PIxel DEtector), covering a 10 m^2 area with about 12.5 billion pixels. The main objective was to increase the readout capabilities to allow the readout and recording of Pb–Pb minimum bias events at rates in excess of 50 kHz, the expected Pb–Pb interaction rate at the LHC after LS2.

The core of the detector is the ALPIDE chip, which is the result of an intensive R&D effort carried out in the last decade and has led to a quantum leap in the field of MAPS for single-particle detection, reaching unprecedented performance in terms of signal/noise ratio, spatial resolution, material budget and readout speed. There is still a lot that can be done to further improve MAPS for high-energy physics detectors by fully exploiting the rapid progress that this technology is making in the field of imaging for consumer applications. One of the features offered recently by CMOS imaging sensor technologies, called stitching, will allow developing a new generation of large size MAPS with an area of up to 26 $cm \times 26$ cm using wafers that are 300 mm in diameter. Moreover, the reduction of the sensor thickness to values of about 20–40 μm will open the possibility of exploiting the flexible nature of silicon to implement large-area curved sensors. In this way, it will become possible to build a cylindrical

detector layer made only by silicon sensor, significantly reducing the material of the supporting structures.

In that direction, an upgrade of the ITS2 is planned, aiming for a new detector installation during the third LHC Long Shutdown (LS3) between 2026 and 2028. The proposed upgrade, which is called ITS3, is based on a new vertex detector consisting of three truly-cylindrical layers composed by curved wafer-scale stitched silicon sensors, featuring a material budget of 0.05 % X_0 per layer, with the first layer positioned at a radial distance from the interaction point of 18 mm. In the ongoing development of ITS3, the CMOS technology node was changed from 180 nm to 65 nm, which gives further advantages in integration density and power consumption. The new detector is intended to replace only the three innermost layers of the ITS2. With the first detection layer closer to the interaction point (from 23 mm to 18 mm) and a reduction of the material budget close to the interaction point by a factor of six, the new vertex detector will significantly improve the spatial precision and tracking efficiency at low p_T. The improvement of the vertexing performance and the reduction of material budget will have a dramatic impact on the measurement of charm and beauty hadrons at low transverse momentum as well as on the measurement of low-mass and low p_T dielectrons [3].

The content of this thesis is dedicated to: 1) the investigation of the ALPIDE features, specifically to verify that its performance does not change when it is positioned in flat and bent configuration; 2) the development of techniques dedicated to bend large size sensors and 3) the study on the feasibility of the newest MAPS prototypes fabricated in 65 nm CMOS process, to fulfill the mechanical and detection performance requirements for the future ITS3.

In this chapter an overview of the physics addressed by ALICE and a general description of the detector complex is given. The concept and improvement strategy for the next ALICE Inner Tracking System upgrade will be explained. In chapter 2, the properties of silicon detectors and the principles of operation of MAPS will be discussed. The ALPIDE chip used currently in the ALICE ITS2 is presented in chapter 3. The procedures developed to bend these devices down to the foreseen ITS3 layer radii will be reported and possible consequences of such bending will be investigated through the results of electrical tests in laboratory and studies on its performance in different test-beam experiments. In chapter 4, the "super-ALPIDE" project will be introduced by a description of the super-ALPIDE design, including the mechanical structures for readout and support. The techniques and tools employed to perform sensor bending, components assembly and wired interconnections will be depicted as well. Chapter 5 will be dedicated to give an overview of new sensor prototypes based on TPSCo 65 nm CMOS process, the first choice of the technology for the implementation of the wafer-scale sensor in the ITS3, focusing the attention on the results of an Analogue Pixel Test Structure characterization. Finally, the conclusions of the work presented in this thesis will be summarized in chapter 6.

1.1 Ultra-relativistic heavy-ion collisions

The collisions of ultra-relativistic heavy-ions allow us to study the strongly interacting matter under extreme conditions of high temperature and/or high nuclear matter density. The matter that fills the Universe today is done of hadrons, e.g. protons and neutrons that could further aggregate into nuclei, and the hadrons are made from quarks, the basic building blocks of matter, and held together by gluons, the carriers of the strong (nuclear) force. Under normal conditions, i.e. the ones of our Solar System in the Universe, the quarks are bound ('confined') into composite objects called baryons (bound states of three quarks, with the proton and the neutron being the most prominent examples) and mesons (bound states of a quark and anti-quark pair). Lattice computations of Quantum Chromodynamics (QCD), the theory of strong interactions, predict that at a sufficiently high energy density, there will be a transition from ordinary nuclear or hadronic matter to a plasma of free ('deconfined') quarks and gluons— a transition that took place in the early Universe a few microseconds (~ 10 μs) after the Big Bang and which might still play a role today in compact stellar objects. The characterization of this new high-temperature phase of matter, called the 'quark–gluon plasma' (QGP), requires a sufficiently large volume of hot matter and is therefore pursued in collisions of heavy nuclei at the highest possible energy. The coupling constant α_S of the strong interaction depends on the momentum scale of the interaction. Processes corresponding to large momentum transfers between partons [4], where α_S is small, can be investigated via a perturbative approach (pQCD). This involves an expansion to higher orders in α_S [5], with only lower orders contributing significantly. For calculations concerning low-momentum transfer processes, the perturbative approach of QCD breaks down, as higher orders become dominant [6]. Such a non-perturbative regime, where the coupling constant becomes large, is of the utmost interest for the description of several essential features of the strong interaction. In particular, in this regime confinement[1] is observed [7]. This leads to the fact that quarks and gluons, the elementary particles carrying the "charge" of the strong interaction, known as colour, cannot be isolated, and therefore not directly observed. Only composite objects, hadrons, which do not possess a net colour charge, can be detected. Another key feature of strong interaction in the non-perturbative regime is the spontaneous breaking of chiral symmetry[2] [8, 9, 10]

The study of the QCD matter in its phase where quark and gluons are deconfined, i.e. not bound into hadrons, can be performed by producing a many body system of quarks and gluons, under the conditions of large energy density. Heating such a system, with zero (or very small) net baryon density, to a temperature exceeding 150–160 MeV (equivalent to more than 10^{12} K), leads

[1] Quark confinement is the technical term describing the observation that quarks do not occur isolated in nature, but only in hadronic bound states as mesons and baryons.

[2] Chiral symmetry breaking expresses the fact that quarks confined in hadrons do not appear as nearly massless constituents but are endowed with a dynamically generated mass of several hundred MeV.

to the creation of the quark–gluon plasma (QGP). This is a state of matter where two of the basic features of low-temperature QCD, confinement and chiral symmetry breaking, are no longer present [11, 12, 13]. Below the ideal gas limit, the QGP manifests itself as a strongly interacting system of quarks and gluons. It behaves as an almost perfect liquid [14], with the possible presence during its evolution of bound states of quasi-particles (qq, gg, qg) [15].

1.1.1 The evolution of a heavy-ion collision

Collisions of heavy ions with ultra-relativistic energies are used to create the QGP in the laboratory. The evolution of a heavy-ion collision is commonly described in terms of a series of stages, which can in principle be factorised. They include: (i) an initial state, defined by the wave-functions of the projectiles, which are universal and independent of any specific scattering process; (ii) large-Q^2 interactions of partons drawn from the projectiles[3]; (iii) smaller-Q^2 interactions generating a pre-equilibrated parton gas; (iv) equilibration and expansion of the QGP; (v) hadron formation; (vi) chemical freeze-out of hadrons; (vii) hadronic interactions that subsequently freeze-out kinetically; (viii) free-streaming of stable particles to the detector.[4]

During a heavy ion collision the very large number of small-Q^2 interactions between the nucleons of the two colliding nuclei determine the overall energy density and entropy deposition in the initial state. Immediately after the collision, other smaller-Q^2 interactions occur in the context of a weakly coupled pre-equilibrium phase. This is followed by the creation of even softer partons in these processes, which enable the formation of a strongly coupled QGP phase. The hard processes from large-Q^2 interactions, which take place in the very early stage of the collisions, enable the creation of high momentum gluons and high momentum/mass quarks. As they have short wavelengths, they will interact with other quarks and gluons on a microscopic level, leading to energy loss effects (the energy being transferred to the medium), and therefore they offer information on the opaqueness of the QGP. The interactions of high-momentum partons with the QGP can be radiative as well as elastic. The amount of energy loss will depend on the colour charge, momentum, mass, type of process (inelastic or elastic), the distance traversed (path length) of the hard scattered parton, and is subject to stochastic processes. The heavy quarks produced via hard processes can also form quarkonia (bound heavy quark-antiquark states), with their production rate being suppressed with respect to scaled expectation from ordinary pp collisions, because the binding force between the quark and anti-quark is weakened (screened) by the presence of the colour charge of quarks and gluons. In addition, the parton fragmentation processes lead to jets, partonic showers that arise from these high energy partons, and that fragment into experimentally observable hadrons.

The evolution of the QGP for most processes involved in soft interactions after $\sim 1\ fm/c$ can be depicted as follows. Since the mean free path of the

[3] Q^2 being the 4 momentum transfer squared

[4] For a more detailed conceptual description of heavy-ion collisions, see Ref.[16]

vast majority of the QGP constituents is expected to be much smaller than the size of the formed QGP (assuming these constituents are strongly coupled), multiple interactions drive the expansion. This expansion is highly influenced by the non-uniform energy distribution in the initial state as a function of space, which creates pressure gradients in the QGP stage. A radial flow occurs due to a greater pressure at the centre of the QGP compared to the outskirts, and this leads to a common velocity field outwards, which is superimposed to a thermal motion. The rate of the hydrodynamic expansion is influenced by the QGP's bulk viscosity, which is its resistance to volume growth. Anisotropic flow is the result of a directional dependence to these pressure gradients. This occurs due to spatial anisotropies in the initial state. These arise, e.g. if the collision zone is almond shaped ($b^5 > 0$, non central collisions). Such spatial anisotropies are converted to momentum anisotropies via the hydrodynamic response. This is influenced by the QGP's shear viscosity, which quantifies the resistance to fluid deformation.

Since the nuclei are charged, the movement of the incoming beams sets up a very large magnetic field (from the protons in the nuclei with relativistic energies), which can also have an influence on motion of quarks during the QGP stage. The Chiral Magnetic Effect (CME) leads to a splitting among positive and negative quarks in the QGP along the direction of the extremely large magnetic fields produced from the colliding ions. The decaying magnetic field will also induce an electric field, which can move the electrically charged quarks accordingly. The high-temperature QGP also produces thermal radiation, in the form of both photons and lepton-antilepton pairs. The electromagnetic radiation does not interact via the strong force in the QGP, and can therefore be used to gain information about the temperature from early to late times in the collision processes. In addition, strange quarks, which have masses below the deconfinement temperature (therefore the QGP temperature) can also be produced easily in the QGP stage, with the dominant mechanism coming from gluon interactions.

During the evolution, the parts of the QGP that cool below the transition temperature T_{pc} ("pseudocritical temperature") will hadronise. The temperature and energy density of the medium will diminish with increasing distance from the collision centre and the transition to normal nuclear matter is a smooth cross-over. It is possible that hadronisation will happen at different times at different places in the phase-space.

The formation of hadrons from the QGP can proceed as follows. Hard partons in a jet will fragment and hadronise in the same manner as in elementary collisions. For partons at lower momenta, which are the majority, if they share a similar space and momenta as other partons, they can combine into hadrons via coalescence. Heavy (charm or beauty) quarks can also combine with heavy antiquarks, thus forming quarkonia and give rise to an additional production mechanism for closed heavy-flavour hadrons.

[5]Impact parameter: distance between the centers of two colliding nuclei at their closest approach during a collision.

After the hadrons materialise out of quarks and gluons, primarily pions (π), kaons (K) or protons (p), the energy density may be large enough to allow for inelastic interactions, with a consequent evolution of their "chemical" composition, in terms of particle species. A loosely bound state such as the deuteron (d) is particularly sensitive to such interactions, as it can be easily formed or destroyed. These interactions cease at the chemical freeze-out temperature T_{chem}, fixing the particle composition. Elastic interactions can still continue and halt at the kinetic freeze-out temperature T_{kin}, which is achieved at the time of $\sim 15\ fm/c$. At this point, the particle momenta are fixed. These particles travel towards the ALICE detector, where they will be measured $\sim 10^{15} fm/c$ after the initial collision.

1.1.2 Observables in heavy-ion collisions

There are various experimental probes used to investigate all phases of heavy-ion collisions: the initial state, the QGP phase, and the final hadronic phase. Each of these probes has varying sensitivity to each phase. To start, a fundamental quantity for many of these probes is the Lorentz-invariant differential yield of final state particles, given by:

$$E\frac{\mathrm{d}^3N}{\mathrm{d}p^3} = \frac{1}{2\pi p_T}\frac{\mathrm{d}^2N}{\mathrm{d}p_T\mathrm{d}y} \tag{1.1}$$

which is the number density of the particle three-momentum scaled by the particle energy (E). This will depend on the measured particle species in question, the transverse momentum, p_T, and rapidity, y. When the particle species is not known, pseudorapidity η is used instead[6], and both are equivalent when the particle energy is much greater than its mass. Different p_T ranges will probe different physical processes, and for clarity in this review, one can define low-p_T as $p_T \lesssim 2\ GeV/c$, intermediate-p_T as $2 \lesssim p_T \lesssim 8\ GeV/c$, and high-$p_T$ as $p_T \gtrsim 8\ GeV/c$.

Initial state: For a given collision, the multiplicity can be determined, which is the number of charged hadrons produced in a broad momentum range. It plays a critical role in providing a selection on a range of impact parameters b. It will be large when b is small, which leads to large numbers of N_{part}[7] that will correspondingly produce large numbers of particles. Such small-b collisions are referred to as central i.e. head-on, whereas collisions with large impact parameters and small numbers of N_{part} (and fewer produced particles) are referred to as peripheral. The multiplicity therefore provides an experimental handle on the centrality of a collision. Such a handle is extremely useful, as many of the system properties such as the energy density or lifetime depend upon the centrality. The multiplicity can also provide a measure of the initial state entropy.

[6] Defined as $\eta = -\ln(\tan(\theta/2))$, θ being the polar emission angle of the particle.

[7] Number of nucleons in the colliding nuclei that participate in an inelastic interaction (at least once)

Ultra peripheral collisions (UPCs) provide another handle on the initial state. These are heavy-ion collisions that typically have very large b values (greater than the nuclear diameter). The nuclei in these collisions are sufficiently separated such that short-range strong interactions are highly suppressed (and therefore no QGP formation). Measurements of the differential cross sections of light (e.g. ρ_0) and heavy (e.g. J/Ψ) vector mesons from these collisions are especially sensitive to the momentum distribution of partons (i.e. quarks and gluons) inside the nucleons. These distributions play a critical role for understanding the production rate of hard processes in the initial stages of heavy-ion collisions.

QGP phase: the amount of the hydrodynamic collective motion developed in the QGP phase can be explored using measurements sensitive to radial and anisotropic flow. Radial flow boosts higher mass produced hadrons to higher p_T in the low-p_T region, and measurements of identified particle mean transverse momentum $\langle p_T \rangle$ can therefore be used to extract information on the amount of radial flow that has developed. In order to investigate anisotropic flow, Eq. 1.1 can be expanded into the azimuthal angular dependence of the transverse momentum vector direction. The most widely used decomposition expresses the φ (the angular component of this vector) dependence of the produced particle density as a Fourier series as follows [17]:

$$\frac{\mathrm{d}N}{\mathrm{d}\varphi} \propto 1 + 2\sum_{i=1}^{\infty} \nu_n \cos[n(\varphi - \Psi_n)] \tag{1.2}$$

The ν_n terms, which can be determined experimentally, are referred to as anisotropic flow coefficients. They depend on the particle species, p_T and y. The term n corresponds to the order of anisotropic flow, and Ψ_n is the corresponding symmetry plane angle, which is the angular direction of anisotropic flow for the order n of interest. For the vast majority of produced hadrons, non-zero ν_n coefficients arise mainly from the QGPs hydrodynamic response (with a relatively small contribution from the hadronic state).

Hard probes are produced in the earliest times of the collision. One of the ways they can be explored is by extracting their nuclear modification factor R_{AA}. This observable is constructed to be sensitive to changes of the dynamics of hard processes in heavy-ion collisions with respect to expectations from elementary pp collisions. Within a rapidity (y) or pseudorapidity (η) interval, R_{AA} as a function of transverse momentum is defined as follows:

$$R_{AA}(p_T) = \frac{1}{\langle T_{AA} \rangle} \frac{\mathrm{d}N_{AA}(p_T)/\mathrm{d}p_T}{\mathrm{d}\sigma_{pp}(p_T)/\mathrm{d}p_T} \tag{1.3}$$

The average nuclear overlap function $\langle T_{AA} \rangle$ is obtained from the average number of N_{coll}[8] divided by the inelastic nucleon–nucleon cross section for the centrality range of interest. For hard processes, the yield N_{AA} in heavy-ion collisions is

[8]Total number of inelastic nucleon–nucleon collisions

expected to scale with the average nuclear overlap function $\langle T_{AA} \rangle$ when compared to the production cross section σ_{pp} in pp collisions, in the absence of any QGP or initial state nuclear effects. That being the case, if $R_{AA}(p_T) = 1$, production from heavy-ion collisions can be considered as a superposition of nucleon–nucleon collisions, assuming the QGP is not formed in pp collisions. Any departure from unity reveals how these processes are modified in heavy-ion collisions. In particular, R_{AA} is expected to be below unity at high-p_T for hadrons that originated from partons undergoing energy loss. Jet-finding algorithms can also be applied in heavy-ion collisions. The internal structure of jet showers is governed by quantum interference effects, resulting in the phenomenon of "angle-ordering", whereby the highest p_T hadrons in the shower are on average most closely aligned with the nominal jet axis. A key parameter in jet measurements is the "jet radius" or "resolution parameter" R, which is effectively the size of the aperture through which the jet shower is viewed. The measured jet yields can be used to determine R_{AA}, while the jet radius dependence of the R_{AA} or jet substructure measurements provide information about the medium modifications of the quark and gluon radiation patterns. The heavy-ion jets can be compared to jets in pp collisions, where no such modifications are expected.

A specific class of hard probes, heavy quarks, gives access to several QGP features that can be accessed via differential measurements of R_{AA} and ν_n. A key difference is that the corresponding hadrons are associated with early-stage probes across their entire p_T range. D mesons, which carry the vast majority of the charm quarks produced, can be investigated differentially as function of p_T via measurements of ν_n and R_{AA}. Measurements of ν_n at low-p_T are sensitive to degree to which heavy quarks take part in the collective expansion of the QGP and approach thermalisation. Measurements of R_{AA} at high-p_T provide insights on the energy loss processes of heavy quarks, which may be smaller than for light quarks or gluons, due to different Casimir coupling coefficient (with respect to gluons) and the dead cone effect (with respect to light partons) [18]. Measurements of quarkonia (e.g. J/Ψ or $\Upsilon(1S)$) R_{AA} and ν_n are also carried out as a function of p_T. The R_{AA} studies allow the investigation of the aforementioned suppression and regeneration processes, with the latter expected to be dominant in the low-p_T region due to larger heavy quark multiplicity in that kinematic range. Anisotropic flow measurements give complementary information to those from open heavy flavours, with the observed effect not involving the contribution of the flow of light quarks.

Electroweak probes e.g. photons, leptons or Z/W bosons provide another class of measurements. If they arise from hard processes in the initial stages, measurements of R_{AA} at high-p_T are expected to be unity, and any deviations from this reveal the influences of non-QGP processes that affect this measurement (these can be studied using other means that will be described in the next section). On the other hand, measurements of direct photon (i.e. not from hadron decays) spectra at low-p_T are expected to arise mainly from the softer processes in the QGP involving charged quarks, and their slope can be related to the temperature when they are created. Direct photons can also be produced from hadronic interactions, so that the observed yield is a convolution of their

emission along the whole collision history. Since photons are produced in an expanding QGP medium, they will be blue-shifted, which has to be taken into account regarding their connection to the temperature of the QGP.

Hadronic phase: a variety of methods are available to investigate the hadronic phase of the heavy-ion collisions. Measurements of identified hadron spectra and ν_n at intermediate-p_T are sensitive to the coalescence of quarks upon hadronisation from the QGP phase, which will compete with hadronisation from the well known fragmentation processes observed in QCD interactions. Such a coalescence mechanism enhances baryon production rates and ν_n values compared to mesons in this p_T range. The chemical freeze-out temperature T_{chem} can be explored by measuring the total yields of specific particles ($\mathrm{d}N/\mathrm{d}y$) via an integration over p_T in Eq. 1.1 for hadrons that do not decay quickly via the strong or electromagnetic interactions (e.g protons, kaons, charged pions, hyperons etc). Measurements of the production of strongly decaying resonances can offer insight into the duration of the hadronic phase. Their decay products for temperatures both above and below T_{chem} are subject to elastic scattering. If these effects are strong enough, they will lead to a depletion in the resonance yields, as the kinematics of the particles resulting from a resonance decay are altered in such a way that the resonance cannot be reconstructed anymore. On the other hand, regeneration processes from hadron–hadron interactions may form resonances during this phase. Final state hadron–hadron interactions can also be studied via the pair-wise femtoscopic correlation function, which can be extracted experimentally as:

$$C(k^*) = \frac{N_{pairs}(\text{same event})}{N_{pairs}(\text{background})} \tag{1.4}$$

where k^* is the invariant momentum difference between the pairs. The numerator is extracted from pairs produced in the same collision, and the denominator is obtained from a background hypothesis that assumes no correlated pairs, and is normalised such that $C(k^*) = 1$ in the absence of same-event correlations. Femtoscopic correlations arise from quantum correlations which are sensitive to the size of the system at freeze-out, or final state hadron–hadron interactions, both of which predominately occur at low k^*.

1.2 A Large Ion Collider Experiment

The ALICE detector stands 16 m tall, 16 m wide and 26 m long, and weighs approximately 10 000 tonnes. It was designed and upgraded over almost three decades by a collaboration that currently includes over 2005 scientists and engineers from more than 171 institutes in 40 different countries. The experiment consists of 12 different detection systems, each with its own specific technology choice and design constraints. A schematic view of ALICE after the Long Shutdown 2 (2019 – 2021), is shown in figure 1.1.

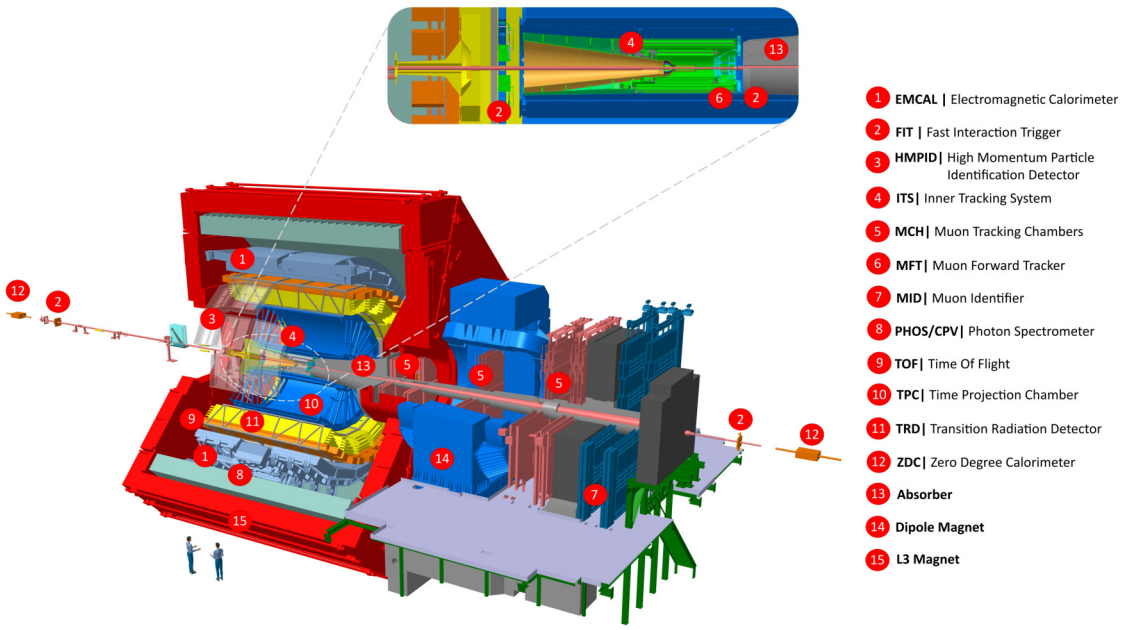


Figure 1.1: ALICE 2 detector systems [19].

It consists of a central part, which measures hadrons, electrons and photons, and a forward single arm spectrometer that focuses on muon detection. The central 'barrel' part covers a region spanning from 45° to 135° (angle from the beam direction) and is located inside a huge solenoid magnet, which maximum field at the nominal power of 4 MW reaches 0.5 T. The central barrel contains a set of tracking detectors, which record the momentum of the charged particles by measuring their curved path inside the magnetic field. These particles are then identified according to mass and particle type by a set of PID detectors, followed by two types of electromagnetic calorimeters for photon and jet measurements. The forward muon arm ($2° - 9°$) consists of a complex arrangement of absorbers, a large dipole magnet and 14 planes of tracking and triggering chambers.

1.2.1 Tracking detectors

An ensemble of cylindrical detectors measures the passage of each particle carrying an electric charge at many points, so that its trajectory is precisely known.

The ALICE tracking detectors are embedded in a solenoidal magnetic field, produced by a huge red magnet, bending the trajectories of the particles; from the particles curvature we can find their momentum. The ITS provides the spatial precision in the tracking such that particles generated by the decay of other particles with a very short life time can be identified by seeing that they do not originate from the point where the interaction has taken place (the "primary vertex" of the event) but rather from a point at a distance of as small as a tenth of a millimeter.

1.2.1.1 Inner Tracking System (ITS)

The current ITS2 consists of seven cylindrical, concentric detector layers (see figure 1.2) of 50-μm thick CMOS monolithic active pixel sensors (MAPS), with a pixel pitch of about 30 x 30 μm^2 for a total of 12.5 Giga pixels covering an active surface of about 10 m^2. A key feature of the new ITS is the very low mass of the three innermost layers, which have a material budget of 0.3% X_0 per layer. It is also able to handle the hit densities resulting from Pb–Pb collisions at 50 kHz interaction rate. The radii of the inner, middle and outer layers range from 2 to 43 cm and their lengths are 30 cm, 90 cm and 150 cm, respectively.

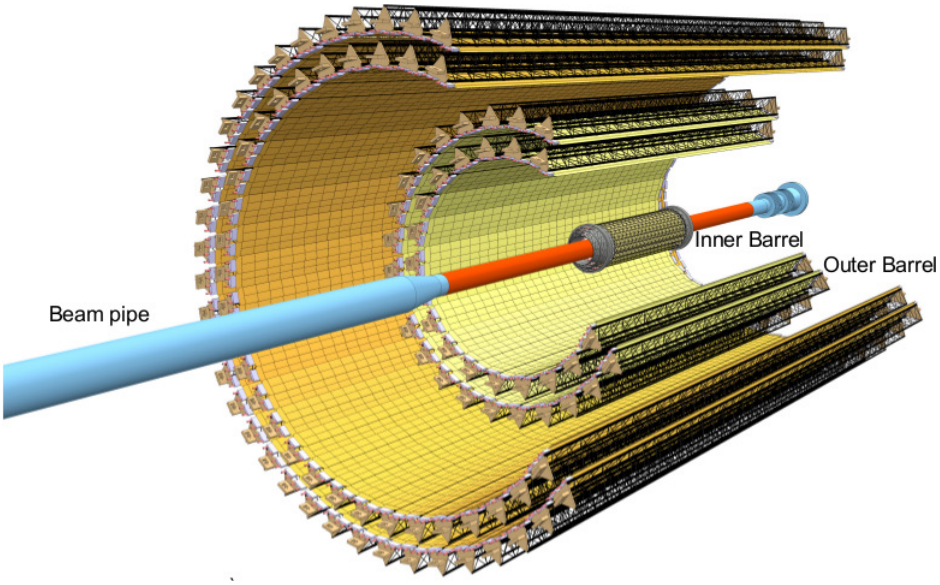


Figure 1.2: Layout of the current ALICE Inner Tracking System ITS2, showing seven cylindrical, concentric layers of monolithic active pixel sensors: the Inner Barrel, surrounded by the middle and outer layers of the Outer Barrel [2].

1.2.1.2 Time Projection Chamber (TPC)

In the radial direction, the ITS is followed by the Time Projection Chamber (TPC), the main tracking detector of the central barrel, extending from 0.85 m to 2.5 m in radius over a length of 5 m. The ALICE TPC is a 90 m^3 cylinder filled with gas and divided in two drift regions by the central electrode located at its axial centre. The field cage secures a uniform electric field along the $z-$axis. Charged particles traversing the TPC volume ionise the gas along their path, liberating electrons that drift towards the end plates of the cylinder. The necessary signal amplification is provided by Gas Electron Multiplier (GEM) detectors. There is a particle identification capability in the TPC, obtained from the measurement of the specific energy loss in the TPC gas.

1.2.2 Particle identification detectors

1.2.2.1 Transition Radiation Detector (TRD)

The Transition Radiation Detector (TRD) extends from 2.8 m to 3.5 m in radius and provides additional space points for tracking, which are also used to determine the size of the distortions due to space charge effects in the TPC, as well as dE/dx measurements for particle identification, and the detection of transition radiation for electron identification at $p_T > 3\ GeV/c$.

1.2.2.2 Time of Flight (TOF)

The subsequent Time-of-Flight detector (TOF) allows the identification of hadrons over a wide momentum range and electrons at low momentum. It measures the flight time of particles from the collision point out to the detector with precision of 80 ps; together with the momentum, this time determines the mass. It covers the central barrel over an area of 140 m^2 with 150 000 individual cells at a radius close to 4 m.

1.2.2.3 High Momentum Particle Identification Detector (HMPID)

The High Momentum Particle Identification Detector (HMPID) is a ring-imaging Cherenkov detector that adds hadron (pions, kaons and protons) identification capabilities at large transverse momenta over a limited acceptance. It has two main parts: a radiator medium, where the Cherenkov light is produced and a photon detector used to generate images of the Cherenkov light (rings). It has a surface of about 12 m^2 and consists on 7 proximity focusing Ring Imaging Cherenkov (RICH) counters of 1.4 m x 1.3 m each. Cherenkov photons, emitted when a fast charged particle traverses the 15 mm thick layer of liquid C_6F_{14} (perfluorohexane), are detected by a photon counter, which exploits the technology of a thin layer of CsI deposited onto the pad cathode of a multi-wire proportional chamber (MWPC).

1.2.3 Calorimeters

A large part of the acceptance in the central barrel is covered by electromagnetic calorimeters. The ElectroMagnetic Calorimeter (EMCal) is realised as Pb-scintillator sampling calorimeters with avalanche photon detector (APD) readout, whereas the PHOton Spectrometer (PHOS) uses $PbWO_4$ crystals with APD readout.

1.2.4 Muon detectors

The muon detectors cover the forward pseudorapidity range $-4.0 < \eta < -2.5$ and use a system of absorbers to remove hadrons and identify muons. The background of secondary muons from pion and kaon decays in the muon system is small at high p_T, thanks to the so-called 'muon plug' absorber, which is placed at $z = 90\ cm$ from the interaction point. The main muon detector stations use multiwire proportional chambers (muon tracking chambers, MCH), and resistive plate chambers (muon identifier, MID), both of which were equipped with new front-end electronics during the LHC LS2. The Muon Forward Tracker (MFT) has been added after the Run 2 and consists of tracking stations with the ALPIDE silicon pixel sensors that are installed in front of the muon plug to improve mass resolution and pointing resolution for the detection of secondary charmonia and muons from B-meson decays.

1.2.5 Forward and trigger detectors

A set of forward detectors form a Fast Interaction Trigger (FIT), which is used for triggering, event selection and determination of the collision time. The FIT system consists of two arrays of fast Cherenkov radiators placed on both sides of the interaction point (FT0), complemented with 3 sets of scintillator detectors. The interaction trigger is provided by the FT0 together with a large azimuthally segmented scintillator detector placed on the opposite side of the muon detectors, which is also used to determine the reaction-plane orientation in Pb–Pb collisions. Two additional scintillator detectors, FDD, are placed on opposite sides of the interaction point at large distances to cover $4.7 < \eta < 6.3$ and $-6.9 < \eta < -4.9$ to select diffractive and ultra-peripheral collisions with rapidity gaps. The Zero-Degree Calorimeters (ZDC) are installed at $\approx 100\ m$ on either side of the interaction point to help determine the centrality and event plane orientation. The readout electronics of the ZDC were upgraded to increase the readout rate to match the rest of the system.

1.2.6 ALICE upgrade prospect

In spite of the huge progress made so far by ALICE in the last years (for a detailed assessment of the main findings of the ALICE experiment in Run 1 and Run 2 see [20] and [21]), several fundamental questions still remain open and need to be addressed experimentally. For instance, low-p_T collective flow measurements are modeled successfully in terms of QCD hydrodynamics operating

at sub-fermi time scales, and the generic suppression of high-p_T hadron yields can be accounted for by jet quenching models in terms of medium-modified parton showers. However, the microscopic dynamics that gives rise to the fast hydrodynamisation is untested experimentally, and so is the question of how efficiently partons lose momentum, hydrodynamise, and thermalise in dense QCD matter. Also, the dynamical mechanisms that lead to a characteristic, event-multiplicity dependent hadrochemical composition remain unknown, and our understanding of heavy-flavour hadronisation, in particular in the baryon sector, is still incomplete.

ALICE in particular, aims, in Runs 3 and 4, to improve substantially the precision of heavy flavour production measurements and to measure for the first time the thermal emission of dileptons in heavy ion collisions at the LHC, both to determine the temperature in the early phase of the collisions and to study the mechanisms for chiral symmetry restoration in the regime of vanishing baryon density. Improved precision at high mass as well as larger statistics are also required to perform differential measurements of dielectron production and elliptic flow to trace the evolution of the emission of electromagnetic radiation with the temperature and to disentangle emission from the QGP and from the hadronic phase.

Existing measurements at the LHC have already shown that charm quarks lose energy as they propagate through the QGP and that charm baryon production is enhanced in central Pb-Pb collisions, suggesting that new hadronisation mechanisms, for example via combination of independently produced quarks, are at work, and that charm quarks may largely equilibrate in the plasma. The improved track impact-parameter resolution and readout rate of the upgraded apparatus will allow one to measure baryon to meson ratios in the charm sector, as well as dual flavour mesons, like the D_s, with sufficient precision to disentangle charm transport and hadronisation effects and to constrain the transport coefficients of charm quarks in the QGP.

With the next Inner Tracking System upgrade, referred to as ITS3, ALICE will perform first measurements of fully reconstructed beauty hadron decays. Beauty quarks are not expected to reach full thermal equilibirum, thus providing not only an important qualitative test of heavy quark transport, but also of hadronisation away from equilibrium. The study of QGP hadronisation via the combination of uncorrelated quarks is key to understand the dynamics connecting hadronisation and collective flow.

An important upgrade to ALICE to be realized during LS3 for data-taking in 2026–2028 at the LHC, is the installation of the forward electromagnetic and hadronic calorimeter (FoCal). The FoCal is a highly granular Si+W electromagnetic calorimeter combined with a conventional sampling hadronic calorimeter covering pseudorapidities of $3.4 < \eta < 5.8$. It will significantly enhance the scope of ALICE for inclusive and correlation measurements with mesons, photons, and jets to explore the dynamics of hadronic matter at small x (momentum fraction) down to about 10^{-6}.

The ALICE target integrated luminosities for the future physics program are listed in the following. For Pb–Pb collisions at $\sqrt{s_{NN}} = 5.3\ TeV$: $L_{int} =$

13 nb^{-1}; p–Pb collisions at $\sqrt{s_{NN}} = 8.5\ TeV$: $L_{int} \approx 0.5\ pb^{-1}$; pp reference samples at the same centre-of-mass energies as the Pb–Pb and p–Pb samples: $L_{int} \approx 6\ pb^{-}$ for each energy; pp collisions at top LHC energy $\sqrt{s} = 13.6\ TeV$: $L_{int} = 200\ pb^{-1}$; $^{16}\mathrm{O} - {}^{16}\mathrm{O}$ collisions at $\sqrt{s_{NN}} = 6.8\ TeV$ (few days in Run 3 with $L_{int} \approx 0.5 - 1\ nb^{-1}$).

The detector design was guided considering a series of test-bench observables to be measured in order to address the physics goals. These required charged particle tracking providing a transverse momentum resolution in the order of 1 – 2 % with excellent efficiency over a large pseudorapidity range. Besides the gain in the size of the statistical sample, an extended rapidity coverage is crucial for the low p_T reach as well as studies of long-range correlations and for studies of the rapidity dependence of the yields. In addition, ultimate track impact-parameter resolution is required ($\sim 2\ \mu m$ at $p_T = 1\ GeV/c$). This is essentially for an effective reconstruction of secondary vertices and decay chains, e.g. of heavy-flavoured hadrons. It is also a key ingredient for the rejection of heavy-flavour background for the dielectron measurements.

For what concerns the particle identification, electrons must be identified in a p_T range from $\sim 50\ MeV/c$ up to about $\sim 2\ GeV/c$. For muons the challenge is to achieve an efficient identification starting from $p_T \sim 1.5\ GeV/c$ at $\eta = 0$. The combination of electrons and muons at low and high transverse momenta, respectively, ensures access to leptonic channels over a wide p_T range. Hadron identification must span the range up to a few GeV/c to identify the products of heavy flavour decays, while retaining high efficiency. For the detection of photons, the requirements are different for different measurements. While the measurement of photon-jet correlations calls for a large acceptance, the reconstruction of states such as the χ_c relies in addition on the measurement of photons at low energies. The detection and energy determination of very low-p_T ($\geq 1\ MeV/c$) photons calls for a specially optimised detector in the forward direction to exploit the longitudinal boost.

1.3 Upgraded Inner Tracking System (ITS3)

1.3.1 Physics performance

As stated in the previous section, an improved vertex resolution will be beneficial for the measurement at low transverse momentum of key observables studied by ALICE. A refined impact parameter resolution and a reduced material budget of the ITS layout will significantly extend the physics capability for the study of heavy-flavour production, notably in the baryon sector, and for the study of low-mass dielectrons, two of the main objectives of the ALICE physics programme in the next decade. In this section the expected improvements of the performance will be outlined considering the case of Λ_c and thermal dielectrons production.

Measurement of Λ_c production

The measurement of the production yields and flow of charm and beauty baryons is of particular interest to study the thermalization and the mechanism of hadronization of c and b quarks in the QCD medium. In particular, if in Pb–Pb collisions heavy quarks can thermalize and hadronize via recombination with light-flavour quarks present inside the quark-gluon plasma (QGP) or at the QGP phase boundary, the production of charm and beauty baryons is expected to be significantly enhanced (compared to measurements in pp collisions) in the low and intermediate momentum region, say below 10 GeV/c. An enhancement of the baryon-to-meson ratio was measured in the light-flavour sector (p/π [22] and Λ/K^0_S [23]) for $1 < p_T < 6$ GeV/c and can be described in terms of light-quark recombination and hadron-mass-dependent radial collective flow [22, 23].

A precise measurement in the charm sector would provide crucial information on the charm quark thermalisation and hadronisation in the Quark-Gluon Plasma, as well as on the roles of recombination and radial flow. In addition, such measurements can provide further evidence for the deconfinement of charm quarks in the fireball, see, e.g., [24, 25, 26]. First measurements in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV [27, 28] and preliminary measurements in Au–Au collisions at $\sqrt{s_{NN}} = 200$ GeV [29] indicate a possible enhancement of the ratio Λ_c/D_0 for $p_T < 10$ GeV/c and no significant modification for higher p_T. These measurements have limited statistical precision and prevent a clear conclusion.

From the experimental point of view, the main issue for the measurement of charmed baryons is their rather short lifetime: the $c\tau$ of the Λ_c is about 59 μm, a factor of 2 smaller than that of the D_0 meson. In addition, unlike the D_0, the Λ_c does not have two-prong decays into charged hadrons. The most convenient decay channels are $\Lambda_c^+ \to pK^-\pi^+$, which has a large three-prong combinatorial background, and $\Lambda_c^+ \to pK^0_s$, for which a precise determination of the decay vertex is not possible because of the long decay length of the neutral kaon. The former is the benchmark channel considered for the ITS3 upgrade. The measurement requires very precise tracking and impact parameter resolution, because the decay tracks are typically displaced from the main interaction vertex by only a few hundreds of microns. This makes the measurement of the Λ_c a powerful benchmark to generally assess the improvement provided by the ITS3 in terms of physics performance for heavy-flavour physics.

A comparison of the performance between ITS2 and ITS3 is shown in figure 1.3 for the statistical significance $S/\sqrt{S+B}$ (left) and signal-to-background S/B ratio (right). S and B are the signal and background yields in an invariant-mass range of $\pm 3\sigma$ around the Λ_c nominal rest mass, where σ is the invariant-mass resolution (about 5 to 15 MeV/c^2 depending on p_T). The statistical significance is given for 8 billion events in the 0 – 10 % centrality class, corresponding to $L_{int} = 10nb^{-1}$. An improvement is found with ITS3 of a factor about four for the significance and ten for the S/B ratio, because the better track impact-parameter resolutions allows for a larger rejection of the combinatorial background and a larger efficiency for the signal selection.

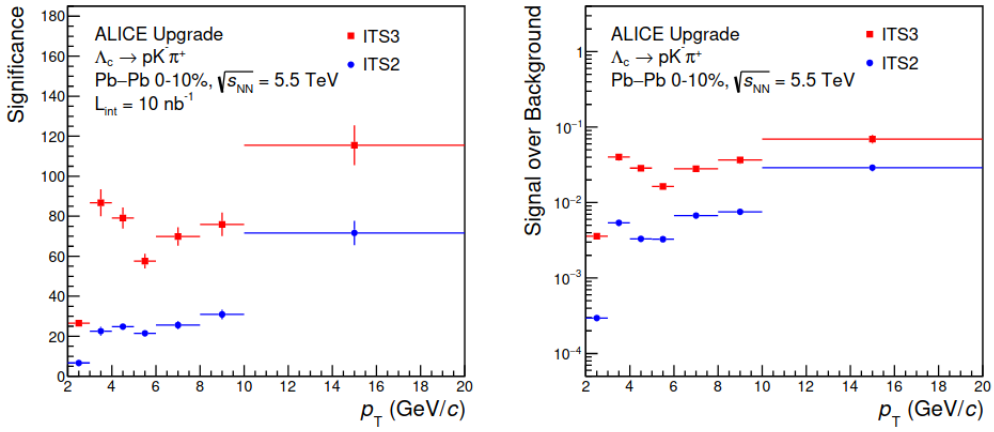


Figure 1.3: $\Lambda_c^+ \rightarrow pK^-\pi^+$ in central Pb–Pb collisions at $\sqrt{s_{NN}} = 5.5\ TeV$ ($L_{int} = 10nb^{-1}$): statistical significance (left) and S/B ratio (right) as a function of p_T [3].

Measurement of Thermal Dielectrons

Electromagnetic radiation produced by the high-temperature system formed in heavy-ion collisions can be detected using real direct photons with very low momentum or virtual photons yielding low invariant-mass dilepton pairs (M_{ee}). In this section we discuss the measurement using dielectron pairs in the ALICE central barrel, which requires acceptance for e^+e^- pairs at invariant masses and transverse momenta as low as possible. This implies electron detection down to $p_T < 100\ MeV/c$. Since the production rate of thermal dileptons is low, very good electron identification is mandatory to suppress the combinatorial background arising when one of the particles of the pair is a hadron misidentified as an electron. Moreover, electrons from π^0 Dalitz decays and photon conversions (mainly from $\pi^0 \rightarrow \gamma\gamma$) produce a large combinatorial background. This calls for low material budget specially of the innermost detection layer and for the capability to reconstruct e^+e^- pairs from photon conversions and Dalitz decays for rejection before the definition of dielectron candidates.

The proposed ITS3 upgrade will enhance the performance of the thermal dielectrons measurement with respect to the ITS2 because:

- The reduction of the material budget of the beampipe by 30% and of the first detection layer by a factor of about six determines a reduction by a factor of about three of the number of electrons from photon conversions before the first track measurement point.

- The enhanced low-p_T tracking capability of the ITS3 substantially improves the reconstruction efficiency of photon conversions in which one of

the particles of the dielectron pair often has very low momentum. This reduces the combinatorial background.

- The improved track impact-parameter resolution of the ITS3 enables efficient tagging of electrons from semi-leptonic charm decays, which form a substantial physical background to the pair-yield and dominate the invariant-mass spectrum for $M_{ee} > 1.1\ GeV/c^2$.

1.3.2 Tracking performance

An important measure of the achieved tracking precision is the track impact-parameter resolution, defined as the dispersion of the distribution of the Distance of Closest Approach (DCA) of the reconstructed (primary) tracks to the interaction vertex. It is the parameter that defines the capability of a vertex detector to separate secondary vertices of heavy-flavour decays from the interaction vertex. A comparison of the impact-parameter resolution of ITS2 and ITS3 is shown in figure 1.4 for pions at typical pseudorapidity $\eta = 0.5$. The solid lines show the resolution for tracks fitted in the seven ITS layers only, while the dashed lines show the resolution for tracks fitted in both the TPC and the ITS. For $p_T \sim 1\ GeV/c$, the impact parameter resolution of ITS3 plus TPC is a factor of about two better than that of ITS2 plus TPC.

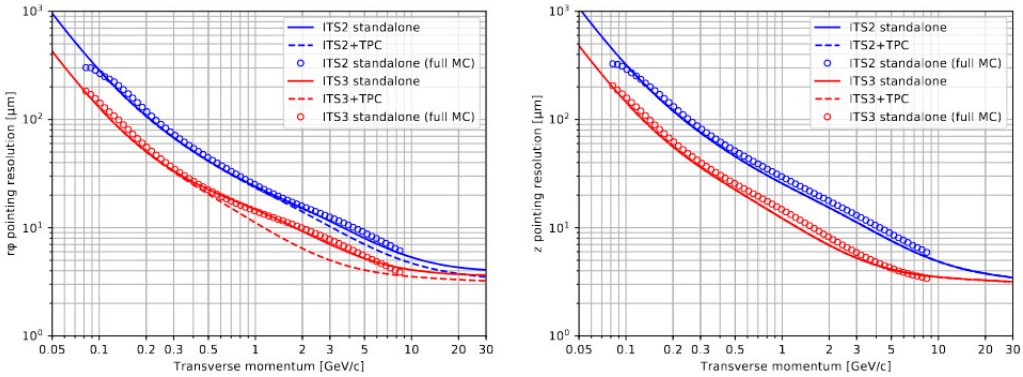


Figure 1.4: Impact parameter resolution for primary charged pions as function of the transverse momentum for the ITS2 upgrade (blue) and the ITS3 upgrade (red) in the transverse plane (left panel) and in the longitudinal direction (right panel). All solid lines show the results with FMCT and ITS only, all dashed lines show the results with FMCT and ITS+TPC, all open circles show the results with full MC and ITS only. [3].

A Fast Monte Carlo Tool (FMCT) was used in detector design phase to optimize the number of layers, their radial position, the material budget and the detector resolutions. The latter, for the ITS-only case, were also estimated with

a full MC simulation, where pions were transported through the detector material using the GEANT3 package [30] and their trajectories were reconstructed using the Cellular Automaton ITS tracker [2]. The Full MC resolutions are shown by the circles for both ITS2 and ITS3 in figure 1.4, while FMCT results are shown as solid or full lines. The two results are found to be very similar, validating the conclusions drawn from the fast simulation tool.

The left panel of figure 1.5 shows the track-finding efficiency in the environment of a single central Pb–Pb collision. The plot shows the efficiency only for "correct tracks", i.e. tracks that contain all clusters produced by the same particle, while at low p_T there will be a significant contribution from tracks with one or more clusters from other particles. The p_T resolution of the ITS3, shown on the right panel of figure 1.5, benefits significantly from the lower material budget of the Inner Barrel in the ITS-only mode. In the ITS–TPC combined tracking mode, at p_T below 0.5 GeV/c, the ITS3 configuration essentially gives the same p_T resolution as ITS2.

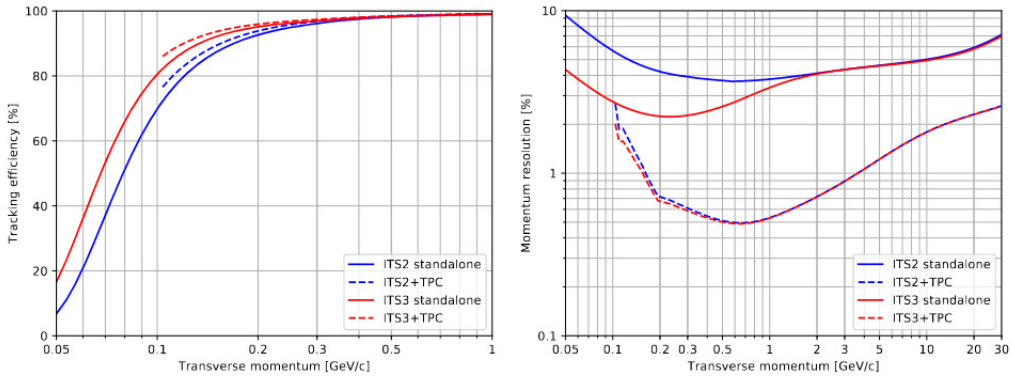


Figure 1.5: Track finding efficiency in central Pb–Pb collisions (left panel) and transverse momentum relative resolution (right panel) in the ITS-only (solid lines) and ITS–TPC combined (dashed) reconstruction modes as function of transverse momentum for the ITS2 and two different ITS3 detector configurations. All solid lines show the results with FMCT and ITS only, all dashed lines show the results with FMCT and ITS+TPC. [3].

1.3.3 Upgrade concept

The ITS2

The current ITS2 is formed by seven layers, grouped into two separate barrels. The Inner Barrel (IB) consists of the three innermost layers (Layers 0 to 2), while the Outer Barrel (OB) contains the two middle and two outer layers (Layers 3 to 6) (See figure 1.2). The three Layers of the IB are also referred to as the Inner Layers, while, for the OB, Layers 3 and 4 are referred to as Middle Layers

and Layers 5 and 6 as Outer Layers. In the azimuthal direction, each Layer is segmented in elements called Staves. The Stave, which extends over the whole length of the respective Layer, is the basic building block of the detector. The Stave contains all structural and functional components, thus making it the smallest operable part of the detector.

The three Inner Layers are built with identical Staves, which are based on the following elements:

- Space Frame: a carbon fibre support structure providing the mechanical support and the necessary stiffness.
- Cold Plate: a sheet of high thermal-conductivity carbon fibre with embedded polyimide cooling pipes, which is integrated into the Space Frame. The Cold Plate is in thermal contact with the Pixel Chips to remove the generated heat.
- Hybrid Integrated Circuit (HIC): an assembly consisting of a polyimide Flexible Printed Circuit (FPC) onto which the Pixel Chips and some passive components are bonded.

Each IB Stave is instrumented with one HIC, which consists of nine ALPIDE chips in a row connected to the FPC, covering an area of 15 mm $\times$ 271.2 mm including 100 μm gaps between adjacent chips along the longitudinal direction (z). The interconnection between Pixel Chips and FPC is implemented via conventional aluminum wedge wire bonding. The HIC is glued to the Cold Plate with the Pixel Chips facing it in order to maximize the cooling efficiency. A schematic drawing of the IB Stave is shown in figure 1.6. An extension of the FPC, like shown in figure 1.7, connects the Stave to a patch panel that is served by the electrical services entering the detector only from one side.

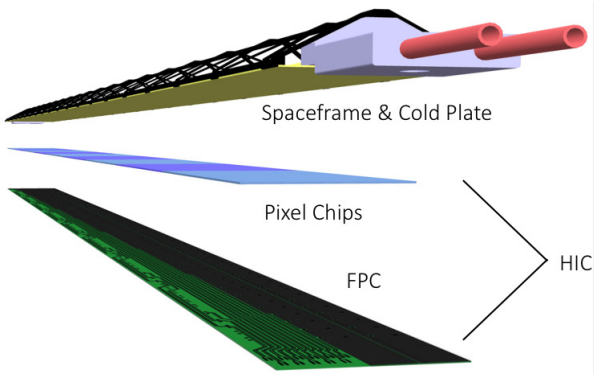


Figure 1.6: Schematic layout of the ITS2 Inner Barrel Stave. Nine pixel sensors are mounted on a flexible printed circuit (FPC) to form a hybrid integrated circuit (HIC). The HIC is glued on a carbon fibre support structure (Space Frame), which integrates a water cooling circuit (Cold Plate) [3].

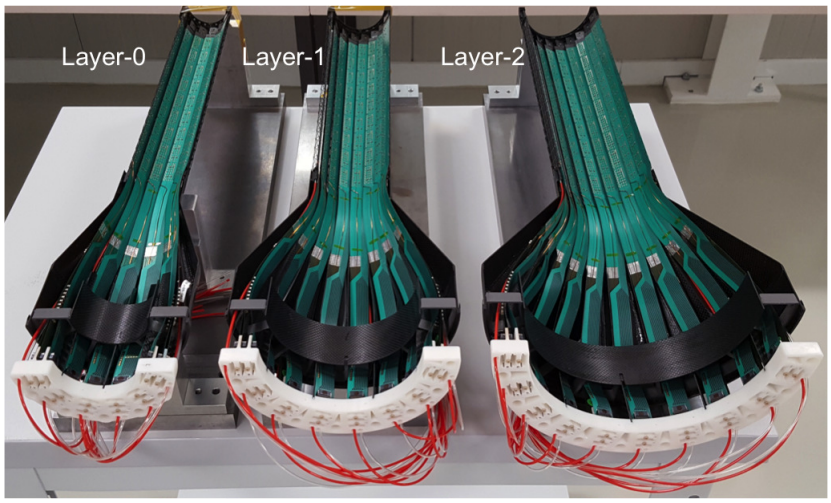


Figure 1.7: Photograph of the ITS2 IB half-layers (Layer-0, Layer-1 and Layer-2). [3].

Figure 1.8 shows the azimuthal distribution of the material of Layer 0 traversed by particles with $|\eta| < 1$. Adjacent Staves are partially overlapping to ensure detector hermeticity. The highest peaks correspond to the overlap of the reinforced structures along the edges of the Space Frame, which guarantee the required stiffness, while the narrow spikes correspond to the reinforcement of the upper central vertex. The peaks at a value of about 0.5% X_0 are due to the polyimide cooling pipes embedded in the Cold Plate, which have an inner diameter of 1.024 mm and a wall thickness of 25 μm and are fully filled with water during the operation of the detector. The average value (0.35% X_0) is unprecedentedly low, which is of crucial importance to achieve high impact parameter resolution at low transverse momentum.

The table 1.1 summarizes the layout parameters of the ITS2.

Table 1.1: Layout parameters of the ITS2. [31]

	Inner Barrel			Outer Barrel			
	Layer 0	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5	Layer 6
Length [cm]		27.1		84.3		147.5	
Radial position [cm]	2.3	3.1	3.9	19.4	24.7	35.3	40.5
N. staves	12	16	20	24	30	42	48
N. HICs	12	16	20	384	480	1176	1344
N. pixel chips	108	114	180	2688	3360	8232	9408

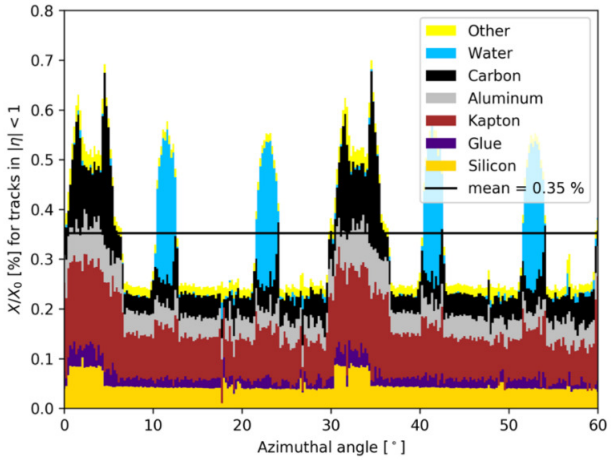


Figure 1.8: Azimuthal distribution (averaged over $|\eta| < 1$) of the material budget of ITS2 Layer 0. The angular interval in the figure corresponds to two staves, with $\varphi = 0$ on the horizontal plane, where two staves from the upper and lower half barrels partly overlap. [3].

Improvement strategies

As shown in the material breakdown of the ITS2, the silicon sensor, which has a thickness of 50 μm and ideally is the only component that would be needed in the detector acceptance, contributes only 15 % to the total. The rest is due to the material of the electrical substrate (Flexible Printed Circuit, FPC), which represents the main contribution (50 %) including the passive components and the glue, of the cooling circuit (20 %) and of the carbon spaceframe (15 %). In order to significantly reduce the material budget, the electrical, mechanical and cooling material have to be reduced. The electrical substrate is the flexible printed circuit that distributes the supply and bias voltages as well as the data and control signals to the pixel sensors. In standard CMOS circuit manufacturing, the maximum size of a chip is limited to the reticle area defined by the field of view of the photolithographic process, which is typically a few centimeters in both directions. For this reason, the Stave surface is covered by tiling several sensors. The reticle is just stepped and repeated across the wafers to create multiple identical images of the same circuit or group of circuits. Stitching is a technology that allows the fabrication of an image sensor that is larger than the field of view of the lithographic equipment [32]. In this technology, the reticles which fit into the field of view of that equipment are placed on the wafer with high precision, achieving a tiny but well defined overlap. In this way, wafer-scale sensors can be manufactured, allowing the fabrication of sensors of arbitrary dimensions, the only limit being the size of the wafer. This opens the possibility of fabricating a pixel sensor with the dimension of an entire stave. The distribu-

tion of power and electrical signals could then be done entirely inside the silicon chip and the electrical substrate would terminate close to the chip edge, where the interconnections to the chip would be realized.

Concerning the elimination of the material associated with the cooling circuit, the possibility of using a low-speed ($< 2\ ms^{-1}$) air flow to remove the heat produced by the ITS Inner Layers, in combination with peripheral liquid cooling, was experimentally demonstrated as a viable option for sensors with a power density below 20 $mWcm^{-2}$ [3]. It should be noticed that the higher the air flow, the stiffer the mechanical support structure has to be, in order to keep the vibrations of the sensor, caused by the air flow, smaller than the intrinsic spatial resolution.

Each ALPIDE chip contains approximately 5×10^5 pixels, each measuring about 27 $\mu m \times 29\ \mu m$, arranged in 512 rows and 1024 columns, for a total active area of 30 $mm \times 15\ mm$. The total power consumption is 180 mW, which corresponds to a power density of about 40 $mWcm^{-2}$. However, most of this power (about 150 mW) is dissipated by the digital interface circuitry and the high-speed output data links, which are located in a small area of about 30 $mm \times 1.5\ mm$ close to one edge of the chip. Only about a sixth of the total power is dissipated in the pixel matrix, which corresponds to a power density of about 7 $mWcm^{-2}$. This would be well within the range for which air cooling represents a viable and effective solution. On the other hand, to place the pixel sensor digital periphery at the end of the sensor would be possible if the sensor had the dimensions of a stave.

As mentioned before, at present small sensors are mounted edge-to-edge on top of an FPC module that provides the power distribution and data bus. The key new idea is to make use of the stitching technology avoiding the need of an FPC and realizing a single large sensor, where power distribution is managed internally, confining the interconnections to the outside world to the sensor edge.

A first step in these directions would be to develop a chip with stitching in 1 dimension according to the scheme outlined in figure 1.9 of size 15 $mm \times 280\ mm$, which corresponds to the full length of a stave. The columns run along the short side of the sensor, and have the same length as in the current ALPIDE sensor. Data are extracted from the matrix as in the ALPIDE chip with hit-driven circuitry based on a priority-encoder addressing scheme [33]. Groups of 16 double columns are read out sequentially, but in principle all of them could be read out in parallel[9]. Data are transmitted from the bottom of the columns along one long side of the sensor to the periphery, as indicated by the green arrow at the bottom of figure 1.9.

The periphery, shown on the left part of the sensor in figure 1.9, contains the control logic to steer the priority encoders, the interface for the configuration of the chip and the serial data transmitters. Since the chip periphery is outside the detector acceptance, there is no particular constraint on the size of the periphery that will be adjusted on the basis of the requirements.

The data bus is conservatively estimated to take an area of up to 50 μm

[9] A more detailed description of the ALPIDE chip readout will be given in a next chapter.

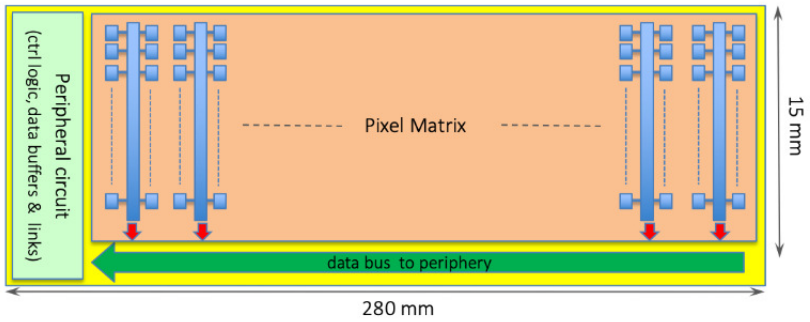


Figure 1.9: Diagram of stitched sensor in one direction (horizontal and vertical dimensions not to scale). Stitching in the vertical direction is also possible. [3].

$\times$ 280 mm, which amounts to a dead area that corresponds to about 0.3 % of the detector acceptance. As a comparison, the ITS2 has gaps in the z-direction between adjacent chips (matrix to matrix) of 100 μm, which result in a total dead area of 0.5 %.

The distribution of power inside the sensor also has to be considered carefully. In the ALPIDE, there are two metal meshes with pads all over the chip surface. In principle, the technology offers the possibility to add a metal layer (called re-distribution layer) underneath the chip passivation layer. The meshes or distribution layer are then connected to the power pads located in the chip periphery. A flexible PCB, connected to the digital periphery of the chip, will provide power to the sensor.

A 2D-stitched wafer-scale sensor can be obtained replicating the 1D-stitched sensor chip (including the data bus) several times along the periphery side. A 300 mm wafer can contain a rectangular matrix of about 280 mm $\times$ 100 mm, or about 280 cm^2. For example, the sensor for Layer 0 would consist of three instances of the 1D circuit all integrated in the same silicon chip. In this way, it would be possible the realization of a sensor chip of the same size as that of a half-layer of the Inner Barrel.

The large sensor will be thinned to values of about 20–40 μm, which correspond to a radiation thickness of 0.2–0.4 ‰ X_0. The possibility to bend and operate ultra-thin sensors to a curvature radius of about 20 mm seems very promising [34], opening the way to the construction of a silicon only cylindrical layer.

Technology options

The baseline technology for the development of the proposed sensor is the 65 nm CMOS process[10] of Tower Partners Semiconductor Co., Ltd. (TPSCo), which

[10]The feature size of any semiconductor technology is defined as the minimum length of the MOS transistor channel between the drain and the source.

offers a number of advantages with respect to the 180 nm process, which is the fallback option to 65 nm and used for the development of ALPIDE:

- The process, which uses wafers that are 300 mm in diameter, allows the realization of the entire half-cylinder as a single chip, while the 180 nm process uses wafers that are 200 mm in diameter, which would not only complicate the mechanical and electrical integration, but also increase the material budget as it would require additional electrical interconnections.
- Owing to the smaller feature size of the transistors, the pixel pitch can be reduced by a factor larger than two. This allows a significant reduction of the charge collection time and also a better position resolution.
- Moreover, the smaller the channel length, the faster the transient response of the transistors, which would translate into better performance.

The 180 nm option is estimated to be slightly cheaper. The 65 nm option, however, will not only allow a simplification of detector layout and assembly, but will also enable the realization of a circuit with significantly enhanced performance, also in terms of time resolution and of rate capabilities.

1.3.4 Layout of the future ITS3

The ITS3 will consist of two separate barrels, referred to as Inner Barrel and Outer Barrel. The Outer Barrel, containing the four outermost layers (Layer 3 to Layer 6), will remain that of ITS2. A completely new Inner Barrel, consisting of the three innermost layers (Layer 0 to Layer 2), will instead replace the current Inner Barrel of ITS2. The ITS3 IB will consist of two halves, named half-barrels, to allow the detector to be mounted around the beampipe. Each half-barrel will consist of three half-layers. The half-layers are arranged inside the half-barrel as shown in figure 1.10. Each of these consists of a single large pixel chip, which is curved to a truly (half-) cylindrical shape.

The main structural components of the new Inner Barrel are the End-Wheels and the outer Cylindrical Structural Shell (CYSS), both made of Carbon Fibre Reinforced Plastic (CFRP) materials, and a series of ultra-lightweight half-wheel spacers, made of open cell carbon foam, which are inserted between layers to define their relative radial position (figure 1.11).

The main layout and geometrical parameters of the ITS3 Inner Barrel are summarized in table 1.2.

The heat dissipated by the sensors is removed by convection through a forced airflow between the layers; in addition, the carbon foam rings in thermal contact with the sensors act as radiators reducing the thermal gradient along the layer. The enhancement of convective heat transfer due to the carbon foam rings results from the passage of the air through the open-interconnected void structure.

A combination of excellent thermal conductivity with low density has made carbon foam materials prime candidates for use in modern heat exchangers.

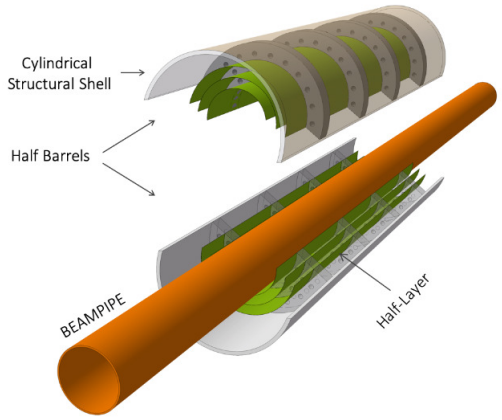


Figure 1.10: Layout of the ITS3 Inner Barrel. The figure shows the two half-barrels mounted around the beampipe [3].

Materials such as KFOAM feature a thermal conductivity similar to that of aluminum at one-fifth of the density and with a coefficient of thermal expansion that is close to that of silicon. POCO HTC, which is a new porous graphite material specifically designed for high thermal performance, shows two-thirds of the thermal conductivity of copper at only one-tenth of the weight [36, 37]. During the design process, the airflow speed will be optimised to guarantee the thermal performance and mechanical stability by avoiding airflow-induced vibrations.

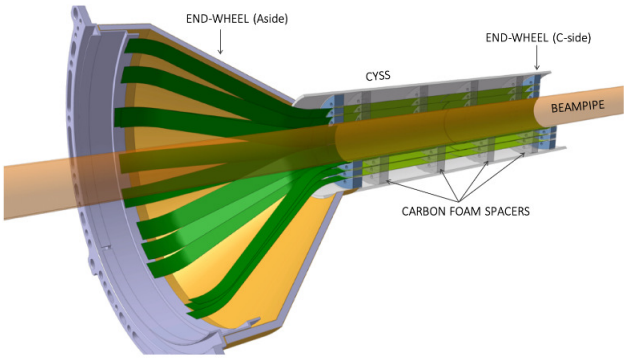


Figure 1.11: Layout of the ITS3 Inner Barrel for the baseline scenario. Two end-wheels and the CYSS provide precise position of the detector relative to the beampipe. [3].

Table 1.2: Geometrical parameters of the upgraded ITS. [35]

Beampipe inner/outer radius	16.0/16.5		
IB Layer parameters	Layer 0	Layer 1	Layer 2
Radial position (mm)	18	24	30
Length (sensitive area) (mm)	270	270	270
Pseudo-rapidity coverage[a]	$\pm$2.5	$\pm$2.3	$\pm$2.0
Active area (cm^2)	305	408	508
Pixel sensors dimensions (mm^2)	280 x 56.5	280 x 75.5	280 x 94
Number of pixel sensors / layer		2	
Pixel size (μm^2)		O(15x15)[b]	

[a] The pseudorapidity coverage of the detector layers refers to tracks originating from a collision at the nominal interaction point ($z = 0$).
[b] For the fallback solution the pixel size is about a factor two larger (O(30x30)μm^2).

Chapter 2

Silicon Pixel Detectors in High Energy Physics

An overview of the interaction mechanism of radiation with matter, the working principle of a pn junction, i.e. the sensor key structure, and the properties of Silicon is given in Appendix A. The physics principles that make up pixel detectors will be reviewed in this chapter and a brief introduction on the Complementary Metal-Oxide-Semiconductor (CMOS) technology will be also provided.

The process and operation principle of Monolithic Active Pixel Sensors (MAPS) will be described in detail, as well as the main radiation-induced effects which degrade their performance and their implications in High Energy Physics (HEP) experiments.

Given the scope of this work to study the possible effects of bending silicon senors, the property of piezoresistivity in silicon and the effect of mechanical stress on field-effect transistors will be also introduced throughout the chapter.

2.1 Silicon detectors

Silicon is the dominant semiconductor material used in the production of position sensitive detectors for particle physics. Silicon detectors are considered as solid state ionization chambers. They offer a great number of advantages with respect to other materials, e.g. the moderate band gap of 1.12 eV is large compared to the thermal energy at room temperature of 25.9 meV, which allows to operate the detector at room temperature. Compared with gaseous ionization detectors, the density of a semiconductor detector is very high, and charged particles of high energy can give off their energy in a semiconductor of relatively small dimensions. Silicon has a density of 2.329 g/cm^3, and the average energy loss per unit of length allows for building thin detectors that still produce measurable signals.

2.1.1 Signal formation in silicon sensors

The signal of a silicon sensor is defined as the induced current on the electrodes; so, even if we talk about charge collection, signal does not start when the charge is collected, but just when the charge begins to move inside the sensor. Consequently, it stops when the whole charge is collected (the so called charge collecting time). When a charge moves through the sensor, the induced charge on the electrodes changes.

The general method to calculate induced charge on electrodes due to the motion of charge carriers in a detector makes use of Ramo's theorem [38]. He devised a method of computing the signal induced by charge between two conductors with an electric field between them. In 1939 a simple equation for the contribution of a moving electron in vacuum to the current in the lead of an electrode A was proposed, which is also valid in the case of more than 2 electrodes:

$$i_A = e\overrightarrow{E}_u \cdot \overrightarrow{u} \tag{2.1}$$

where i_A is the instantaneous current received by the given electrode due to a single electron's motion, e is the charge of the electron, $\overrightarrow{u}$ is its instantaneous velocity, and E, is the component in the direction $\overrightarrow{u}$ of that electric field which would exist at the electron's instantaneous position under the following circumstances: electron removed, given electrode raised to unit potential, all other conductors grounded. Ramo's theorem can be applied to predict the amplitude of the signal induced on one of the pixel electrodes and with a reverse bias applied on the substrate. For example, considering the case of creating electron-hole pairs from an ionizing event close to the cathode surface and along an axis that is perpendicular to the center of the pixel. The holes will travel only a short distance to the cathode and will not make any or very little contribution, but the induced current observed at the pixel electrode will be entirely due to the motion of the electrons into the volume of the sensor.

The amplitude of the signal read out by the circuitry is proportional to the charge collected by the pixel. The sensor converts the energy deposited/induced by a particle ionizing charge (or photon) to an electrical signal. The signal charge can be quite small; therefore it must be amplified. Usually, the pre-amplifier is configured as an integrator, which converts the narrow current pulse from the sensor into a step impulse with a long decay time. In HEP detectors the primary function of the pulse shaper is to improve the signal to noise ratio. However, the pulse shape should be compatible with the digitizer or device for interconnection. Analogue to digital conversion translates a continuously varying amplitude to discrete steps, which correspond to a unique output bit pattern. Finally, the digital information is ready for subsequent storage and analysis. There are different kinds of detectors, but many of them have some similarities in their basic functions [39].

2.2 CMOS technology

The CMOS (Complementary Metal-Oxide-Semiconductor) is a technological approach allowing the realization of integrated circuits containing two polarities of Metal Oxide Transistors MOS, patterned on the same chip. The sandwich of metal, silicon oxide, and semiconductor, called MOS structure, is the basic building block of the MOS Field Effect Transistors (MOSFETs) [39, 40, 41] which are widely used in microelectronics. Complementary Metal Oxide Semiconductor transistor consists of P-channel MOS (PMOS) and N-channel MOS (NMOS). Figure 1 shows a basic scheme of each MOS polarity.

Figure 2.1: Basic scheme of NMOS (left) and PMOS (right) transistors. [42].

2.2.1 Field Effect Transistors (FETs)

A Field Effect Transistor (FET) (Figure 2.2) is a three-terminal active semiconductor device, where the output current is controlled by an electric field generated by the input voltage. FETs are also known as unipolar transistors because, unlike bipolar transistors, FETs only have either electrons or holes operating as charge carriers. FET uses the voltage applied to its input terminal (called the Gate), to control the current flowing from the source to drain, making the Field Effect Transistor a "Voltage" operated device. The device consists of an active channel through which charge carriers, electrons or holes, flow from the source to the drain. Source and drain terminal conductors are connected to the semiconductor through ohmic contacts. The conductivity of the channel is a function of the potential applied across the gate and source terminals. MOSFET transistors are formed as a sandwich consisting of a semiconductor layer, usually a slice, or wafer, from a single crystal of silicon, together with a layer of silicon dioxide (SiO_2) and a layer of metal. These layers are patterned in a manner which permits transistors to be formed in the semiconductor material substrate.

As mentioned above, CMOS circuits use a combination of PMOS and NMOS to implement logic gates and other digital/analogue circuits. NMOS is built on a p-type substrate with n-type source and drain diffused on it. In NMOS, the majority carriers are electrons. When a high voltage is applied to the gate, the NMOS will conduct. Similarly, when a low voltage is applied to the gate,

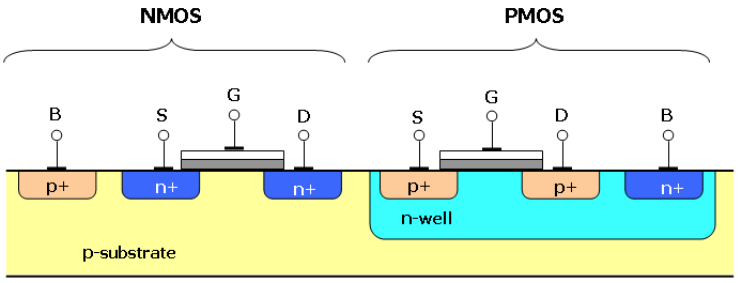


Figure 2.2: Cross-section of NMOS and PMOS field effect transistors. Each MOSFET has 4 terminals called body (B), source (S), gate (G) and drain (D). Note that the gate terminal is electrically insulated from the semiconductor by a layer of oxide (shown in gray) [43].

NMOS will not conduct. NMOS are considered to be faster than PMOS, since the carriers in NMOS, which are electrons, travel faster than holes. P-channel MOSFET consists P-type Source and Drain diffused on an N-type substrate. Majority carriers are holes. When a high voltage is applied to the gate, the PMOS will not conduct. When a low voltage is applied to the gate, the PMOS will conduct. A Complementary MOS circuit has almost no static power dissipation, power is only dissipated in case the circuit actually switches, this allows integrating more CMOS gates on an integrated circuit, resulting in much better performance.

2.2.2 Piezoresistance in MOSFETs

The effect of mechanical stress on field-effect transistors is of particular interest because stress causes changes in the parametric behavior of individual devices as well as analogue and digital circuits. In this section, we explore the basic dependence of MOSFET drain current on mechanical stress using the standard drain current expressions for an n-channel MOSFET.

$$\begin{aligned} &\text{Linear Region} : I_D = K\left(V_{GS} - V_T - \frac{V_{DS}}{2}\right)V_{DS} \\ &\text{Saturation Region} : I_D = \frac{K}{2}(V_{GS} - V_T)^2 \end{aligned} \tag{2.2}$$

$$\text{where} \quad K = \mu_n C_{ox}\frac{W}{L} = \mu\frac{\varepsilon_{ox}}{t_{ox}}\frac{W}{L}$$

In equation 2.2, V_T is the threshold voltage of the transistor, V_{GS} is the gate-to-source voltage, V_{DS} stands for the drain-to-source voltage, μ_n represents the carrier mobility in the channel, C_{ox} is the oxide capacitance, ε_{ox} is the oxide

permittivity or dielectric constant of the oxide layer, t_{ox} is the oxide thickness, and W and L are the width and length of the channel respectively. Similar equations apply to p-channel devices. From equation 2.2, it can be seen that stress may affect the drain current through mobility changes, threshold voltage changes or dimensional changes. The experiments show that the change of I_D is much grater than what can be explained by geometrical distortion, therefore we must look for an explanation based on variation of μ and V_T. Thus, for small variations in the saturation region, the normalized change in drain current can be written as:

$$\frac{\Delta I_D}{I_D} \cong \frac{\Delta \mu}{\mu} - 2\frac{\Delta V_T}{V_T}\left(\frac{V_T}{V_{GS} - V_T}\right) \tag{2.3}$$

Threshold Voltage Invariance

The variations of threshold voltage and mobility under stress can be determined in an independent manner [44, 45] by measuring the slope and intercept of the linear region transfer characteristic as a function of applied uniaxial stress as indicated in figure 2.3.

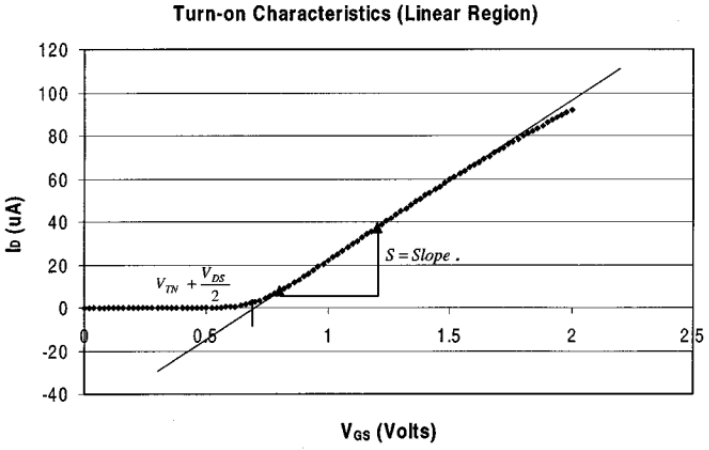


Figure 2.3: Linear region transfer characteristic [46].

From the linear region expression in 2.2, the x-intercept occurs for $V_{GS} = (V_T + V_{DS}/2)$, and the slope of the characteristic is $S = \mu_n C_{ox}(W/L)V_{DS}$. Experimental values of threshold voltage versus stress have been extracted from linear region measurements for devices with channel lengths (at the mask level) ranging from 15.1 μm down to 0.3 μm and a sample set of results are presented in figure 2.4 for NMOS devices from the IBM process. These results

demonstrate that threshold voltage is essentially independent of stress, which is consistent with the theory of Mikoshiba [47]. Similar results have been obtained for PMOS devices over the full channel length mentioned above [46]. The threshold voltages are also independent of the channel orientation.

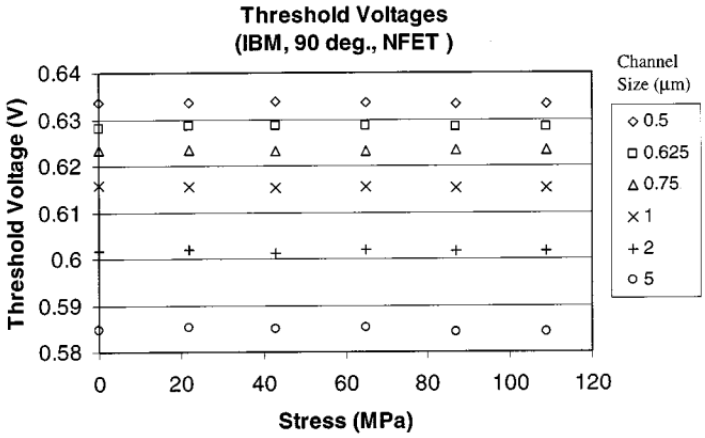


Figure 2.4: Example of threshold voltage versus applied uniaxial stress for NMOS devices in the IBM process [46].

Mobility Variation

Based upon these results, the second term in 2.3 can be neglected, and it will be assumed from here on that the variation in MOS drain current under stress results from mobility variations alone:

$$\frac{\Delta I_D}{I_D} = \frac{\Delta \mu}{\mu} \tag{2.4}$$

The piezoresistive response of FETs is dependent on the orientation of the device channel with respect to the crystallographic axes [46]. Here we are assuming that the stresses are resolved in a natural coordinate system aligned parallel (x_1', $\phi = 0^o$) and perpendicular (x_2', $\phi = 90^o$) to the wafer flat as indicated in figure 2.5 for a (100) silicon wafer. The axes correspond to edges of rectangular die as normally positioned on the wafer. The third axis, x_3', is perpendicular to the surface of the wafer.

When operating in strong inversion in either the linear or saturation region, the current in the field effect transistor is controlled by the resistive portion of the channel, and based on direct analogy to resistor variations [44, 45, 48], the normalized variations in drain current for MOSFETs with channels oriented at 0 and 90 are given by:

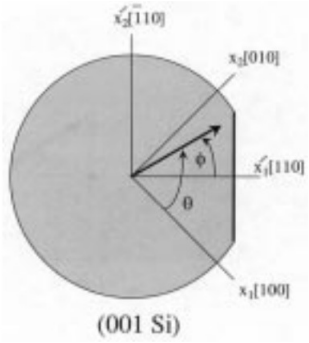


Figure 2.5: Coordinate system for a (100) wafer [46].

$$\left.\frac{\Delta I_D}{I_D}\right|_0 = \left.\frac{\Delta\mu}{\mu}\right|_0 = \frac{\Pi_S}{2}(\sigma'_{11} + \sigma'_{22}) + \frac{\Pi_{44}}{2}(\sigma'_{11} - \sigma'_{22}) + \Pi_{12}\sigma'_{33}$$
$$\left.\frac{\Delta I_D}{I_D}\right|_{90} = \left.\frac{\Delta\mu}{\mu}\right|_{90} = \frac{\Pi_S}{2}(\sigma'_{11} + \sigma'_{22}) - \frac{\Pi_{44}}{2}(\sigma'_{11} - \sigma'_{22}) + \Pi_{12}\sigma'_{33} \tag{2.5}$$

where σ'_{11}, σ'_{22} and σ'_{33} are the three normal stresses resolved in the primed coordinate system indicated in the figure 2.5. The Π_{ij} represent the piezoresistive coefficients of the FET. These devices do not exhibit a dependence on the in-plane shear stress σ'_{12} that would appear in devices with channels oriented in directions other than 0^o or 90^o (e.g. $\pm 45^o$).

The piezoresistive coefficients of NMOS devices have been found by applying a sequence of controlled uniaxial stresses to the transistors and measuring the corresponding changes in the slope of the turn-on characteristic in figure 2.3. Under uniaxial stress with $\sigma'_{11} = \sigma$, the normalized mobility variations in 2.5 reduce to:

$$\left.\frac{\Delta\mu}{\mu}\right|_0 = \left(\frac{\Pi_S + \Pi_{44}}{2}\right)\sigma \quad \text{and}$$
$$\left.\frac{\Delta\mu}{\mu}\right|_{90} = \left(\frac{\Pi_S - \Pi_{44}}{2}\right)\sigma \tag{2.6}$$

From 2.6, the effective value of piezoresistive coefficient determined from the measurements depends on device orientation. For a device with a 0^o orientation, $\Pi_{eff} = (\Pi_S + \Pi_{44})/2$ and for a 90^o orientation, $\Pi_{eff} = (\Pi_S - \Pi_{44})/2$. The following steps are used to determine the effective values of the piezoresistive coefficients:

1. Measure the I_D-V_{GS} turn-on characteristics by sweeping the gate voltage and measuring the drain current;

2. Extract the slope and intercept from the linear region characteristics below the inflection point to determine the threshold voltage and apparent mobility;

3. Repeat steps 1)-2) for the desired range of stresses;

4. Determine the effective piezoresistance coefficients by plotting the normalized changes in mobility versus applied stress;

5. Repeat steps 1)-4) for all device sizes.

The value of $\Delta\mu/\mu$ is plotted versus the applied uniaxial stress, and the slope is determined from a least squares fit to the data. Using the 0^o and 90^o device pairs, the individual values of Π_S and Π_{44} can be determined from the sum and difference of the two slopes.

Four-point bending testing has been used to generate the required uniaxial loadings [49]. In the four-point-bending method, rectangular strips containing a row of chips are cut from the wafer and placed in the bending fixture to generate the desired uniaxial stresses. A portion of a typical family of transfer characteristics as a function of uniaxial stress are shown in figure 2.6 (left) for a 90^o NFET with channel length of 0.625 μm and the values of the normalized mobility changes appear in 2.6 (right). The slope of the characteristic is 328 $(TPa)^{-1}$. For a similar NFET oriented at 0^o, the slope of the characteristic is 455 $(TPa)^{-1}$.

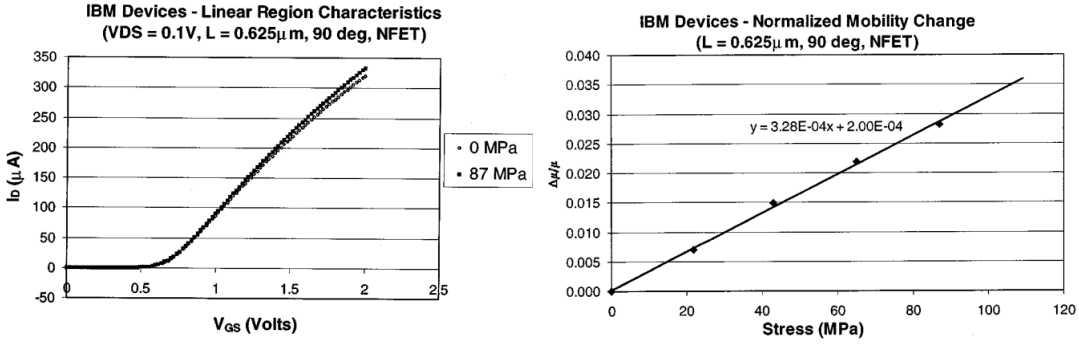


Figure 2.6: (left) Linear region transfer characteristics for two different values of applied uniaxial stress (right) normalized mobility change versus applied uniaxial stress [46].

According to the results presented above, it has been demonstrated that MOSFET threshold voltage is independent of stress and that mobility changes dominate the piezoresistive response of the FET channel region.

2.3 Monolithic Active Pixel Sensors (MAPS)

Monolithic Active Pixel Sensors (MAPS) detectors have been postulated for charged particle tracking in the early 2000s [50] stimulated by a growing interest at that time in developing lepton colliders with an aim on the International Linear Collider (ILC).

During the last decades, MAPS which integrate the sensor and the readout chip in one silicon wafer, have been explored for particle physics experiments with promising results [51, 52].

These have several advantages over traditional hybrid technologies with sensors bonded to readout ASICs. Integrating sensors and front-end electronics on the same die removes the need of interconnections, thus reducing complexity and mass. The close connection of sensor and front-end amplifier reduces input capacitance which reduces the achievable noise floor and thus for the same S/N ratio allows for a reduction in signal and therefore sensor thickness, which also reduces mass. CMOS devices are made in standard commercial technologies with small feature size allowing fine readout pitch. Furthermore designs in standard commercial technologies can be produced quickly and inexpensively, even in large sizes of megapixel arrays, thinned as needed, offer individual pixel readout, radiation hardness, while operating at high speed with low power consumption, from a single and low voltage supply.

Finally, traditional sensors collect signal charge by drift, achieving a tolerance to non-ionizing radiation in excess of 10^{15} 1 MeV n_{eq}/cm^2. In MAPS, signal charge is often collected primarily by diffusion, and therefore may be trapped more easily, resulting in a lower tolerance to displacement damage. MAPS are nowadays a viable solution for HEP experiments, after the successful experience of STAR at RHIC and ALICE at the LHC, in particular when the radiation requirements are modest[1].

2.3.1 MAPS process and principle operation

Monolithic Active Pixel Sensors are typically fabricated using a standard CMOS process version and they consist of three layers. A highly p-doped (p++) substrate which acts as mechanical support, with a very low resistivity[2] , typically in the range of a few tens of $m\Omega cm$, over which a p-doped (p-) epitaxial layer is

[1]Since the Inner Layers of the ITS3 will be closer to the interaction point than in the case of the ITS2, the flux of particles crossing the innermost layer, and correspondingly the radiation load, will increase. The innermost layer will see a particle flux of about 2.2 $MHz\ cm^{-2}$ for a Pb–Pb interaction rate of 50 kHz, which is well within the rate capabilities of the ALPIDE sensor. Here, it should be noted that ALPIDE has a configurable integration time with a minimum value of 1 μs and thus would have sufficient margin to also operate at interaction rates well beyond 100 kHz. The radiation load will also increase by 60–70 %, but still remains well below 1×10^{13} 1 MeV $n_{eq} cm^{-2}$ (NIEL) and 10 kGy (TID), values for which the ALPIDE sensor was verified to fully preserve its performance [3].

[2]Silicon wafers manufacturers usually specify the resistivity ρ of the material, which is related to the doping concentration N by $\rho = 1/N\mu q$, where μ is the mobility of the respective carrier. High resistivity silicon is defined as monocrystalline silicon having a bulk resistivity larger than 1 $k\Omega cm$.

grown. This layer, whose thickness is typically up to 20 μm with a resistivity of the order of 10 Ωcm , represents the sensing volume. The electronics are built in the last micron or so of this layer, with nMOS (pMOS) transistors occupying heavily doped p-wells (n-wells). As a detecting element, the most commonly used structure is the one formed by an N-doped well created in the epitaxial layer, for example the n-well diode.

In the MAPS implementation, this epitaxial layer is used as a detector radiation sensitive volume, with a diode n-well/p-epi working as a charge collecting element. The detector is only partially depleted in the vicinity of the n-well/p-epi junction, so the charge is collected mainly through a thermal diffusion mechanism. However, due to the particular doping profile (p++substrate/p-epi/p+well), there is a potential minimum in the middle of the epitaxial layer, limiting the volume spread of diffusing electrons created by the absorption of radiation. The electrons generated are deflected by the substrate due to a potential barrier formed between the lightly doped p-type epitaxial layer and the heavily doped p-type substrate. The substrate - epi creates a potential barrier preventing the signal charge from diffusing into the substrate.

Similarly, a potential barrier exists between the lightly doped epitaxial layer and the heavily doped p-wells of the nMOS transistors. This results in the containment of the majority of the electrons within the epitaxial layer, which diffuse randomly in the epitaxial layer and are guided towards the n-well of the collection diode. These electrons may move only along the plane parallel to the surface and are rapidly collected when passing close to the collecting pn diode junction, with a typical collection time of the order of 100 ns. The detector active volume is limited in depth to the epitaxial layer only, because of the small lifetime of charge carriers inside a p++ substrate. Electrons lifetime depends on doping concentration and is ~ 50 ns[3] and > 100 μs[4] in the substrate and epitaxial layer respectively. Therefore, the total amount of available charge created by an impinging minimum ionizing particle amounts to a few hundreds electrons only, for a typical epitaxial thickness of the order of 10 μm.

As far as the charge collection mechanism in MAPS is concerned, the presence of the n-well of the pMOS transistor in a standard CMOS process poses a significant problem. This n-well that is at a positive potential, could also collect the diffusing electrons and thus compete with the collection diode. This would result in the reduction of signal charge collected and hence contribute to a reduction of the charge collection efficiency. One way of solving competition between the charge collection between n-wells is to use a deep p-well under the n-well where the pMOS transistors are placed (Figure 2.7). Charged particles can traverse the metal layers and any other material, generating a thin trail of electron-hole pairs in the silicon. Provided that there is only one pn junction in the pixel, the entire amount of radiation-generated electrons will be collected, thus making the sensor able to detect particles regardless of where they hit the sensor [54].

[3]Given the doping concentration of $\sim 10^{18} cm^{-3}$ [53]
[4]Given the doping concentration of $\sim 10^{13} cm^{-3}$ [53]

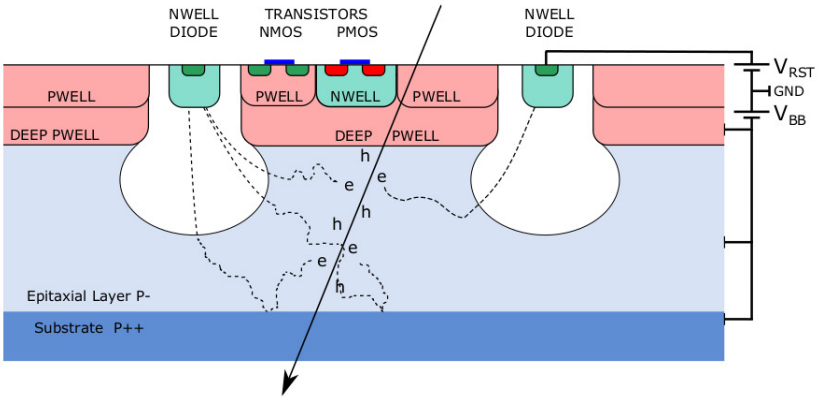


Figure 2.7: MAPS cross section [2]. The deep p-well underneath the n-well is a feature of TowerJazz 180 nm CMOS process.

2.3.2 Charge Generation and Collection in MAPS

A part of the energy deposited by particles crossing silicon is used for the generation of free electron-hole pairs. The remaining energy goes into the excitation of phonons, which dissipate thermally. In contrast to hybrid pixel sensors, which feature an active layer of about 200 μm to 300 μm thick, the epitaxial layer used in the MAPS for the upgraded ITS is between 18 μm and 30 μm. This leads to a comparably small average charge generated by a traversing MIP of 1000 e- at an epitaxial-layer thickness of 18 μm. Due to the fluctuations, a single particle can deposit less energy. The smaller collected charge Q leads to a reduction in terms of signal voltage $\Delta V_{IN} = Q/C$ for a given pixel capacitance C. In order to achieve a high signal voltage, the combination of charge collected by the central pixel in a cluster[5], called seed, and the capacitance of the pixel C needs to be optimised in order to increase the Q/C ratio. A high Q/C ratio leads to an improved signal-to-noise ratio and as a consequence reduced power consumption of the circuitry [55]. The signal induced on the electrode can be described using the Ramo-Shockley theorem [56, 38]. The small diameter between 2 μm and 3 μm of the collection electrode compared to the epitaxial-layer thickness leads to a weighting field only non-negligible in the close vicinity of the collection electrode [57]. Consequently, only for charge travelling close to the collection electrode a signal is induced. After the arrival of the charge at the electrode, the integral of the induced signal corresponds to the collected charge. The small size of the collection n-well (2 μm - 3 μm) compared to the pixel pitch of the order of 30 μm leads to a small capacitance of the electrode of a few fF [58].

As shown in figure 2.7, in MAPS the charge is collected by two mechanisms.

[5]If the charge of a traversing particle is shared by several pixels, all pixels collecting charge form a cluster. The pixel collecting the majority of the charge is called the seed pixel.

In the vicinity of the collection n-well, charge is collected by drift in the electric field of the depletion zone of the pn-junction formed by the collection n-well and epitaxial layer. In the undepleted part of the epitaxial layer, the charge is collected by diffusion. In the first MAPS for the ITS upgrade, full depletion of the epitaxial layer was not reached. Potential barriers at the borders between the deep p-well ($N_A \approx 10^{18}\ cm^{-3}$) and the epitaxial layer ($N_A \approx 10^{13}\ cm^{-3}$) as well as between the substrate ($N_A \approx 10^{16}\ cm^{-3}$) and the epitaxial layer keep the charge contained in the epitaxial layer. The potential V_{bi} at the borders can be calculated using:

$$V_{bi} = \frac{k_B T}{q} \ln\left(\frac{N_{A,p^+}}{N_{A,p^-}}\right) \tag{2.7}$$

where N_{A,p^+} and N_{A,p^-} are the doping concentrations of p-wells or substrate and the epitaxial layer, respectively [59, 60]. The factor $k_B T/q$ is the thermal potential which is about 25 mV at about 300 K. The resulting potential barriers are about 7 and 12 times the thermal potential at the border of the substrate and the p-wells, respectively. The charge-collection efficiency depends on the charge-carrier lifetime and thus indirectly on the charge-collection time. The charge-carrier lifetime is mainly determined by impurities causing energy levels within the band gap [59]. Additionally, charge is also produced in the substrate which, however, depending on the lifetime and the path length in the substrate stays mainly uncollected [61].

A possibility to increase the amount of collected charge is increasing the epitaxial layer thickness leading to more initial charge generation. But a thicker epitaxial layer and a constant depletion volume, lead to an increased diffusion volume. The increased diffusion leads in turn to more pixels sharing the generated charge and therefore less relative signal in the seed, reducing the charge collection efficiency and the radiation hardness. To further increase the depletion the doping concentration of the epitaxial layer should be kept as low as possible.

An approach to increase the signal voltage at the seed pixel is to reduce the charge sharing and concentrating it in fewer pixels. This can be achieved by enlarging the depletion volume. The depletion volume can be increased by increasing the reverse bias voltage V_{RB} between the collection electrode and the substrate as the depletion volume monotonically increases as function of V_{RB}. The total reverse bias voltage is $V_{RB} \approx V_{RST} - V_{BB}$, measuring both voltages relative to the chip ground potential. In the MAPS for the upgraded ITS, the reset voltage[6] V_{RST} cannot exceed the supply voltage of 1.8 V. The reverse-substrate bias V_{BB} is applied via a p-guard ring surrounding the pixel matrix to the epitaxial layer and the substrate. The input capacitance C is the sum of the circuit capacitance and the junction capacitance. The circuit

[6] After the detection of a particle event, the sensor may need to be reset or cleared to prepare for the detection of the next event. This resetting process often involves applying a reset voltage to bring the sensor back to its initial state or baseline.

capacitance consists of the routing line capacitance of the connection from the electrode to the input transistor, the input capacitance of the input transistor and the capacitance of the reset mechanism. The junction capacitance is built up between the n-well and the undepleted epitaxial-layer volume as well as the surrounding p-well. As the capacitance is inversely proportional to their distance, the junction capacitance can be decreased by increasing the depletion volume and increasing the spacing between the n-well diode and the surrounding p-well. The maximum spacing is limited mainly by the surface needed by in-pixel and general readout circuitry.

2.3.3 Radiation effects

Radiation induced effects in MAPS are usually subdivided into bulk and surface damage. Bulk defects are caused by the displacement of lattice atoms due to non-ionizing energy loss, while the latter include the effects of ionizing radiation on the dielectrics and the silicon-dielectric interface at the sensor surface.

Bulk damage

Bulk damage is caused by the inelastic displacement of nuclei in the sensor, producing imperfections in the crystal structure that introduce additional energy levels within the silicon bandgap that may change the properties of the material. This type of radiation damage is related to hadrons. To allow a comparison of the damage caused by different particle types with different energies, bulk radiation damage is scaled using the non-ionizing energy loss (NIEL)[7].

The radiation induced lattice defects lead to an increase of the volume-related leakage current[8] across a p-n junction that is strictly proportional to the equivalent fluence Φ_{eq}

$$\Delta I = \alpha \Phi_{eq} V \tag{2.8}$$

where V is the depleted volume and α the current-related damage rate[9] [57]. An increase of the leakage current leads to an increase of the shot noise, as well as a larger power dissipation of the sensor.

If the impurity atoms used as doping are involved in the radiation induced crystal defects, it may happen that they become electrically inactive, hence

[7] Neutrons of 1 MeV are commonly used as reference particles. The damage due to a fluence Φ_{phys} of another particle has therefore to be related to the equivalent fluence Φ_{eq} of 1 MeV neutrons causing the same damage. The energy dependent conversion factor is called hardness factor [57].

[8] Leakage current: if free charge carriers, as generated by ionizing radiation, enter the depletion region of a reversely biased diode, they will be collected by the respective electrodes by means of drift in the electric field. However, also in absence of radiation there always is a steady current referred to as leakage current. It stems from various components, one being the diffusion of free charge carriers from the undepleted sensor volume into the depletion region. The more important contribution, however, is the thermal generation of charge carriers at generation-recombination centers in the depleted volume.

[9] For silicon at 20^oC α is about 3.9×10^{-17} A/cm.

losing their function as donors or acceptors. At the same time, new acceptor or donor states may be created during the radiation induced generation of defects. Crystal defects introduce states that act as trapping centers for charge carriers. As a carrier moves through the sensor material, the probability of encountering a trap is proportional to the elapsed time. A parameter to describe trapping is therefore the carrier lifetime τ, defined as the average time it takes for a minority carrier to recombine.

Surface defects

Ionizing radiation is scaled in with the total ionizing dose (TID). It is an accumulating effect which becomes worse the longer a device is exposed to ionizing radiation. It essentially affects the surface oxide layers, as well as the lateral isolation oxides of MOSFET transistors. The critical material is silicon dioxide (SiO_2) which is used for the gate oxide, for device isolation, and for separating the routing metals. In the oxide, charged particles and sufficiently energetic photons produce electron-hole pairs. Electrons and holes in SiO_2 have very different mobility; hence if an electric field is applied, the two are quickly separated. In absence of electrons, the holes can not recombine and are trapped in the oxide, where they can originate two types of defects. In the oxide itself, the trapped holes generate positive charges. At the interface between the oxide and the crystalline silicon they can generate the so-called interface states. The build-up of interface states is slower than the direct hole trapping; therefore the device characteristics still evolve after the irradiation has stopped. Interface states are amphoteric, so they can behave as donors or acceptors. This will shift the threshold voltage of CMOS devices and can give rise to leakage currents between drain and source of NMOS transistors and in between neighbouring transistors. A reduction of the TID sensitivity has been observed when moving to smaller technology nodes. This improvement is mostly attributed to the thinner gate oxides. [62, 63, 64].

Single Event Effects

Ionizing radiation in digital structures can induce Single Event Effects (SEE) as the result of an instantaneous impact of radiation affecting the state of the electronics, and can occur either as a Single Event Transient (SET) or a single-event upset (SEU). The former causes a transient change of voltage in one of the capacitive nodes of a logic gate or a memory cell. The likelihood of a SET decreases with increasing node capacitance. If this change is captured by a memory device, it becomes a persistent effect. In general, this is a threshold phenomenon that depends on the Linear Energy Transfer (LET) of the impinging particle, because the deposited charge must be sufficient to change the status of a circuit node. If a bit-flip happens in a state register of a Finite State Machine (FSM) or in a configuration register, a full system reset or reconfiguration may be needed to restore the system into a properly functioning state [65, 66]. The ionization can also cause single event latchup (SEL), a type of short circuit

that triggers parasitic structures which can disrupt proper functioning of the element, or possibly can even lead to its destruction. During a SEL, the device current exceeds the maximum specified for the device. Unless power is removed, the device will eventually be destroyed by thermal effect.

Chapter 3

The ALICE PIxel DEtector of the ITS

As a high performing MAPS, ALPIDE is the advanced sensor developed for the particle tracking of ALICE experiment at CERN [67].

The ALPIDE development is an alternative to the traditional rolling-shutter[1] readout architecture adopted by the Mistral, Astral [69] and Cherwell [70] prototypes that have been designed in the framework of the ITS Upgrade. It was pursued aiming at reducing the power consumption and integration time by an order of magnitude below the specifications for the ITS after the LS2.

The ALPIDE pixel chip is manufactured using the commercial 180 nm CMOS imaging sensor process by TowerJazz [71]. The sensor includes the option to implant an n-well on top of a deep p-well (see figure 2.7), separating the active volume, i.e. the epitaxial layer, from all n-wells except for the collection diodes. This makes possible to implement PMOS transistors inside the pixels, without deteriorating the charge collection efficiency. The chips are produced on wafers with a high resistivity ($> 1k\Omega\, cm$) epitaxial layer of 25 μm thickness on a p-type substrate. This allows for a bigger depletion area in the active layer and improves the signal to noise ratio of the sensor as well as its resistance to non-ionizing radiation.

In the next sections, the ALPIDE design and its operational features will be introduced. In order to verify if the electronic circuitry on the ALPIDE preserves its functionalities after the mechanical stress of bending it, extensive test campaigns dedicated to the electrical characterization of bent single ALPIDE chips and the evaluation of their performance in different test beam environment, were carried out and the results will be presented.

[1] Rolling shutter readout is the standard for MAPS based on processes limited to NMOS transistors inside the pixels [68].

3.1 ALPIDE Pixel Matrix Architecture

The ALPIDE chip (Figure 3.1) contains approximately 5×10^5 pixels, each measuring 29.24 $\times$ 26.88 μm^2, arranged in 512 rows and 1024 columns, for a total area of 30 x 15 mm^2 and thickness of 50 μm. A periphery circuit region of 1.2 $\times$ 30 mm^2 including the readout and control functionalities is present.

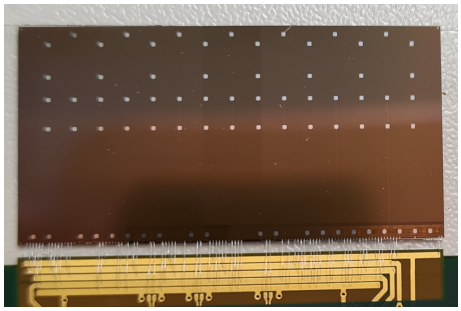


Figure 3.1: Picture of an ALPIDE chip wire bonded to a flexible printed circuit. The die size is 30 $\times$ 15 mm^2 and the total thickness is 50 μm.

Each pixel cell contains a sensing diode, a front-end amplifying and shaping stage, a discriminator and a digital section. In every pixel there is a pulse injection capacitor for injection of test charge in the input of the front-end. The readout of pixel hit data from the matrix is based on a circuit named Priority Encoder. There are 512 instances of this circuit, one every two single column, which provides to the periphery the address of the first pixel with a hit in its double column [72] (Figure 3.2).

3.2 Principle of operation of in-pixel circuitry

The ALPIDE in-pixel front-end circuitry consists of an analogue part and a digital part [58]. The analogue part mainly shapes and discriminate the signal and digital part contains the hit register and further logic to test and mask (disable) the pixel. A simplified layout of the in-pixel circuitry and the signal-flow inside the pixel circuitry of ALPIDE are shown in figure 3.3.

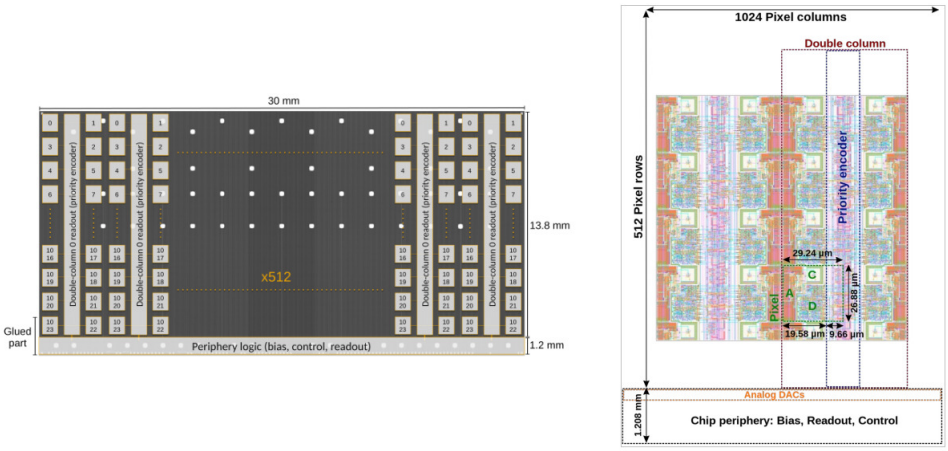


Figure 3.2: (left) Layout of the ALPIDE pixel matrix. The pixels are organised in double-columns, each featuring a priority encoder circuit which propagates the addresses of the hit pixels to the periphery logic. The aluminum pads providing the electrical interface to the chip are located on the top of the periphery logic. (right) Schematic illustration of the ALPIDE chip architecture on top of a closeup showing the pixel matrix circuitry. The single pixel layout is highlighted in green with the locations of the collection diode (C), the analogue front-end (A) and the digital pixel section (D). At the bottom there is the chip periphery containing the analogue DACs and a mostly digital section.

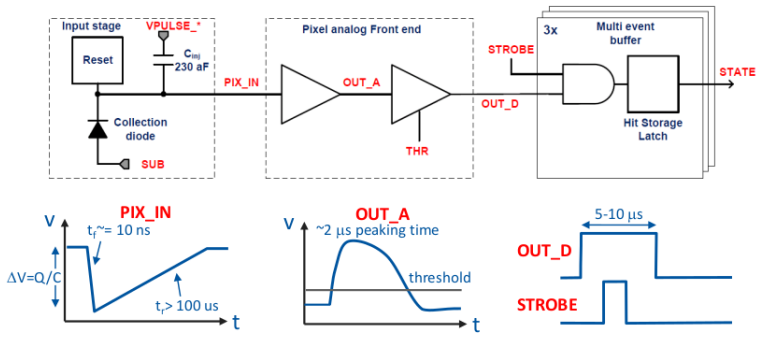


Figure 3.3: In-pixel hit discrimination block diagram and signal flow [33].

The input stage contains a collection diode, a continuous reset of the input node and a pulse injection capacitance (C_{inj}) used to inject test charges into the analogue front-end. When a charge is collected, there is a fast potential drop at the input node of the amplifying circuit. The reset mechanism slowly restores the potential to its nominal value. At the amplifier input, the voltage drop is shaped to a signal with a duration up to 10 μs and peaking time of $\sim$ 2 μs.

The peaking time and pulse length are chosen to maintain an optimized power consumption and the required physics performance [2]. Several bias voltages and currents are adjustable via on-chip 8-bit digital-to-analogue converters (DAC). As long as the amplitude of the signal is over the threshold, the discriminator output stays asserted. If the strobe signal, given by the trigger (common to all pixels), is asserted at the same time, the discriminator output state is stored into a register. Each pixel contains three registers as multi-event buffer memory. By the hit-driven architecture, the data is read out only when there are stored hits.

3.2.1 Pixel Analogue Front-End

A scheme of the pixel front-end circuit is shown in figure 3.4.

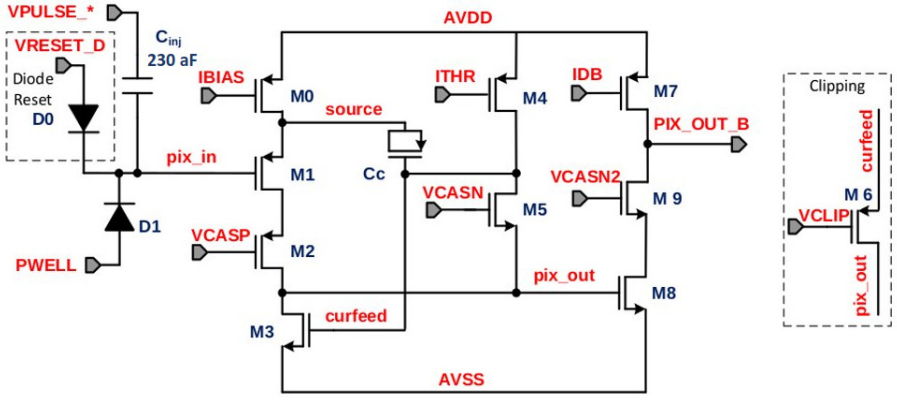


Figure 3.4: Pixel analogue front-end schematic for ALPIDE [72].

Diode D1 is the sensor p-n junction. The input node is continuously reset by diode D0. VRESET_D establishes the reset voltage of the charge collecting node (pix_in). A particle hit will lower the potential at the pixel input pix_in by a few tens of mV. This will cause the source follower formed by M1 and the current source M0 to force the source node to follow this voltage excursion and dump charge associated with the voltage change and the capacitance at the source node onto the analogue output node pix_out. In addition the coupling of the source node to the curfeed node by Cc will cause the current in M3 to be reduced. Both effects add and cause the potential on the output node pix_out to swing upwards by several hundreds of mV. This forces M8 into conduction and if the charge deposit from the particle hit is sufficiently large to overcome the current setting IDB on M7, M8 will drive the PIX_OUT_B node to zero. The baseline value for pix_out and the reset to baseline after particle hit is defined by voltage bias VCASN and current bias ITHR. The gate of M8 transistor is controlled by pix_out and baseline is defined in a way that $I_{M8} < IDB$. The charge threshold is defined by the distance between baseline voltage of pix_out

and the point where $I_{M8} = IDB$. Therefore, this parameter is controlled by ITHR, VCASN and IDB: the effective charge threshold is increased by increasing ITHR or IDB and it is decreased by augmenting VCASN. When pix_out has a higher voltage level than the threshold point $I_{M8} > IDB$, PIX_OUT_B receives a discriminated active low output. Voltage bias VCLIP controls the gate of the clipping transistor M6, which limits the pulse duration in case of very large collected charges. The active low PIX_OUT_B signal is applied to the digital section of the pixel where it is used to set the hit status register. It is possible to electrically inject a test charge into the input node. This is achieved by applying a voltage pulse of controllable amplitude to the VPULSE pin of the C_{inj} capacitor. The design value of C_{inj} is 230 aF. The amplitude of the applied voltage pulse is defined by the difference between VPLSE_HIGH and VPLSE_LOW, both set in the DAC unit. The pulse generation is controlled by the digital section of the pixel and the periphery.

3.2.2 Pixel digital Front-End

The digital section of the pixel is shown in figure 3.5. The pixel contains three State Registers. Each State Register is a Set-Reset Latch that keeps the hit information. The State Register is normally set by the front-end discriminated output PIX_OUT_B if the corresponding STROBE_B signal is asserted simultaneously. It can also be set programmatically by the DPULSE signal (digital pulse functionality) in order to test the digital in-pixel circuitry. The registers can be reset either by a PIX_RESET pulse generated by the Priority Encoder during the readout, either by a global FLUSH_B signal. The selected State Register output bit can be masked and the result is the output to the Priority Encoder (STATE signal). The logic provides two programmable functions: masking and pulsing. When control bit MASK_EN is set high, the STATE output is forced to 0, effectively masking the pixel output to the Priority Encoder.

3.3 Address-Encoder and Reset-Decoder Circuit

In order to achieve a power density and an integration time one order of magnitude below the minimum specifications for the ALICE ITS2, the ALPIDE data-driven readout architecture has been developed with zero-suppression in the matrix. The architecture contains an Address-Encoder and Reset-Decoder (AERD) circuit based on an arbitration tree with the priority logic, only reading hit pixels. The zero-suppression technique allows a sufficient reduction of the readout time and the power consumption.

The front-end works like an analogue memory: it will generate a pulse with a shaping time of about 4 μs. If a strobe is applied during this time after a hit arrives at the pixel, this hit data will be latched in the in-pixel state register as a STATE signal. The AERD circuit which reads and resets the hit pixels is arranged in double columns inside the matrix. The I/O signals of the AERD block are described in figure 3.6. The design takes advantage of a gating

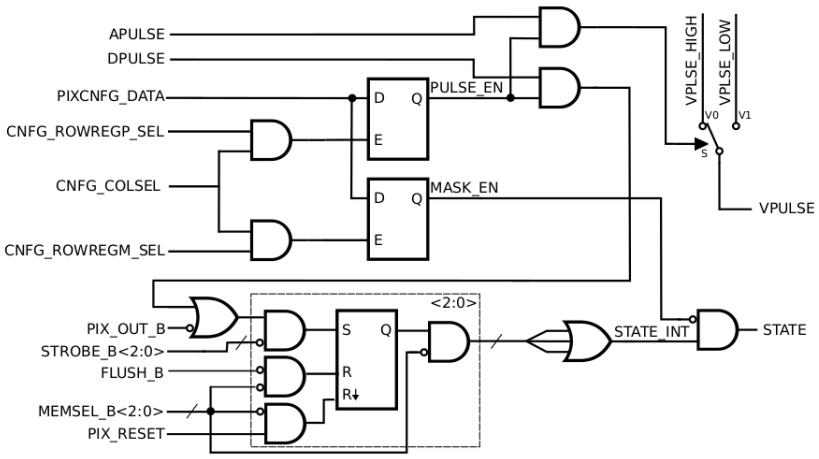


Figure 3.5: Diagram of the ALPIDE in-pixel digital front-end [72].

technique to reduce power: SYNC is a gated clock signal propagated from the chip periphery, used to select the highest priority pixel to be read and reset. A VALID signal is a flag which is activated when there is at least one hit in a pixel.

The AERD readout circuit has been implemented based on an arbiter tree scheme with hierarchical address encoders and reset decoders. The block diagram of the tree structure to decode 16 pixels, as an example, is shown in figure 3.7. The VALID signal propagates from the lower hierarchical level of the arbiter tree to the top. If pixel 4 is hit, through the FAST OR chain, VALID will be active. During the readout phase, while the VALID signal is active, a synchronous signal (SYNC) propagates back into the hit pixel to read its address. Combined with priority logic, at the lowest hierarchical level of the tree, the SYNC signal resets only the pixel with the highest priority during the same clock cycle after the address of that pixel has been read. During the propagation of the SYNC signal, also the ADDRESS[3:0] of the pixel being reset propagates down to the End of Column (EoC). The ADDRESS lines are managed by tri-state buffers, which are enabled when SYNC is high, and in high impedance when SYNC is low. Therefore, to ensure that the value of ADDRESS is available at the EoC, half a SYNC cycle is set aside for the propagation of the ADDRESS signals. On the falling edge of the SYNC, the state register of the read pixel is reset, and when SYNC is low, the new configuration of VALID and internal signals propagate such that the next pixel is going to be read in the subsequent SYNC cycle.

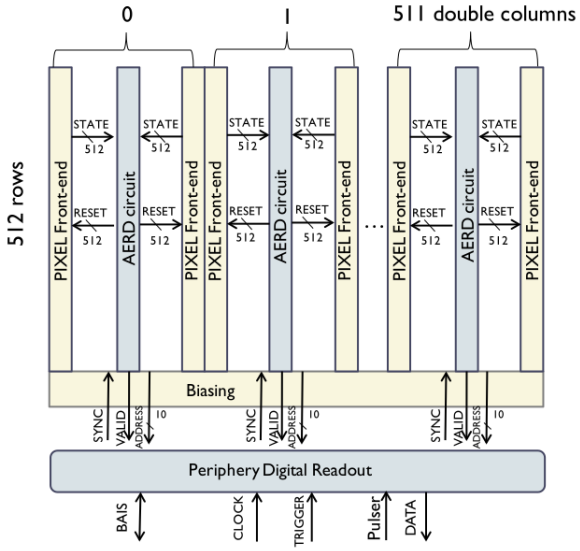


Figure 3.6: Readout structure of the ALPIDE chip [73].

3.4 Electrical characterization of an ALPIDE chip in laboratory

An extensive test campaign, dedicated to the electrical characterization of a bent single ALPIDE chip, has been realized. The results of the measurement and evaluation of the main parameters steering the chip functioning, are presented and discussed in this section. The setup arranged in the laboratory was composed by a single ALPIDE chip wire bonded to a flexible printed circuit (FPC). This component is extended by coupling it to an I-board, used both as noise filter and to release mechanical stress from the setup. A decoupling of 10 μF was set on its analogue and digital lines. The I-board is connected through a cable to an adapter board (Flex2daq) for the communication between the chip and the DAQ board (Figure 3.8). At this stage, decoupling capacitors of 100 μF were also used for the analogue and digital current channels.

Concerning the DAQ board (rev3b, [74]), it is powered with a voltage of 5 V. It contains potentiometers to vary the analogue and digital power (AVDD[2] and DVDD[3], respectively) and their nominal value is 1.8 V. Other present elements are the adapter for the Flex2daq board and the LEMO connector for the application of a back bias to the chip substrate. For this study, the tests

[2]Supply net of the analogue domain. This includes the pixel front-end circuits, the analogue biasing circuits (DACs), the ADC block.

[3]Supply net of the digital domain. This includes the in-pixel configuration registers, the matrix readout circuits, the peripheral readout circuits and the chip input and output buffers and transceivers.

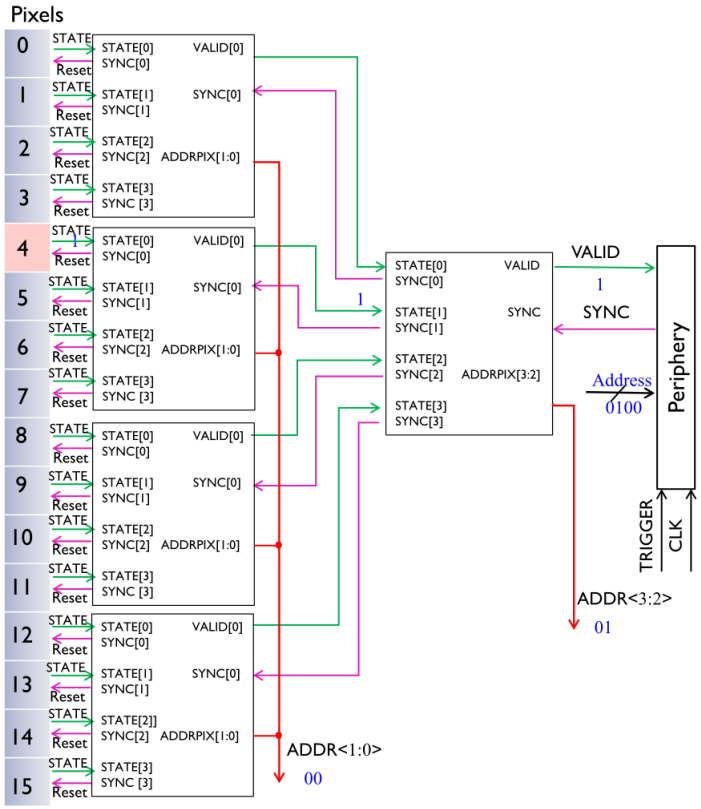


Figure 3.7: Two hierarchical levels of AERD to decode 16 pixels [73].

were done at 0 Vbb, so a 0 Ω terminator was coupled to the back-bias line to prevent it from having a floating ground, which could cause detrimental effects on the circuit performance and operating capabilities or increase the risk of damaging the chip. The whole setup was placed inside a black box in order to isolate it from any light source.

The tests have been performed with the ALPIDE both in flat and bent positions. For the bending procedure a cylindrical jig made of Aluminum was designed and realized to assure a good control of the shape and it is attached to the platform holding the setup (Figure 3.9 (left)). The radius of the jig is $r = 18\,mm$, this corresponds to the intended radius for the innermost layer of the ITS3. The chip is positioned on top of this tool, such that the bent axis is parallel to its long side and is bent carefully forcing it to follow the curved shape by using a piece of kapton foil. Starting at one of the short sides and then moving on towards the other one in a rolling fashion, the chip bends gradually and is fixed in the final position by screwing the kapton to the jig at both extremes (Figure 3.9 (right)).

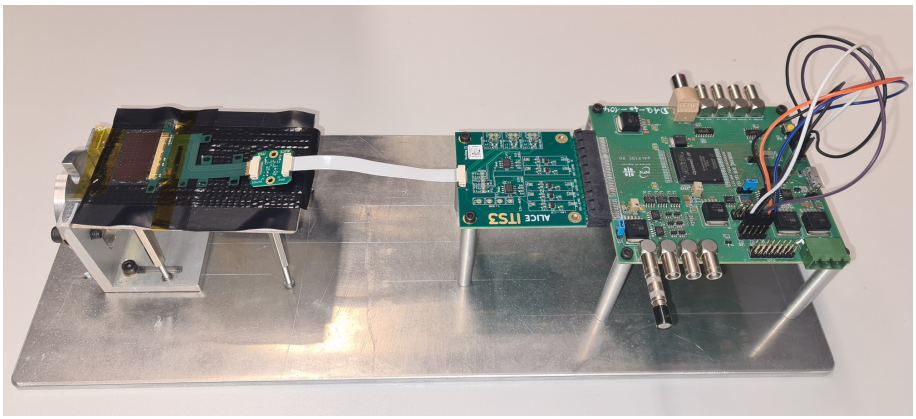


Figure 3.8: Setup assembled for the electrical tests with flat ALPIDE.

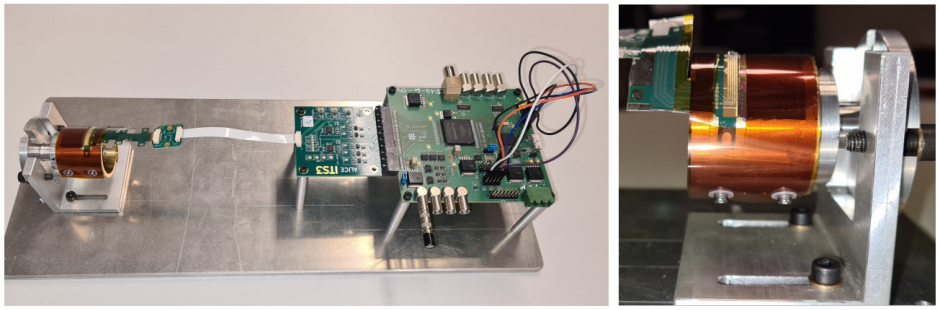


Figure 3.9: (left) Setup assembled for the electrical tests with the ALPIDE bent over a cylindrical jig. (right) Closed view of the ALPIDE bent along its long side over the cylindrical jig and held in position with a kapton foil screwed at the edges.

3.4.1 Threshold and Noise

The key parameter leading the chip performance is the charge threshold. It is defined for every pixel by three front-end parameters: ITHR, VCASN and IDB. The effective value is increased by increasing ITHR or IDB and decreased by augmenting VCASN. For a given set of these parameters, pixel charge threshold can be determined using the analogue pulsing. The measurement consists in injecting N_{inj} times a range of test charges Q_{inj} via pulsing capacitance C_{inj} such that:

$$Q_{inj} = C_{inj}(VPLSE\_HIGH - VPLSE\_LOW) \tag{3.1}$$

$VPLSE\_HIGH$ and $VPLSE\_LOW$ maximum value is 1.8 V and are set via on-board 8-bit DACs. The minimum voltage step equivalent to 1 DAC unit

is 7 mV. Consequently, the minimum charge step is 10 e^-, considering the nominal value of the pulsing capacitance (230 aF). For each test charge Q^i_{inj}, out of N_{inj} injections, N^i_{hit} pass the threshold, i.e. result in pixel hits. The plot of N^i_{hit}/N_{inj} as a function of Q^i_{inj} (Figure 3.10) is often referred to as s-curve measurement as its shape resembles the letter S. The threshold is defined as charge at which pixel fires in 50% of cases (injections). Assuming a Gaussian distribution of the noise, the response function of the pixel is fitted with an error function:

$$f(x) = \frac{1}{2}\Big[1 + \ erf\Big(\frac{x-\mu}{\sqrt{2}\sigma}\Big)\Big] \tag{3.2}$$

where μ is the charge threshold and σ is the temporal noise of a pixel. By fitting the noise component with the error function, one can estimate the spread of the noise distribution, typically represented by the standard deviation. This allows the identification of systematic effects that might be present in the measurements, such as electronic noise, variations in sensor response, or non-uniformities.

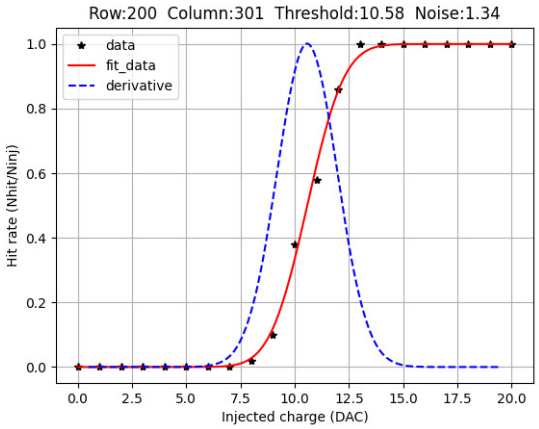


Figure 3.10: S-curve measurement. N_{hit}/N_{inj} as a function of Q^i_{inj} for an ALPIDE pixel. The red line is the error function fit and the blue line is its derivative representing electronic noise of the pixel.

S-curve measurement is performed on all pixels in a sensor; the distribution of thresholds and noise are shown in figure 3.11. The tail in the threshold distribution indicates a sub-array of pixels with different behavior. As observed in the map of figure 3.12, there is spatial dependence with respect to the sensor area, given that in this case the measurement was done before the realization of an equalization effort for this specific sensor. Finally, the average of the threshold distribution is defined as the chip threshold. Additionally, the RMS of the threshold contributes to the Fixed-Pattern Noise (FPN). This is assumed

to originate from mismatches in the in-pixel circuitry, leading to variations in the pixel matrix with a stable pattern.

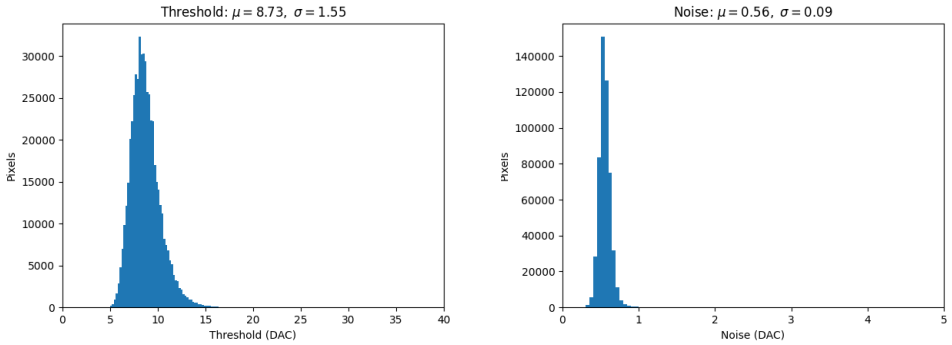


Figure 3.11: Threshold and noise distributions of the pixel matrix.

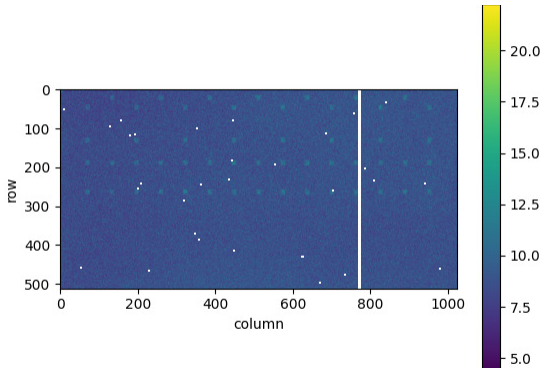


Figure 3.12: Spatial distribution of thresholds over the pixel matrix. The empty points correspond to noisy and/or nonfunctional pixels, which are masked in the analysis stage.

3.4.2 Fake-hit Rate

The fake-hit rate of a sensor is defined as the number of hits per pixel per event in absence of any external stimulus, i.e the rate at which pixels show a hit when there are no ionizing particles. For most of the chips, there are few pixels that do not exhibit a response function as the one shown in figure 3.10. These are the pixels that dominate the fake hit rate if not masked. Due to its strong effect on the performance of tracking device this quantity is considered a very important parameter in pixel chip performance. There are two sources of fake-hits: thermal/shot noise and random telegraph noise (RTN), the latter manifests itself as discrete changes of the pixel output between two or more levels: the

amplitudes are typically well defined but the period is random and may reach from μs up to minutes [75]. RTN depends on various parameters, including the transistor size and type, the applied gate voltage, the oxide thickness and the temperature [58, 76, 77].

The fake-hit rate is measured by sending a number of triggers (N_{trg}) to the chip without providing any charge injection, and recording the number of hits N_{hit}:

$$FHR = \frac{N_{hits}}{N_{pix} \cdot N_{trig}} \tag{3.3}$$

where N_{pix} is the number of pixels in the sensor.

3.4.3 Test results

As listed before, the ALPIDE chip has been characterized mainly in terms of threshold and fake-hit rate. As a first step, the evolution of both variables have been measured with respect to time. For this task, the scan of each parameter is repeated 50 times, running the data taking consecutively under the same conditions, i.e. maintaining invariant the in-pixel front-end DACs (ITHR=51, VCASN=57 DAC). Such repetitions are identified as "runs" in the plots and the total times covered were ∼40 hours in the case of threshold measurements (∼50 min each run) and few minutes for all fake-hit rates.

In the next step the variables were measured over certain ranges of the mentioned DACs. First of all, the ITHR value was varied between 40-160 DAC in steps of 10 DAC, maintaining VCASN=57 DAC. Later, VCASN was changed in the range 45-75 DAC and ITHR remained fixed and equal to 51 DAC. All the tests above were performed and their results were compared for three positions of the chip in the following order:

1. Original flat configuration (flat1).
2. Bent at a radius of r=18 mm, following the procedure described previously (bent1).
3. Flat position, turned back by manual unbending (flat2).

Evolution over time

The comparison of the charge threshold trends over time for all configurations is shown in figure 3.13. In the case of flat1 (black points), the behavior of the threshold values presents some statistical fluctuations, indicating the variation of the noise in the pixel matrix. This could be related to external conditions like the temperature since it has been demonstrated experimentally that exists a correlation between temperature and threshold [78] and our setup was not placed in a temperature-controlled environment. The points vary around 8.7 DAC ($87e^-$) and the difference between the maximum and minimum values in the curve is 0.13 DAC ($1e^-$). The trend, once the chip is bent (red curve), behaves

more stable and is shifted to larger thresholds. This represents a clear effect of the curvature on the integrated (in-pixel) electrical circuitry, which is now in a more decompressed state, potentially leading to differences in the electrical properties (section 2.2.2). In this particular case, where the chip is bent along its larger side and faced up, as reported in [79], the analogue current, measured during chip operation, is larger with respect to flat position. The points in this curve have the mean close to 9.2 DAC ($92e^-$) and the maximum difference of thresholds is 0.13 DAC, equivalent to $1e^-$. The blue curve, corresponding to the last flat configuration, is shifted to lower thresholds with respect to bent case, but the values are still higher and more stable than those in flat1. The trend fluctuates around 8.5 DAC ($85e^-$) and the maximum difference of thresholds is 0.05 DAC. The returning of the sensor to the original behavior would depend on the environment conditions like temperature as mentioned above or on the fact that the stress applied for bending did not cause permanent deformation to the silicon [80]. The time passed between the measurements for flat1 and flat2 was approximately of three weeks. The standard deviation of the threshold distributions have been averaged for all runs at each position and the values are also reported in figure 3.13. These have a correlation with threshold, i.e. increase as the threshold increases. In the case of the bent chip, the increment could be an effect of the stress applied to the sensor matrix.

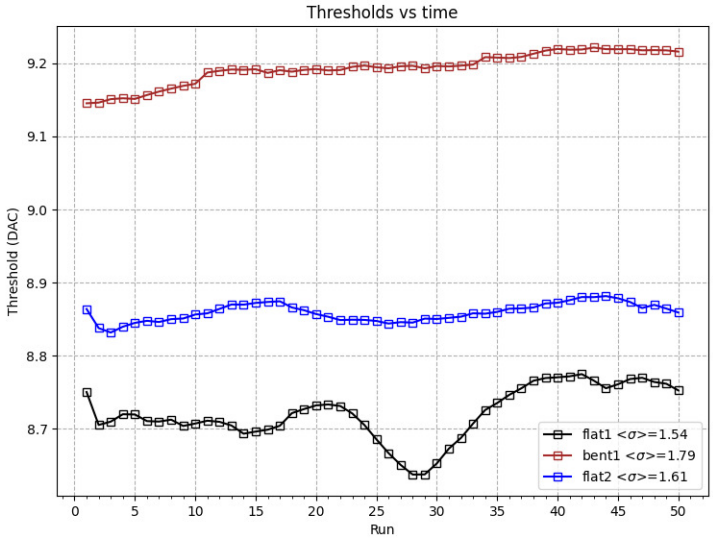


Figure 3.13: Behavior of pixel matrix thresholds as a function of time.

Additionally, the values of current consumed by the analogue (IAA) and digital (IDD) circuitry of the pixel matrix have been measured and can be observed in figure 3.14. The increment of IAA, which is of about $\sim$1.5 mA for

the bent chip compared with the values of flat1 and flat2, is consistent with the results reported previously [79].

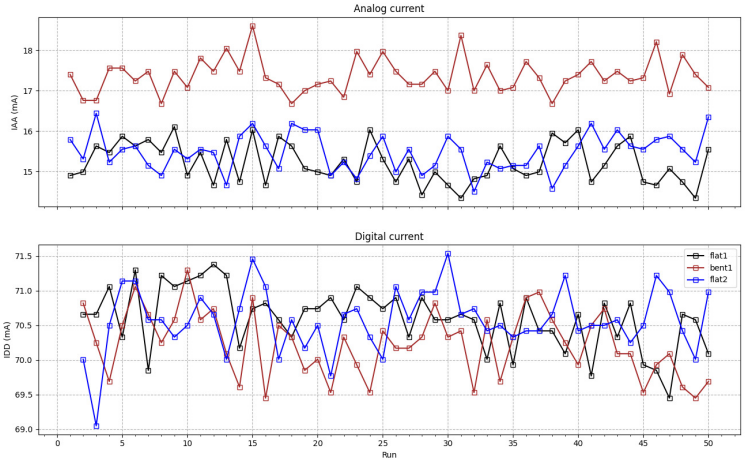


Figure 3.14: Trends of analogue (IAA) and digital (IDD) currents as a function of time.

Figure 3.15 shows the calculation of the noise for every run. As expected, the behavior of noise trends is opposite to thresholds because an increment of the latest contributes to mask part of the signals generated from the different noise sources. It can be noticed that the pixel matrix is less noisy when the chip is bent, product of the higher threshold values and the contrary happens for flat1 and flat2.

The other important parameter evaluated over time is the fake-hit rate, whose evolution is represented in figure 3.16. The red curve, corresponding to bent chip, shows the highest values of noise occupancy (fake-hit rate). Even though this seems to be in contrast with the results obtained in noise measurements, high levels of noise are not necessarily correlated with high noise occupancies. In general the noise occupancy is strongly determined by few noisy pixels with very high fake hit rates (noise much higher than the mean). The measurements were indeed repeated without pre-masking any pixels. However, even without pre-masking noisy pixels, the fake-hit rates are within the same order of magnitude $\sim 10^{-7} pixel^{-1} event^{-1}$ for all positions, which is compatible with the requirements for the ALPIDE chip reported previously [71].

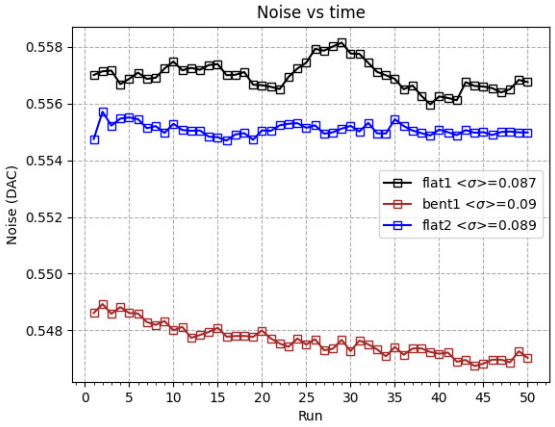


Figure 3.15: Values of the pixel matrix temporal noise as a function of time.

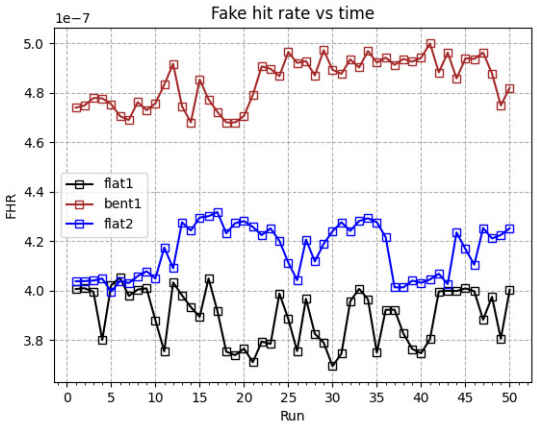


Figure 3.16: Calculated values of noise occupancy (Fake-hit Rate) for the hole pixel matrix as a function of time.

The impact of masking pixels is studied by excluding n pixels with the highest hit count in the fake-hit rate calculation. Masking is justified as long as the fake hits originate from a constant, small group of pixels featuring a significantly higher fake-hit rate than a standard pixel. Figure 3.17 shows the changes on the fake-hit rate with respect to the number of masked pixels (upper axes) featuring fake hits (bottom axis) from 2 up to 25% of the number of triggers sent to the matrix during the measurement ($N_{trig} = 100000$). As can be observed, there were no fired pixels, i.e. pixels with $N_{hit} = N_{trig}$. For the three positions there was just one pixel hitting the 10% of the times the

matrix was read and after masking it the fake-hit rate remained on the order of $10^{-7} pixel^{-1} event^{-1}$. There were a larger number of pixels recording from 2 to 5 hits, specially when the chip was bent. The effect of masking them was a significant reduction of the fake-hit rate down to $\sim 10^{-10} pixel^{-1} event^{-1}$.

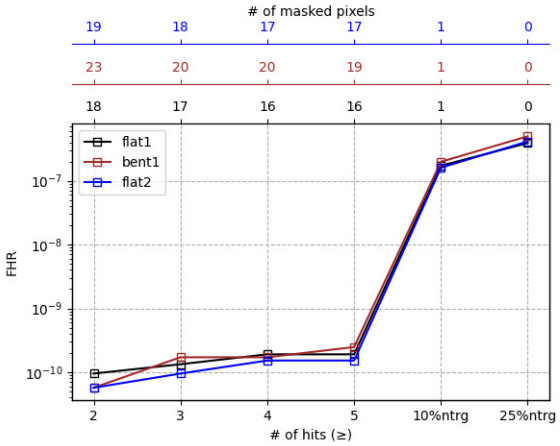


Figure 3.17: Evolution of the Fake-hit Rate values with respect to masked pixels (three upper axes) recording hits from 2 up to 25% the number of triggers sent to the matrix during the readout (bottom axis, $N_{trig} = 100000$).

Influence of ITHR and VCASN

Figure 3.18 exhibits the trends of threshold as a function of ITHR (left) and VCASN (right). The plots show the expected behavior of the pixel matrix for each parameter range. Indeed, as demonstrated in several studies [67, 81], the threshold increases almost linearly with increasing ITHR and it should decrease as VCASN increases. It is important to note that similar thresholds can be obtained with different combinations of ITHR and VCASN[4]. The standard deviations shown in figure 3.18 are represented as a function of ITHR (left) and VCASN (right) in the trends of figure 3.19. These change according to the threshold variation, as demonstrated in [31], which indicates the uniformity of the sensor response.

[4] As mentioned in a previous section, VCASN determines the baseline voltage at the pix_out node within the analogue front-end. If the baseline voltage at the pix_out node, connected to the gate of transistor M8 (Figure 3.4), exceeds a certain critical voltage $V_{c,M8}$ for which the current through M8 surpasses IDB, the PIX_OUT_B node of the pixel is driven continuously low (active). This represents a limitation when trying to decrease the threshold by increasing VCASN.

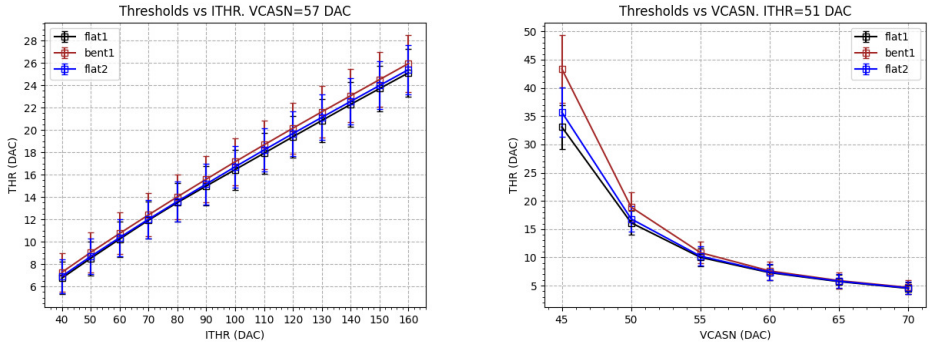


Figure 3.18: Threshold values as a function of ITHR (left) and VCASN (right).

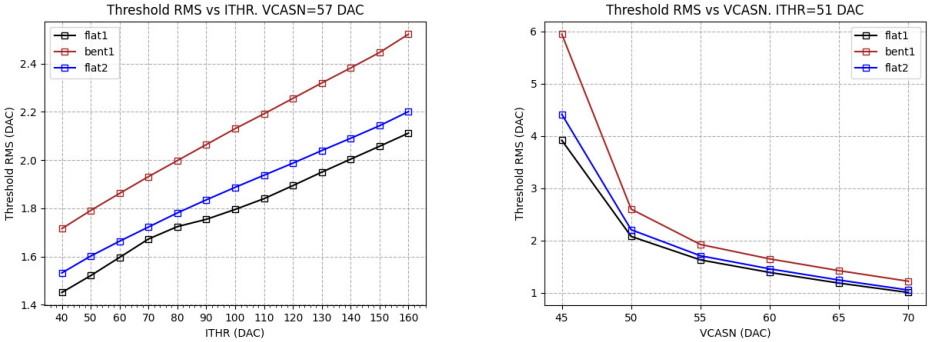


Figure 3.19: Standard deviation of the threshold distributions as a function of ITHR (left) and VCASN (right).

The behavior of the temporal noise over the variation of ITHR and VCASN is shown in figure 3.20. As illustrated, the mean noise increases with both ITHR and VCASN increment. Similar tests were performed with previous ALPIDE prototypes (pALPIDE-1 and pALPIDE-3) [58, 82] where it was observed that the mean noise reduces with increasing ITHR. However, the results of this thesis correspond to the final ALPIDE version and are in agreement with the noise values obtained from measurements with the pALPIDE-3 [82], specially in sector 5, whose circuit design parameters are identical to those used in the final ALPIDE. The differences between older and last ALPIDE circuitry (diode reset, optimized size for M3, M5, M6, and M8, spacing of $3\mu m$, larger pixel size, etc.) contribute to the increment of the noise by manipulating the volume of the current (ITHR) in the front-end transistor which drives the amplification and shaping factor of the signal.

The fake-hit rate has been also evaluated with respect to ITHR and VCASN in figure 3.21. It is clearly visible that fake hit rate drops with increasing ITHR.

Comparing this behavior with the trends shown in figure 3.18, where it can be observed that the ITHR increment results in higher threshold, we can conclude that the highest the threshold the lowest the fake hit rate because pixel has to cross the threshold level in order to record a hit. On the other hand, when VCASN increases, threshold value drops and this leads to higher fake hit rate.

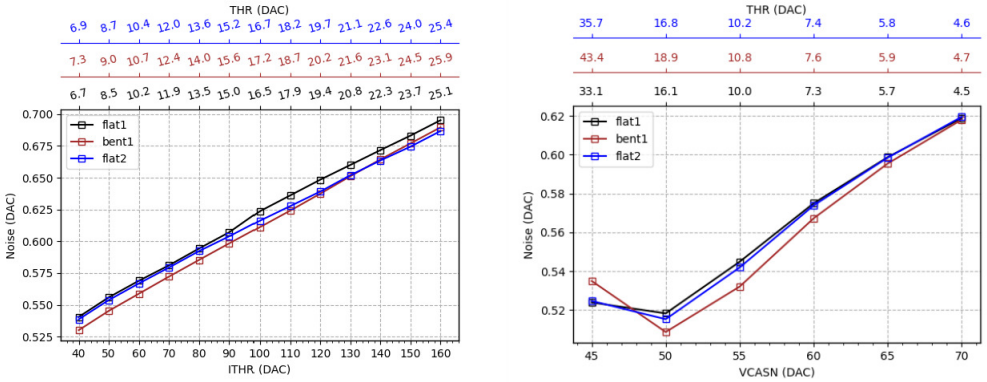


Figure 3.20: Temporal noise calculated as a function of ITHR (left) and VCASN (right).

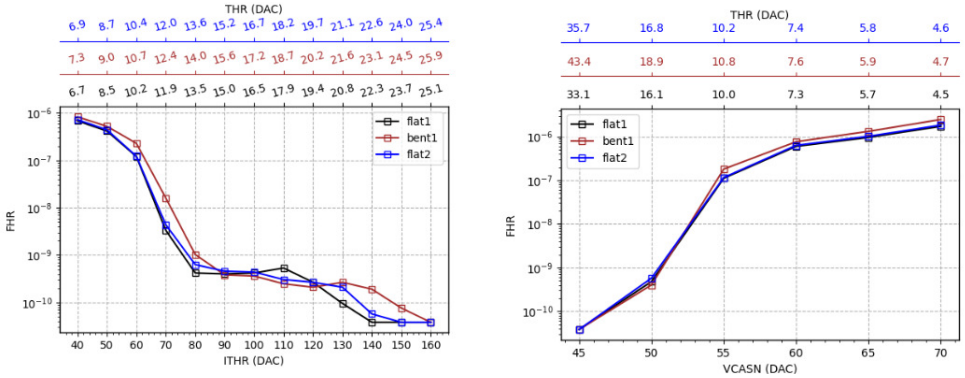


Figure 3.21: Fake-hit Rate values calculated as a function of ITHR (left) and VCASN (right).

3.5 Study on the performance of bent ALPIDEs in test-beam environments

A fundamental step in the investigation of a detector performance is the realization of a so-called test-beam experiment. During such investigation, the

detector to be studied, usually referred to as device under test (DUT), is subject to a particle beam with known properties. The main idea of a test-beam is to use well-known, already characterized detectors as a reference "telescope" and compare the prediction of the charged track trajectories reconstructed with this telescope with the space points obtained from the DUT. These reference detectors are therefore placed in the beam line, before and after the DUT. The position of a traversing beam particle is registered by the detectors of each reference plane. During the analysis of the test-beam data, this information will be used to eventually reconstruct the particle's trajectory called "track", with the further purpose to calculate the detection efficiency and the spatial resolution of the DUT, as well as to evaluate its tracking capabilities. Additional detectors with a sufficient time resolution are usually required for providing an external trigger signal. Plastic scintillators are commonly used for such triggering. In the next sections the analysis of data collected during campaigns realized at the test-beam facility of the German Electron Synchrotron (DESY) located in Hamburg, Germany and at the Super-Proton Synchrotron at CERN will be discussed.

3.5.1 Test-beam facilities

The DESY II

The DESY II synchrotron usually accelerates electrons. After the injection by a linear pre-accelerator at an energy of $E_{min} = 0.45\,GeV$, a bunch of about 10^{10} electrons is accelerated to an energy of $E_{max} = 6.3\,GeV$ inside DESY II. The beam is then stored with an oscillating beam energy ranging between E_{min} and E_{max}. Besides providing the experimental areas with a particle beam, DESY II also serves as an injector for the larger synchrotron PETRA III. The beams provided at the DESY II test-beam facility are produced by so-called parasitic beam generation in contrast to extracting and using the primary beam directly. For this process, a several micrometer thick target wire is placed in the primary beam orbit, such that Bremstrahlung photons are generated at all available beam energies. These high-energy photons tangentially escape the beam pipe of the synchrotron. They hit a secondary conversion target, where electron-positron pairs are generated by the pair production process and then these enter another evacuated beam pipe. The beam particles pass through the magnetic field of a dipole magnet, which forces them to fan out according to their momentum and particle species. An adjustable collimator cuts out and focuses the final beam that is delivered to the experimental area. Finally, a static lead collimator determines the shape of the beam used for the experiment. Due to the oscillating energy pattern of the primary beam inside the DESY II synchrotron, the rate of the parasitically generated test-beam scales with the selected test-beam energy. A higher selected energy leads to a lower particle rate. For this reason an electron beam energy of 5.4 GeV was chosen as a tradeoff between high energy and particle rate for the test-beams performed, whose results are discussed in this work.

The Super-Proton Synchrotron at CERN

A part of the data provided for this study were collected at the Super-Proton Synchrotron (SPS) at CERN. The SPS takes particles from the Proton Synchrotron and, as shown in figure 3.22 (left), is not only used as a pre-accelerator to fill the Large Hadron Collider, but it also provides particles for a range of smaller experiments and it serves as a test-beam facility. Figure 3.22 (right) shows a photograph of the experimental hall at CERN, in which a number of experiments and test-beam setups are located. The SPS has a circumference of around 7 km and accelerates protons up to energies of 450 GeV [83]. Some of these protons are directed onto a target in order to produce secondary particles. The particle beam provided to the test-beam setup normally consists of charged pions with a momentum of 120 GeV.

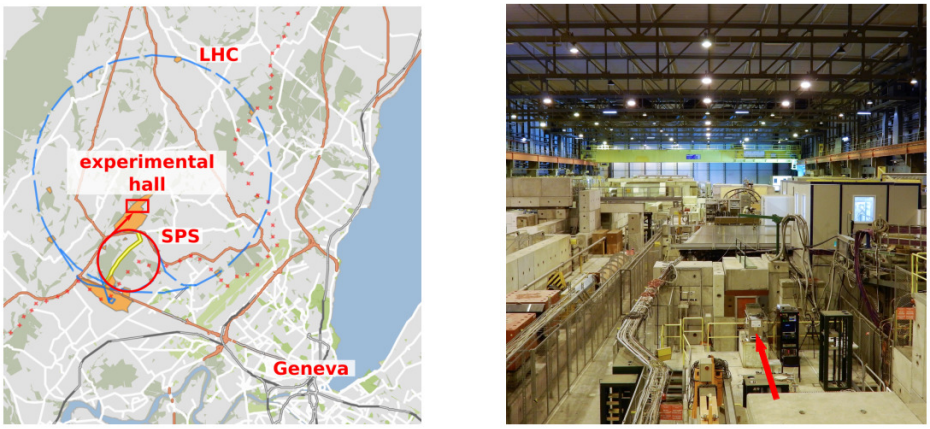


Figure 3.22: (left) Illustration of the CERN accelerator complex. The red arrow indicates the direction of the particle beam towards the experimental hall (red box) in the North Area at CERN. (rigth) Photograph of the experimental hall in the North Area of CERN. The red line indicates the direction of the beam.

3.5.2 Test-beam experiments

Within the scope of this work, the first ever in-beam characterization of a bent MAPS was performed in a test-beam campaign at the DESY II test-beam facility in June 2020 [84]. The data acquired in these measurements with a bent ALPIDE where analyzed and systematically compared with data taken during the same campaign, but with a different telescope featuring a flat ALPIDE as DUT. The study has the main scope to evaluate the sensing performance of a bent chip through the variation of the working point defined by front-end parameters and contrast it with the results of an originally flat sensor.

Another test-beam regarding ALPIDEs bent in a different way was performed in July 2021 at CERN SPS facility. In this experiment a set of DUTs

was composed by six ALPIDE chips, bent at the foreseen ITS3 layer radii, forming the so-called micro-ITS3. The objective is the simultaneous in-beam characterization of several concentrically arranged bent ALPIDE chips, thus modeling the tracking layers of the future ITS3. Additional measurements were carried out with the same detector arrangement by exploring the option of introducing a target in the center of the detector barrel of the micro-ITS3 in order to mimic particles emerging from a real collision. Through this activity the possibility to reconstruct tracks and vertices from hadronic interactions in a geometrical configuration very close to the next full ITS (including ITS3) would be proved. In all the experiments mentioned, a beam telescope featuring six flat ALPIDE chips as reference detectors was used to measure the performance of the flat and bent DUTs. Each of the reference ALPIDE chips is mounted in a separate metal case with a window at the active area of the sensor. Pictures of the mounted setups can be seen at the right of figures 3.24 and 3.25.

All the detectors are connected to a dedicated Data Acquisition (DAQ) board, which delivers the power and reads the corresponding attached chip. If an external trigger signal is provided, all DAQ boards can be daisy-chained in order to distribute this signal to each detector. The same is true for the BUSY-signal that is sent by the ALPIDE chip during readout. While the chip is BUSY no trigger is accepted until the readout process has finished.

Before performing the test-beam experiment, the reference planes are tuned and set to their optimal working points using the procedure and scans described in the previous section. In contrast to the DUTs, the reference planes are always operated with a back-bias voltage of $VBB = -3V$ in order to reduce the noise level and increase the depletion region in each of the pixels. In the following sections, a more detailed description of the experimental setups, specially of the devices under test will be given for each test-beam.

3.5.2.1 Test-beam June 2020

The two setups considered in this study from the test-beam realized in June 2020 were composed by one Device Under Test and 6 reference planes. For what concerns the DUT, as mentioned above, one setup contains a bent ALPIDE and the other arrangement, taken for comparison with the first, was built with a flat chip. Each DUT was placed between two arms, each composed by three reference planes. In the case of the setup with the flat DUT there were 2 cm of distance in between the planes and for the bent device the layers were positioned with different interplanar distances as shown in the figure 3.23. In the global coordinate system the beam direction defines the positive z-axis at the origin of the x-y-plane. The origin along the z-axis is fixed by the placement of the detectors in the geometry of the setup. The local coordinate system is right-handed with the x- and y-axes defining the sensor plane. In the current setups, for all planes the x-axis goes through the chip columns direction and the y-axis goes through the rows.

A picture of the bent DUT can be observed in figure 3.24 (left). As specific procedure to bend the chip, the long edge hosting the bonding pads and the

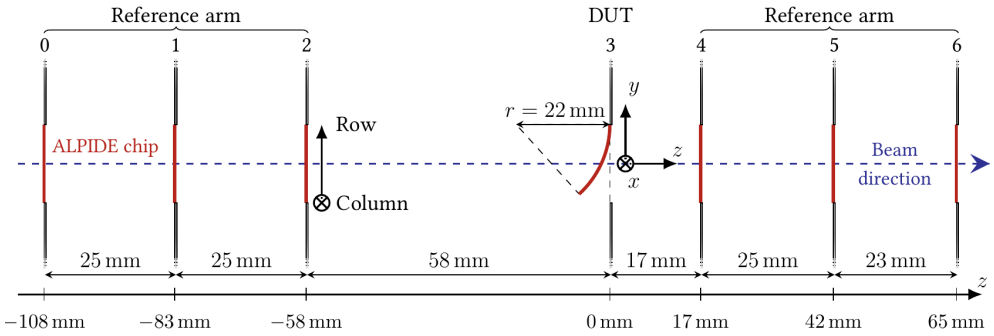


Figure 3.23: Scheme of the ALPIDE beam telescope, featuring a bent sensor as the DUT.

periphery logic was glued onto a carrier board by mean of acrylic adhesive. The carrier card mainly serves as interface between the ALPIDE and the DAQ board, it contains the electronics for signal and power distribution, as well as decoupling elements to filter electronic noise. After gluing, the chip is wire-bonded to the carrier card for electrical connection. The bonding area remains flat, then the rest of the sensor is left unattached and lightly compressed between two layers of polyimide film, which are tied to two lateral wheels that can be moved along the bending direction. By adjusting the position of the lateral wheels and progressively wrapping the polyimide film around them, a curved shape is induced to the chip. Following this method the chip was bent up to curvature radius of less than 18 mm.

A metrological mapping of the sensor surface was performed on a Coordinate Measuring Machine (CMM) and after a fit procedure an approximate average curvature radius of 16.9 mm was calculated. The measurement was repeated with the same CMM setup after the test-beam and resulted in approximately 24.4 mm. The difference between these radii is explained by a relaxation of the polyimide foils with time. Therefore, the curvature radius of the chip during the test-beam is expected to lie in-between the radii measured with the CMM [84]. Figure 3.24 (right) shows the final experimental setup containing the six reference detector planes, arranged in two reference arms of three planes each, and the bent DUT mounted within the ALPIDE beam telescope.

3.5.2.2 Test-beam July 2021

A different approach to test the performance of bent ALPIDE chips was followed at the test-beam carried out at CERN SPS facility in July 2021. Both the setups considered to be studied from this campaign consisted of 12 ALPIDE chips, numbered from 0 to 11. The telescope was configured with six flat sensors as reference planes, three on each side of the DUTs, featuring a fixed minimal distance of 2.5 cm between them. The DUTs (from ALPIDE_3 to ALPIDE_8) were placed in the middle of the reference planes, bent on three concentric 3D-

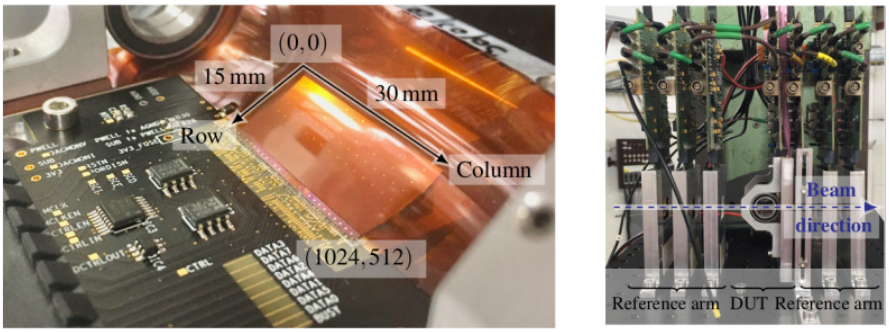


Figure 3.24: (left) Picture of the bent DUT used in the DESY test-beam campaign in June 2020. The ALPIDE chip held in bent position with two polyimide attached to aluminum wheels. (right) Test-beam setup containing the six reference detector planes and the bent DUT mounted within the ALPIDE beam telescope.

printed cylindrical jigs at $r_1 = 18\ mm$, $r_2 = 24\ mm$ and $r_3 = 30\ mm$, which correspond to the intended radii for the ITS3 layers. Due to the mechanical constraints in the connection of the chips the bent DUTs were rotated counterclockwise by about 90^o around the z-axis (beam axis) with respect to the reference planes. Each jig hold two bent chips mounted opposite of each other, i.e placed symmetrically to the center of the cylinder along the beam direction. The measurements were performed in a "single-crossing" configuration, in which the beam is perpendicular to the plane tangent to the chip surface at the chip center, this geometry allows for the simultaneous test of all the sensors in the multi-DUT setup. For the bending procedure of each ALPIDE, the cylindrical tool is attached to a rotational stage with a stepping motor. At the same time, a polyimide foil is attached by one side to the jig and pulled to the other side by a wire sustained at a constant tension. The chip is positioned in-between the tensioned foil and the jig, which automatically rotates in small steps, forcing the ALPIDE to follow the shape of the cylinder. Starting at one of the short sides and moving on towards the other one, the chip bends gradually until it reaches the final position, where it remains fixed by glue previously applied underneath. After the bending procedure has been successful, the polyimide foil is removed. The automatized process allows for better stress control and reproducibility.

Each jig has a thickness of approximately 3.1 mm and is made of polypropylene-like ACURRA-25 material. It constitutes a significant amount of material that can lead to additional scattering. Therefore, a window is cut out of the jigs behind the chips to minimize the material crossed by the beam particles. The dimensions of the windows are $17 \times 9\ mm^2$, which are smaller than the chip dimensions ($30 \times 15\ mm^2$) to ensure the mechanical stability of the fixation of the chip to the jig.

The sensors are first bent and then electrically connected by wire bonds. It

should be noticed that, in contrast with the setup used in June 2020, in this experiment the bent axis is parallel to the long side of the chips. A picture of the resulting DUTs, forming the micro-ITS3 and a top view of the DUTs mounted within the ALPIDE beam telescope can be seen in figure 3.25. Additionally, a schematic of the positions of all the ALPIDEs in the setup is shown in figure 3.26. Since the new sensor that will be used in the future ITS3 is intended to contain a different pixel matrix and circuit structure, the direction of its curvature in the final detector will be diverse from the one described above for the ALPIDEs used in the test-beam. However, this bending procedure allowed to build a structure of single sensors with a geometry similar to the ITS3 and to perform their simultaneous in-beam characterization.

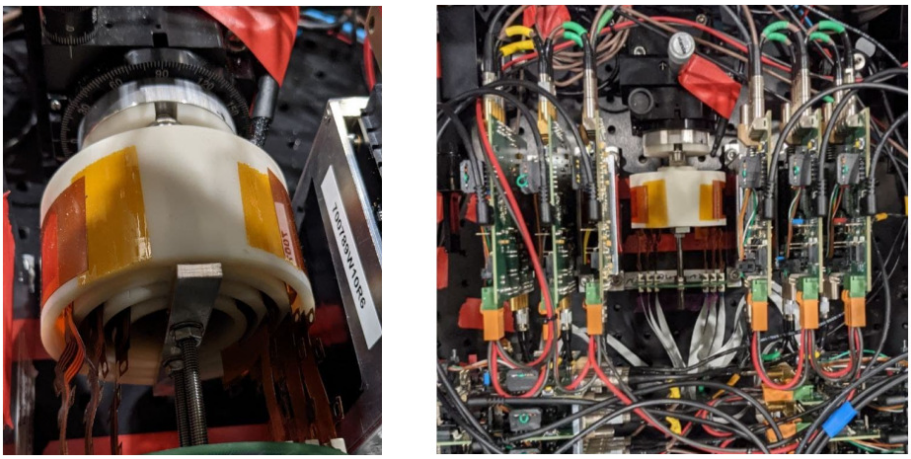


Figure 3.25: (left) Image of the six DUTs used in the July testbeam campaign, bent over three cylindrical jigs at different radii, forming the micro-ITS3. (right) Top view of the bent DUTs mounted within the ALPIDE beam telescope.

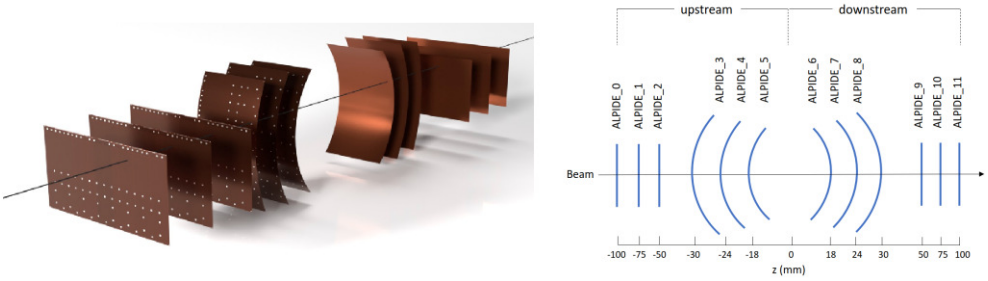


Figure 3.26: (left) 3D render and (right) schematic side view of the beam telescope with the six bent DUTs used at the test-beam campaign in July 2021.

As mentioned above, two experimental setups have been considered for the

study of data from the test-beam campaign in July 2021. The first one is totally described by the schematics shown in figure 3.26 and the second one featured the same detector arrangement, but in this case a cylindrical copper target of 5 mm diameter and 1.16 mm thickness, mounted in a plastic holder, was placed at the center of the micro-ITS3 as shown in the figure 3.27.

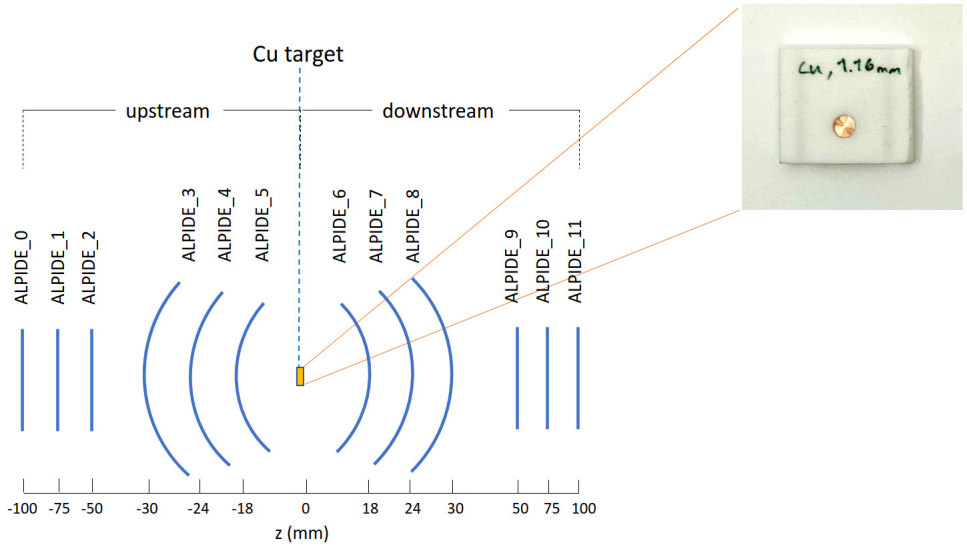


Figure 3.27: Schematic drawing of the beam telescope and the six bent DUTs used at the test-beam campaign in July 2021 with a copper target placed approximately at the center of the arrangement.

3.5.3 Test-beam reconstruction chain with Corryvreckan

For the scope of this work the analysis framework Corryvreckan is used in order to analyze the collected data. Corryvreckan is a flexible, fast and lightweight test beam data reconstruction framework based on a modular concept of the reconstruction chain. It is designed to fulfill the requirements for offline event building in complex data-taking environments. It is written in modern C++ and due to its modular structure it is highly versatile and easily extendable. It reduces external dependencies to a minimum by implementing its own flexible but simple data format to store intermediate reconstruction steps as well as final results [85].

Data acquisition and readout

In order to collect data from all involved detectors during a test-beam, a data acquisition framework is required. The data stream produced by each sensor is assembled, eventually converted to a different data format and written to a raw

data file for further analysis. For the test-beam campaigns considered in this work the software EUDAQ has been used. As a DAQ framework designed for the use in test-beam campaigns, it provides easy integration of high-precision beam telescopes using pixel detectors.

Each data producing instance connected to the DAQ system, e.g. the ALPIDE sensors and DAQ boards, is controlled by a so-called "Producer". Producers are the source of the data taking and writing channel of EUDAQ, these instances are the bridge between the user-specific hardware and the DAQ software and need to be written for every type of detector involved. It is able to initialize the corresponding detector as well as start and stop the sending of data on demand from the Run Control (it acts as the user interface of the DAQ system; a continuous data taking period is referred to as run). An EUDAQ event is built for each trigger, which can be written to storage in a raw file. Consequently, such data file is a serialized collection of detector raw data organized in (EUDAQ) events. Each event consists of a set of status information, such as run number, trigger number and timestamp followed by the unprocessed binary output from the sensor [86, 87]. The status part of the data is referred to as header and provides the means to organise events sent from different Producers. The information of each detector is then sorted and organized according to the header information.

Corryvreckan features modules to read and process data written by EUDAQ. The binary data format written to the raw data file can be converted back to the standardized EUDAQ specific format and then be read into Corryvreckan. The raw data are read plane-by-plane and then event-by-event. Hence, the timing information of the first processed detector plane defines the beginning, the end and the duration of the corresponding Corryvreckan Event. Now the information of each hit pixel, e.g. the position, is copied to the so-called Clipboard, which is a temporary event storage allowing all modules of Corryvreckan to access this data for further processing. The transfer of data to the Clipboard is only made if the timing information of the current plane matches the one of the Corryvreckan event. By default, the event building process for the EUDAQ event loader is based on timestamp information, but the software can be set to only consider the trigger number for the assignment of hit information from different detector planes to a Corryvreckan Event.

Coordinate systems and geometry file

The software relies on various coordinate systems chosen as right-handed Cartesian, where the z-axis represents the beam direction. Primarily, these systems can be divided into two groups, the so-called local and global. Two separate local coordinate systems are attributed to describe points on the surface of each detector and one global system is used in order to relate the local coordinates for each relevant detector in the setup.

From the raw data the position information of a hit pixel is only available as its address, so the first local coordinate system has its origin in the lower left corner of the chip. The column and the row numbers of the hit pixel are

displayed on the x-axis and y-axis respectively. Thus, the described detector surface always lies in the x-y-plane, independently if the real detector is bent or not. For this reason, the z-coordinate is conventionally set to zero independent of the x-y-position of the hit. Given the pixel pitch and the number of pixels in both directions, the column-row coordinates can be converted to a mm-scale. From this conversion, the second local coordinate system centered in the chip middle, is derived.

In order to relate the local coordinate systems for the detectors to each others, a global reference point is passed to the software via a geometry configuration file. The distances between the detectors needs to be measured by hand. The measured z-coordinate of each detector is added to the respective local coordinates in order to add it to the global coordinate system. The definition of the global coordinates for the campaigns in June 2020 and July 2021 are illustrated in the figures 3.23 and 3.26 (right) respectively.

The state of the art official version of the Corryvreckan [85] test-beam data analysis framework does not support curved sensor geometries. Cylindrical coordinates were implemented by the addition of a new detector class "BentPixelDetector" [88] to the standard software. The general idea of implementing the bent sensor is to first transform from the flat local coordinate system, to a local-bent one. In a second step the usual translations and rotations according to the geometry configuration file are applied to the local-bent coordinates thus transforming them to the global frame.

Before running Corryvreckan, a geometry file configuration of the detectors must be set. Inside it, the data related to the intrinsic resolution of the various telescope planes are entered, as well as the relative distances between the planes within the experimental setup. The Corryvreckan execution reconstructs the event and performs the analysis taking the geometry file as a first input. In addition, a so-called Region Of Interest (ROI) can be defined for each detector. It allows tracks or clusters to be marked as within a certain region on the respective detector. This information can be used in analyses to restrict the selection of tracks or clusters to certain regions of the device, e.g. to exclude known bad regions from the calculation of efficiencies. As such, only tracks featuring an intersection point within these (rectangular) regions are considered for further analysis. The ROI is defined as a polynomial in local pixel coordinates of the device using the roi keyword. A rectangle could, for example, be defined by providing the four corners of the shape.

Masking pixels

Due to the manufacturing process or handling of the chips, material damage can be introduced, causing malfunction of single pixels or even full double columns. The pixels affected could show an increased hit rate due to electronic noise, which influences the analysis results, e.g. the track reconstruction. Consequently, such noisy pixels need to be identified and excluded from the study. In Corryvreckan, pixels can be classified as noisy with the frequency method: here a global average pixel hit rate is calculated for each detector and the hit rate of

each single pixel is compared to this sensor average. If a pixel fired a defined amount of times more than the mean global hit rate it is recognized as noisy and its address is written to a mask file, which is produced for each detector and read back by the event loader modules at the beginning of the further analysis stages.

Alignment procedure

An analysis with Corryvreckan requires a configuration file defining which detectors are present in the setup. This file also contains the position and rotation of each detector plane. The z-positions of all planes can and must be measured by hand in the existing test beam setup and entered in this configuration file for the analysis. The x- and y-positions as well as the rotations angles with respect to each coordinate axis of the global coordinate frame cannot be measured precisely by hand. However, these quantities have a strong influence on the tracking since a misalignment of a single plane with respect to another by a fraction of a millimeter can correspond to a shift of several pixel pitches, resulting in reconstructed particle tracks significantly distorted, i.e. they will probably not describe correctly the actual particle trajectory, leading to biased analysis results. Consequently, an alignment procedure is needed. It consists in an iterative operation during which shifts and rotations are applied to the data of all the planes (by adjusting one plane at the time) in order to obtain a track residual, namely the distribution of the spatial distance between the interpolated track intercept and the associated spatial information on the plane, centered around zero with the narrower possible dispersion. Track residuals in x-y plane are measured with respect to a plane used as a reference. The width of the distribution depends on the tracking resolution of the telescope and is influenced by many factors such as the beam energy, the material budget of the detector planes, the distance between the detector planes, etc.. The procedure requires the detector spatial resolution, because this defines the uncertainty on the spatial information and therefore influences the track χ^2.

After each alignment iteration an updated geometry file is produced by Corryvreckan with the updated geometry found during the alignment. At the beginning of each stage of iteration, the initial position of each detector plane is modified (shifted and rotated) based on the results obtained in the previous stage.

Clustering and pre-alignment

Depending on the incident point and angle of a traversing particle, it may create electron-hole pairs in two or more adjacent pixels. In addition, charge can "leak" from one pixel into neighboring cells by lateral diffusion. This effect is called charge sharing. As a consequence, a clustering algorithm is needed to group such pixels into one cluster, since they belong to a single particle crossing. Since no charge-equivalent information is available from the ALPIDE, the cluster center positions are determined by arithmetic mean calculation.

The main scope of the prealignment is to provide a first rough alignment of the test-beam setup. As such, no tracking information is used during this stage. Independently of its role, i.e. reference or DUT, the prealignment can be performed simultaneously for each detector plane. Only translational shifts in the x- and y-coordinates are performed, while rotations are subject to the subsequent alignment stage. The global z-coordinate is kept fixed at the measured value given in the geometry file. To have an initial estimation of the misalignment of the detector planes with respect to each other, the spatial correlations are investigated. In the context of test-beam reconstruction it can be defined at a pixel or cluster level. In the latter case, the spatial correlations are calculated between a specific detector and the reference detector as:

$$x_{correlation} = x_{cluster\ on\ reference\ detector} - x_{cluster\ on\ this\ detector} \tag{3.4a}$$

$$y_{correlation} = y_{cluster\ on\ reference\ detector} - y_{cluster\ on\ this\ detector} \tag{3.4b}$$

One-dimensional hit coordinate correlations are used in order to identify the values of the respective x- and y-shifts. In order to determine the shift for the prealignment, a Gaussian is fitted to the distribution of the hit coordinate differences between the involved detector planes in both, x- and y-direction. The shifts applied to the respective planes is then determined by the mean value of the fitted Gaussian. Other methods like the mean of the 1D correlation histogram or its maximum can be used to determine the translational shifts [85]. The correlations do not necessarily have to be centered at zero because an offset might reflect a physical displacement of a detector plane in x and y with respect to the reference plane.

Tracking and reference telescope alignment

For the next steps the reference arms and the DUT(s) are treated separately to avoid biasing the analysis of the DUT(s). The trajectory of the particles can be reconstructed from the cluster centers in the telescope planes by fitting these spatial points with a certain track model. Corryvreckan supports the straight-line track model, which ignores the effect of multiple scattering and general-broken-line (GBL) track model, which accounts for scattering effects [89]. Given their large momentum (120 GeV) and the low material budget of the test-beam setup, soft multiple-scattering of the beam particles is considered a negligible effect. Therefore, the track finding works as follows. All combinations of clusters in the first and the last hit detector plane are connected to form a straight line. Clusters in further detectors are consecutively added if they are within the spatial cuts (in local coordinates) and time cuts, updating the reference track at each stage. As mentioned, the DUT planes are excluded from the track finding.

The alignment of the beam telescope planes is realized by optimizing the χ^2 distribution of the track fits taking into account a sufficiently large set of

reconstructed tracks. If χ^2 is divided by the number of degrees of freedom (ndof), which is given as the number of fitted points (3 translational directions per point) subtracted by the fit parameters constrained by the model, it is referred to as reduced χ^2, i.e. χ^2_{red}. If the value is close to unity, then the data are acceptably described by the fitted model. Thus the χ^2_{red} distribution can be considered as a measure of the goodness of the setup alignment. Therefore, a good alignment is characterized by a χ^2_{red} distribution peaking at low values.

In the alignment procedure, the detector planes are shifted and rotated iteratively relative to the detector marked as reference to increase the tracking quality. For the scope of this work, the module named Alignment Millepede is used. It features an implementation of the Millepede algorithm, which allows a simultaneous fit of both the tracks and the alignment parameters [85]. The algorithm stops if a specified convergence level is reached, i.e. the absolute sum of all corrections over the total number of parameters, is smaller than the configured value.

Alignment of the Devices Under Test

After the alignment of the beam telescope, also the DUTs need to be aligned with respect to the reference planes. With the track information from the beam telescope, residuals can be calculated, i.e. the distance between the track interception point at the DUT sensor surface x_{track} and the corresponding cluster center position on the DUT. A set of absolute values ($x = 100\,\mu m$ and $y = 100\,\mu m$), describing an ellipse, are defined as the maximum spatial distance in the local XY plane between clusters and a track for the track association. If the reconstructed track and the associated cluster on the DUT originate from the same traversing particle, it is expected that the absolute value of the residual is minimal in the case of a good DUT alignment. Ultimately, the goal of the alignment is to force the residuals to be centered around zero. A deviation is caused by the limited position resolution of both the DUT and the sensors contributing to tracking and if the considered particle is scattered by the detector material.

The residual distributions of the DUT(s) are not expected to be strictly Gaussian, as they would be in case of dominance of the uncertainty related to multiple-scattering over that from the intrinsic resolution of the detector. The figure 3.28 shows the residual distributions in x-(blue) and y-(red) directions after the alignment of the six bent DUTs from a representative run of 304631 events with the setup shown in figure 3.26 (right) and the ALPIDE sensor planes configured as reported in 3.1, corresponding to the July 2021 test-beam. It can be noticed in figure 3.28 that, as expected, the mean of x and y distributions for all the DUTs are almost zero, indicating the achievement of a good alignment in absence of multiple-scattering of the beam particles in the sensors material due to the low material budget of the ALPIDE sensors (0.05 % X_0 except the ALPIDE_0 with 0.1 % X_0) and the high beam energy (120 GeV).

Table 3.1: Configuration of the ALPIDE sensor planes in the run with DUT residual distributions shown in figure 3.28

ALPIDE (position from beam)	Thickness (μm)	ITHR (DAC)	VCASN (DAC)	Threshold (e^-)
0 (REF0)	100	50	112	97.1
1 (REF1)	50	50	112	85.6
2 (REF2)	50	50	112	90.2
3 (DUT0)	50	50	55	109
4 (DUT1)	50	50	58	109
5 (DUT2)	50	50	57	106
6 (DUT3)	50	50	55	105
7 (DUT4)	50	50	55	110
8 (DUT5)	50	50	55	105
9 (REF3)	50	50	105	103.9
10 (REF4)	50	50	105	98.4
11 (REF5)	50	50	105	101.1

Analysis quantities

Having a fully aligned setup, the actual analysis of the data can be performed for the set of reconstructed tracks, using dedicated analysis modules. As a starting point the intersection point for each track with the DUT sensor surface is calculated. Next, clusters within an elliptic search window, whose dimensions are chosen by the user, are associated to the track and if more than one cluster is found within the fiducial area, only the closest cluster is considered. Based on the track intercept point and the associated cluster positions on the DUT a range of analysis quantities are calculated. The ones relevant for the scope of this work are described in the following.

- **Detection efficiency**

  The detection efficiency ε_{det} represents the probability of a sensor to detect traversing particles. It is determined by two effects: first, the charge collected in the seed pixel (and the related distribution) influenced by the geometry of the collection diode and the surrounding region within the single pixel, the reverse substrate bias and the particle type and energy, and second, the charge threshold. For each accepted track in the beam telescope a corresponding signal on the DUT is searched. In the case a cluster can be associated to the track, it is counted as efficient or matched; if not, it is counted as inefficient. The fraction of tracks with associated clusters over the total accepted tracks represents the detection efficiency of the DUT. It can be determined as follows:

$$\varepsilon = \frac{\#\, tracks\ with\ associated\ cluster}{\#\, tracks\ with\ +\ without\ an\ associated\ cluster} \tag{3.5}$$

  The efficiency uncertainty is calculated by applying a Clopper-Pearson

confidence interval of one sigma. Taking into account the respective upper and lower limit of 1 and 0, this corresponds to the central 68.3 % of an assumed binomial distribution for a given efficiency [85, 90].

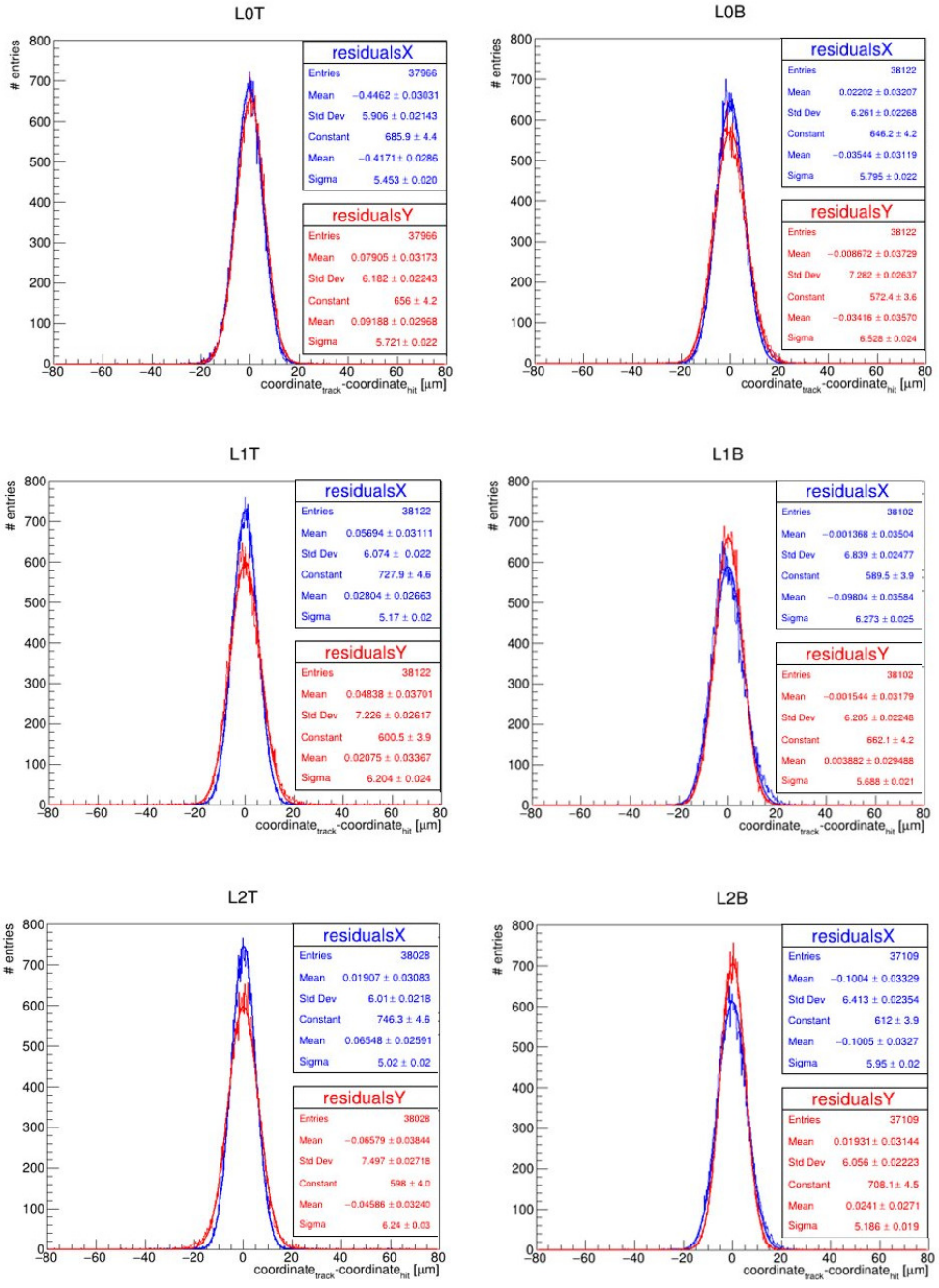


Figure 3.28: Residual distributions in x-(blue) and y-(red) directions after the alignment of the six bent DUTs from a representative run with the setup shown in figure 3.26 (right), corresponding to the July 2021 test-beam. The distributions are centered around zero and resemble a Gaussian shape.

- **Position resolution**

  As the detection efficiency, the position resolution represents a central quantity to characterize the performance of a tracking detector. It is indirectly obtained from the test-beam data by measuring the standard deviation σ_{res} of the reconstructed hit position with respect to the extrapolated track crossing point. This is referred to as the distance between track intersection point and associated cluster position in each coordinate direction, i.e. the respective (track) residual:

$$x_{residual} = x_{track\ intercept} - x_{associated\ cluster} \tag{3.6a}$$

$$y_{residual} = y_{track\ intercept} - y_{associated\ cluster} \tag{3.6b}$$

  σ_{res} contains contributions either from the intrinsic spatial resolution (indicated as "position resolution, σ_{pos}" in the following) of the DUT and the track uncertainty σ_{track}. From the residual distribution, the intrinsic spatial resolution of the DUT can be estimated by:

$$\sigma_{pos}^2 = \sigma_{res}^2 - \sigma_{track}^2 \tag{3.7}$$

  The track resolution σ_{track} at the DUT in turn depends on the intrinsic resolution of the telescope planes and physical uncertainties such as multiple scattering. The tracking uncertainty can be determined by a Monte Carlo simulation of the beam telescope. From previous simulations it follows that $\sigma_{track}^2 < \sigma_{pos}^2$, since for the reference also ALPIDE sensors are used. Considering the six reference planes used to reconstruct the tracks with intrinsic resolution of 5 μm, which is the value required for the ITS2 and also for the ALPIDE (27 μm × 29 μm pitch), obtained with an average cluster size of 3, the track resolution results $\sim$ 2 μm. Taking the variance σ_{res} of the residual distributions of ALPIDE_5 (L0T) from figure 3.28 and applying the equation 3.7, the position resolution of this DUT results $\sigma_{pos_x} = 5.07\,\mu m$ and $\sigma_{pos_y} = 5.36\,\mu m$.

- **Cluster size**

  The cluster size is defined as the number of pixels forming a cluster. It depends on how much signal charge is generated by the traversing particle. The size of the associated cluster is an important quantity to be investigated, specially for the case of bent sensors, since more or less charge is produced depending on the beam incident angle on the sensor surface and thus the traversed distance of the particle within the detector material. If the particle makes a longer path inside the sensor, it can generate charges

in more adjacent pixels, leading to a larger cluster. This kind of measurements do not necessarily require tracking information from a test-beam data set, but as presented in the section 3.5.4.1 they can still profit from it.

3.5.4 Analysis results

The data acquired during the test-beam campaign at the DESY II test-beam facility in June 2020, have been analyzed and systematically compared with data taken with a flat ALPIDE. The results on the sensing performance evaluation of a bent chip through the variation of the working point defined by front-end parameters and their comparison with the results of an originally flat sensor, are presented in this section. On the other hand, the analysis of efficiency, cluster size and preliminary position resolution are discussed for the experiment realized in July 2021 at CERN SPS facility, where a set of six ALPIDE chips were bent at the foreseen ITS3 layer radii, forming the so-called micro-ITS3. Here, the simultaneous in-beam characterization of several concentrically arranged bent ALPIDE chips, has been carried out, thus modeling the tracking layers of the future ITS3.

Additionally, measurements realized by exploring the option of introducing a target in the center of the detector barrel of the micro-ITS3 to mimic particles emerging from a real collision, are also introduced in the section. The results of the reconstruction of tracks and vertices from hadronic interactions in a geometrical configuration very close to the next full ITS (including ITS3) are also briefly discussed.

3.5.4.1 Test-beam June 2020

The data sets analyzed from this campaign for flat and bent DUTs were chosen corresponding to the chip parameters as follows. The bias voltage for both sensors was 0V. For the flat DUT, VCASN = 61 and ITHR values ranging from 50 to 100 with steps of 10 were taken. For the bent chip the selected values were VCASN = 64 and ITHR from 31 to 66 with steps of 5.

Tracks association As a result from the comparison between the clusters and the interpolated track positions, the spatial distributions of the tracks with associated clusters on both flat and bent DUTs for representative runs are shown in figure 3.29 (top) and (bottom) respectively.

The empty areas at the upper and right sides of the maps, i.e. near the row 512 and the column 1024 of the chips, result from the relative position of the DUTs with respect to the reference planes, which defines a “geometrical acceptance” of the tracking telescope given by the limits of the overlapping regions. In the bent chip, the area near the row 512 is glued to the carrier card which impacts the tracking performance due to the additional material, resulting in a reduction of associated tracks. The dark points on the maps, showed as pixels with zero efficiency, corresponds to “dead columns”. Given

the data-driven readout architecture of ALPIDE explained in section 3.3, where all the pixels in a double column are connected to the same Priority Encoder instance, a malfunction in one of this circuits can result in a zero or faulty communication with the entire double column. In the former case, the pixels readout is completely null, so they are "naturally" masked, in the latter case their State Register output bits could be individually masked by setting the control bit MASK_EN as high. In the figure 3.29 (top) three partially empty double columns can be observed. As mentioned, those pixels could be affected by a failure in the AERD circuit, but in this case, instead of being dead, they featured an increased hit rate and were classified as noisy by Corryvreckan, consequently they were excluded from the analysis.

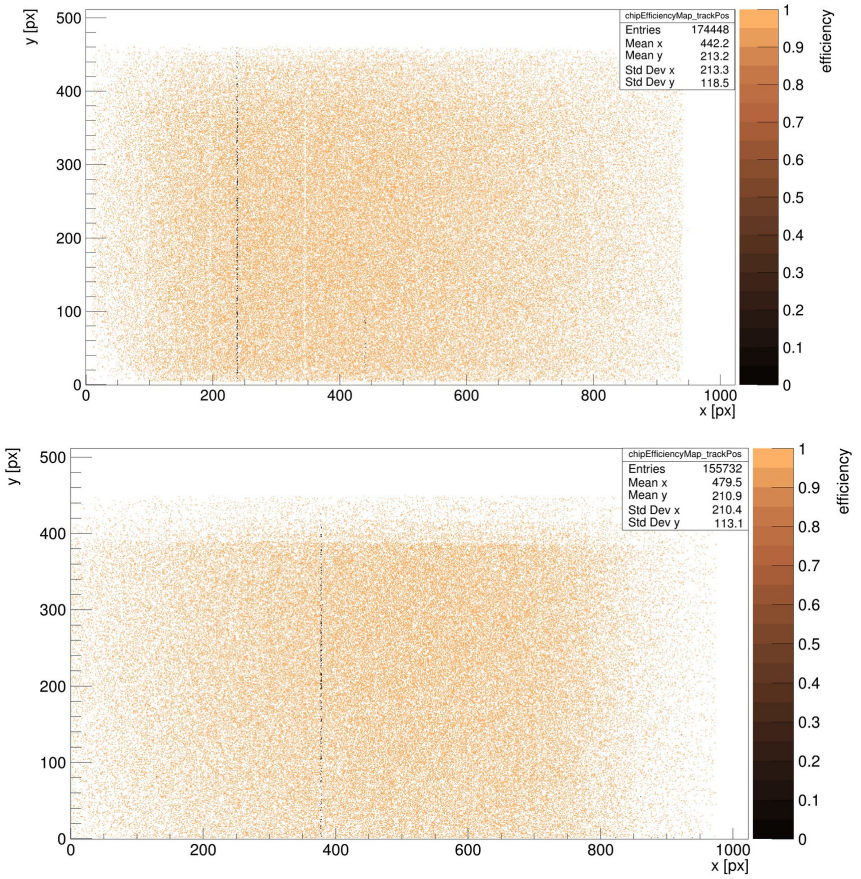


Figure 3.29: Spatial distribution of the tracks with associated clusters intersecting the (top) flat and the (bottom) bent ALPIDE sensors used as DUTs during the June test-beam campaign.

Cluster size The figure 3.30 shows the dependence of the average cluster sizes on the DUT threshold for the flat (black) and the bent (red) sensors. Each point in the plots represents the average value of the cluster size distribution for all the events of the corresponding run. In order to ensure that a cluster is produced by a real particle and not by noise or any other effect, the tracking information was used and only sizes of clusters properly associated to tracks were considered. Having fulfilled the criterion of being in close proximity of the calculated intercept of a track with the sensor surface of the DUTs, it is negligible the probability that such a hit does not originate from a traversing charged particle. As observed in the plot, each data set corresponds to a different ITHR range due to the availability of data.

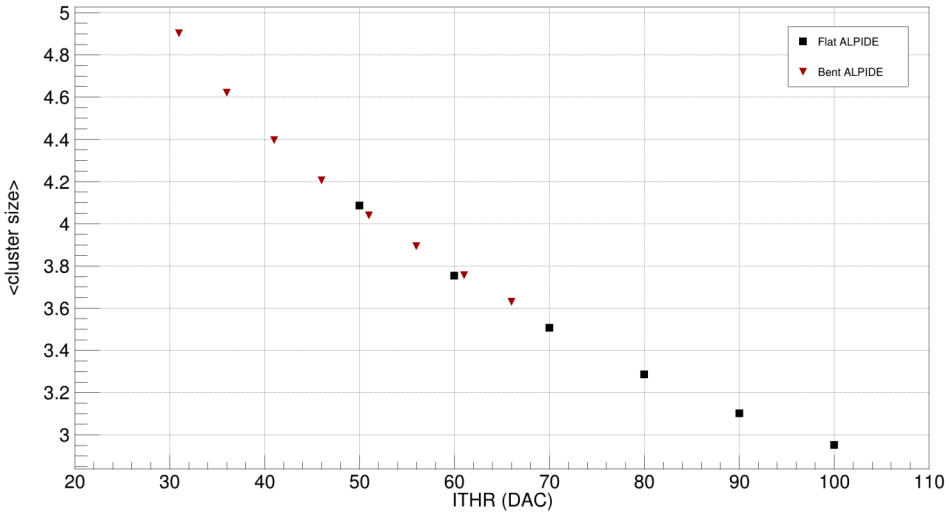


Figure 3.30: Average cluster sizes vs the ITHR values for the flat (black) and bent (red) ALPIDE sensors used as DUTs during the June test-beam campaign.

The largest fraction of the signal charge generated by a particle traversing the sensor is most probably collected in the primarily hit pixel. The charge shared to an adjacent pixel has to overcome the in-pixel threshold in order to be registered as a hit. Consequently, the cluster size is expected to be strongly dependent on the sensor threshold. As such, the decrease of the cluster sizes with the increment of the threshold values is clearly observed for both point distributions in figure 3.30, which are fairly consistent. For lower thresholds even more adjacent pixels show a hit due to the shared charge. Thereby the average cluster size resulting from a charged particle should further increase.

The cluster distributions measured at different ITHR configurations, corresponding to points in figure 3.30, are shown in figure 3.31. For bent and flat sensors a higher frequency of larger clusters is observed at lower ITHR/threshold,

while at higher ITHR/threshold the occurrence of single pixel and smaller clusters increases as predicted from previous simulations [31]. The shape of the distributions for the bent sensor at ITHR=61 DAC and the flat sensor at ITHR=60 DAC are quite equivalent, which indicates a comparable response of both devices at similar conditions.

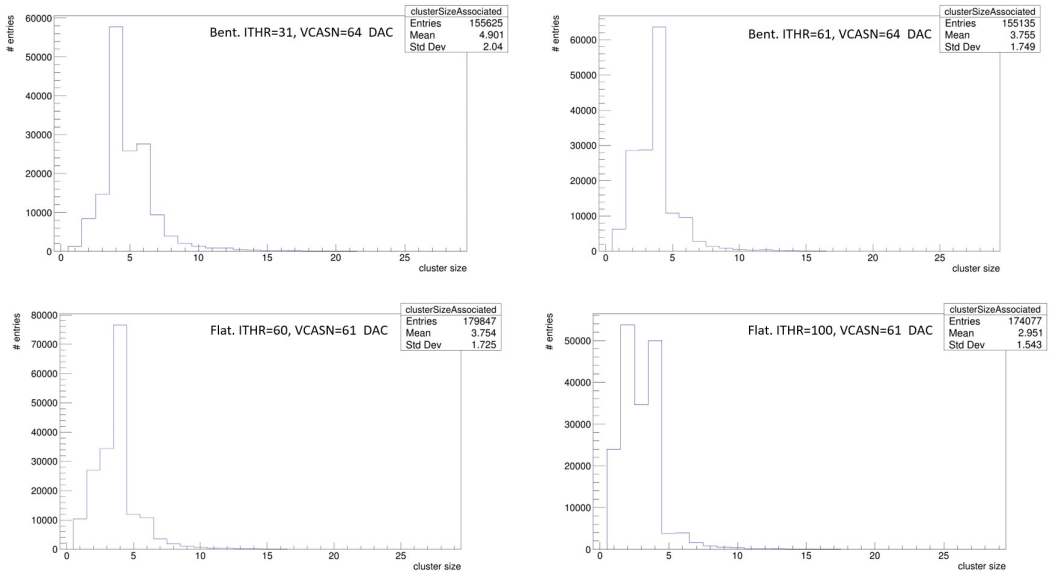


Figure 3.31: Measured cluster size distributions at different ITHR for flat (bottom) and bent (top) sensors.

The particle incident angle is obviously dependent on the position on the bent DUT sensor surface given the test-beam setup geometry displayed in figure 3.23. The particles traversing the bent sensor with an incident angle $\alpha > 0$ with respect to the sensor surface normal will travel a longer distance inside the detector material, generating more free charge by ionization. This leads to an increased total amount of charge shared in adjacent pixels, which could be sufficient to trigger a hit, resulting in larger cluster sizes. The profiles of the cluster sizes along the rows direction are performed for both flat and bent DUTs in the figure 3.32. For every row, the average of all cluster sizes along the columns direction is displayed. For the bent ALPIDE, clusters become larger as increasing the chip curvature, i.e. towards the row 0. These results confirm that particles featuring a larger incident angle produce larger clusters, given their increased path in the active volume of the detector and the generation of more free charge carriers along their trajectories. This behavior has been already anticipated by a previous analysis of test-beam data from June 2020 [84].

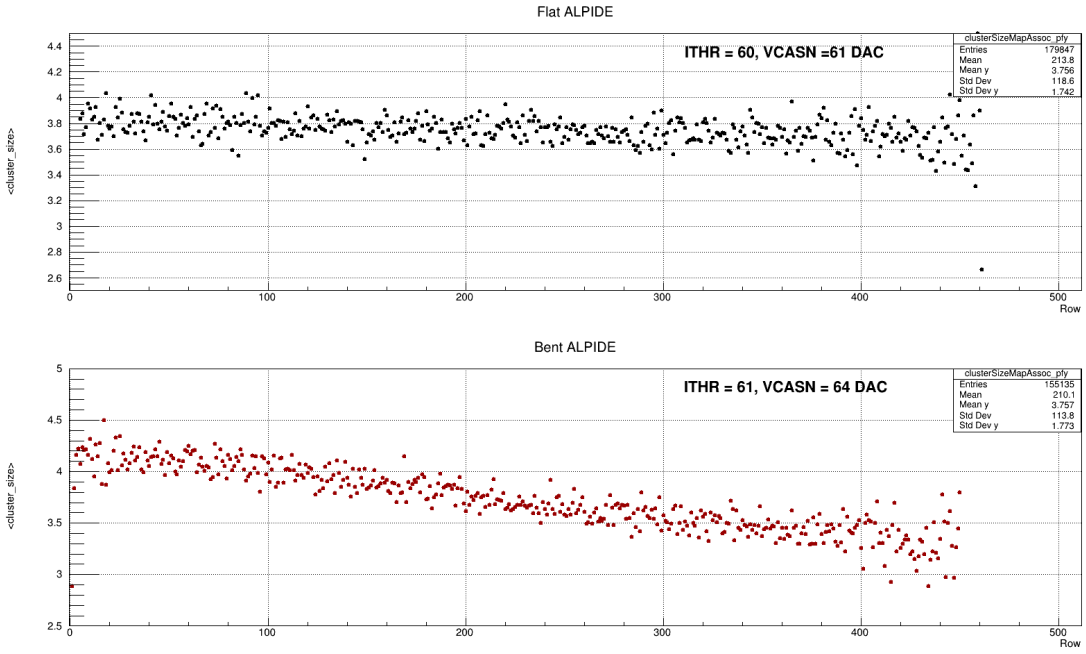


Figure 3.32: Profiles of the cluster sizes along the rows direction for the flat and bent DUTs used during the June test-beam campaign.

Detection efficiency The values of efficiency over the selected threshold ranges for both DUT geometries are shown in the figure 3.33.

A decreasing trend with respect to the threshold increment is observed. This behavior is expected given that for higher in-pixel thresholds the probability that a charge generated by a traversing particle is not sufficient to trigger the pixel to fire, i.e. registering a hit, is increased. Consequently, it might occur that even though a particle passed through the DUT no hits are registered and no cluster can be associated to the corresponding accepted track, thus resulting in a decreased efficiency. The effect gets even larger with further increased threshold values, explaining the trend visible in figure 3.33. Additionally, the higher efficiencies of the bent chip can be result of the larger cluster sizes, which would make easier the association between tracks and clusters. Both chips show an efficiency higher than 99.7% over the full threshold range, which indicates that, in principle, bending an ALPIDE following the procedure described in section 3.5.2.1 has no substantial impact on its performance.

3.5.4.2 Test-beam July 2021

From the test-beam campaign in July 2021 two different data sets were analyzed. The first one corresponds to the setup featuring the ALPIDE telescope

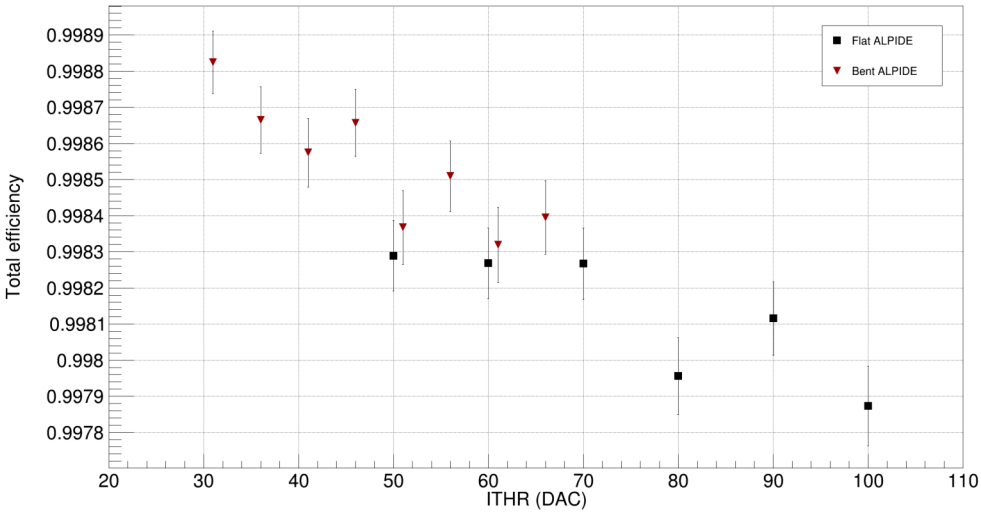


Figure 3.33: Total efficiency as a function of the threshold values for the flat and bent ALPIDEs used as DUTs during the June test-beam campaign.

and the micro-ITS3 as described in figure 3.26. Here, the six bent ALPIDEs used as Devices Under Test were characterized in terms of detection efficiency and spatial resolution over a range of working points. The second setup, as illustrated in figure 3.27, was composed by the same telescope geometry with the addition of a cylindrical Copper target of 5 mm diameter and 1.16 mm thickness, placed at the center of the micro-ITS3. The study is focused on the calculation of the track impact parameter resolution, i.e. the standard deviation of the distribution of Distance of Closest Approach (DCA) for the tracks from the particle trajectories after interactions on the target to the primary vertex reconstructed from the fitted "beam" track.

Telescope with micro-ITS3

Detection efficiency As described in section 3.5.2.2, a window of $17 \times 9 mm^2$ is cut out from the jigs behind the chips to minimize the material crossed by the beam particles. In order to not disturb the efficiency measurement with increased scattering angles resulting from material overlapping in the regions of the DUTs covered by the cylinders material, a region of interest (ROI) was defined. However, due to the rotation of the DUTs with respect to the reference planes by 90^o, there is a "natural" ROI given by the limits of the intersecting regions, which corresponds to the "geometrical acceptance" of the tracking telescope. This effect can be observed in figure 3.34, where the spatial distributions of the associated clusters on all the DUTs in the chips local coordinate system

are shown. The areas defined by the points represent the intersection sector of the corresponding DUT with the telescope planes, while the shaded areas are the regions of interest defined manually.

The detection efficiency for several runs featuring different working points was obtained for all the DUTs and is shown in figure 3.35. It can be seen that between $80e^-$ and $160e^-$ there is a plateau region with sensor efficiencies above 98%. The expected trend is visible for all the sensors given the already discussed reasons. Going towards higher threshold values the efficiency starts to significantly decrease. As reported in [84], this effect becomes sizable above a threshold of about $150e^-$. Around the nominal working point of $100e^-$ the obtained efficiencies are approximately of 99%. This results represent an evidence that the detection efficiency of a system of ALPIDE chips, bent in a configuration similar to that of the future ITS3 layers, as an example for curved CMOS MAPS is retained as compared to flat sensors.

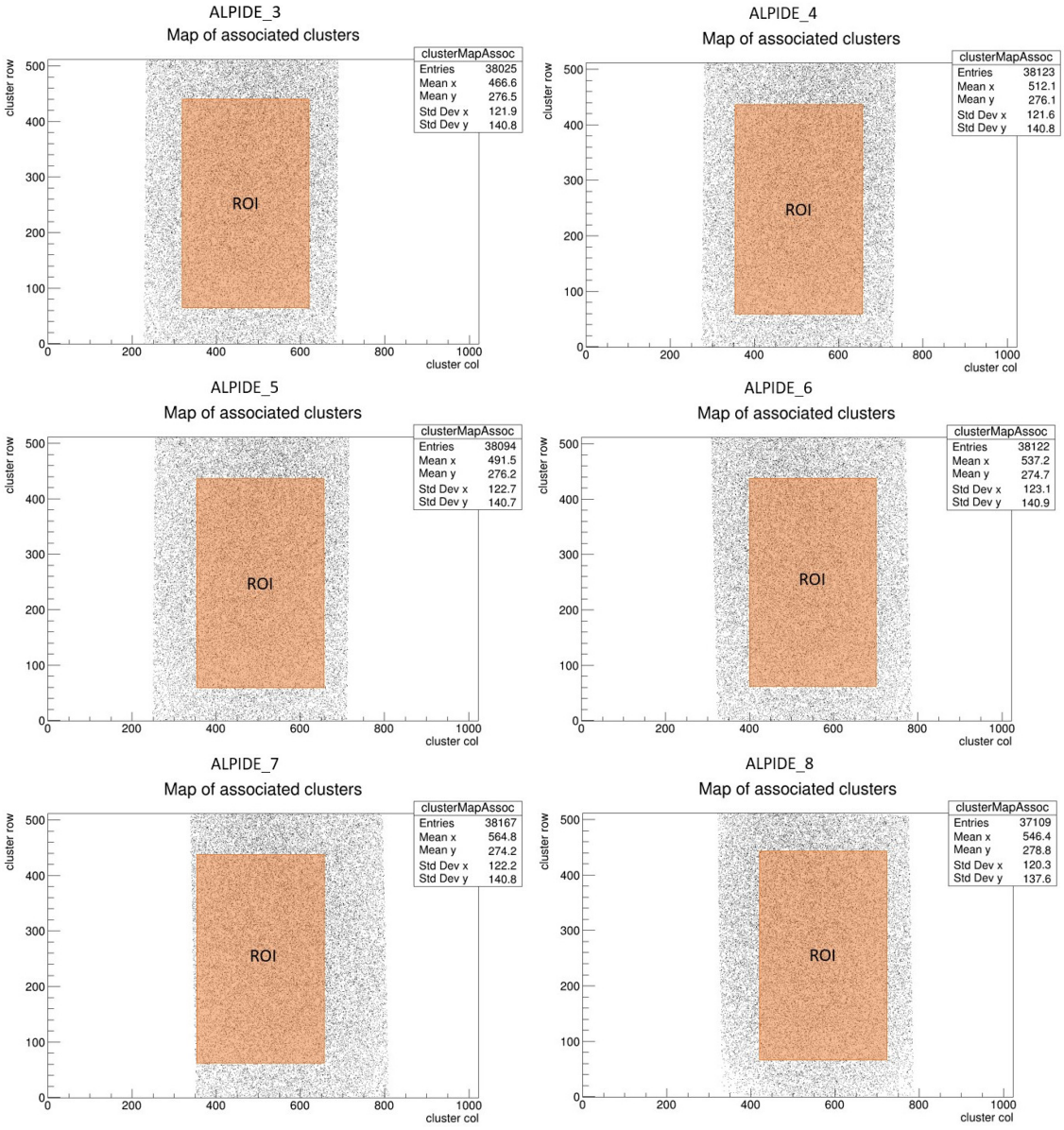


Figure 3.34: Spatial distributions, in local coordinates, of the associated clusters on the DUTs in the first setup used during the July test-beam. The areas defined by the points represent a "natural" ROI given by the limits of the intersecting regions of the corresponding DUT with the telescope planes, due to the 90° rotation of the DUTs with respect to the reference detectors. The shaded areas are the regions of interest defined for the analysis.

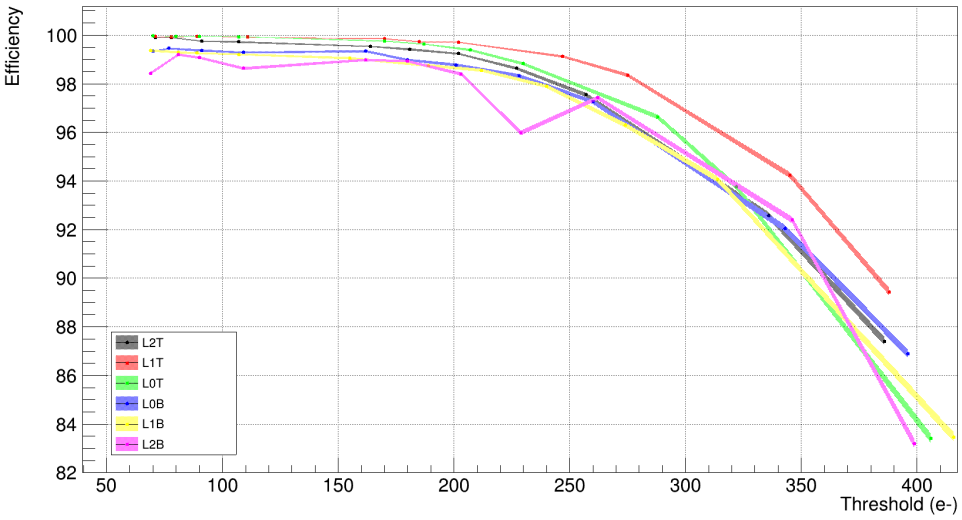


Figure 3.35: Sensor efficiencies for different runs featuring different threshold values calculated for all the DUTs in the first setup used during the July test-beam.

Spatial resolution The track residual distribution for a bent DUT in x- and y-direction is well described by a Gaussian centered at zero proving that a good alignment has been achieved, as shown in figure 3.28. Performing a Gaussian fit to the distributions, the standard deviation σ_{res} and its uncertainty can be extracted for both directions. Doing this for several runs featuring different working points of the DUT, the qualitative dependency of the position resolution on the sensor threshold setting can be evaluated. These quantities have been obtained for all the DUTs and are illustrated in figure 3.36. For all cases there is a visible minimum in the standard deviation of the residual distributions indicating the working point with the optimal position resolution, reached at a threshold value between $100e^-$ and $200e^-$. For the set of runs we have used in this study, the threshold of the DUTs was increased by decreasing their VCASN. As was presented in the section 3.4 and illustrated in figure 3.20 (right), for a certain threshold value, there is a minimum in the noise level; this behavior seems to be resembled by the trends of spatial resolution, which suggests that the optimum in position resolution is directly related to the noise level of the respective DUT. It may happen that the position resolution and thus σ_{res} worsens with an increased noise level. A higher noise level increases the probability of a pixel to fire. Consequently, the probability that a cluster resulting from noise is associated to the track is higher. Since the noise is not correlated with the track intercept, the residual tends to be larger. Over the full pixel matrix and a full run, this leads to a broader residual distribution and

therefore a worse position resolution. This explains the increasing trend of σ_{res} towards lower threshold values due to an increased noise level. On the other hand, the increase of σ_{res} towards higher threshold values can be understood as follows. As the cluster size generally decreases with increasing in-pixel threshold, the precision of the cluster position also decreases and this, in turn, leads to a broader residual distribution.

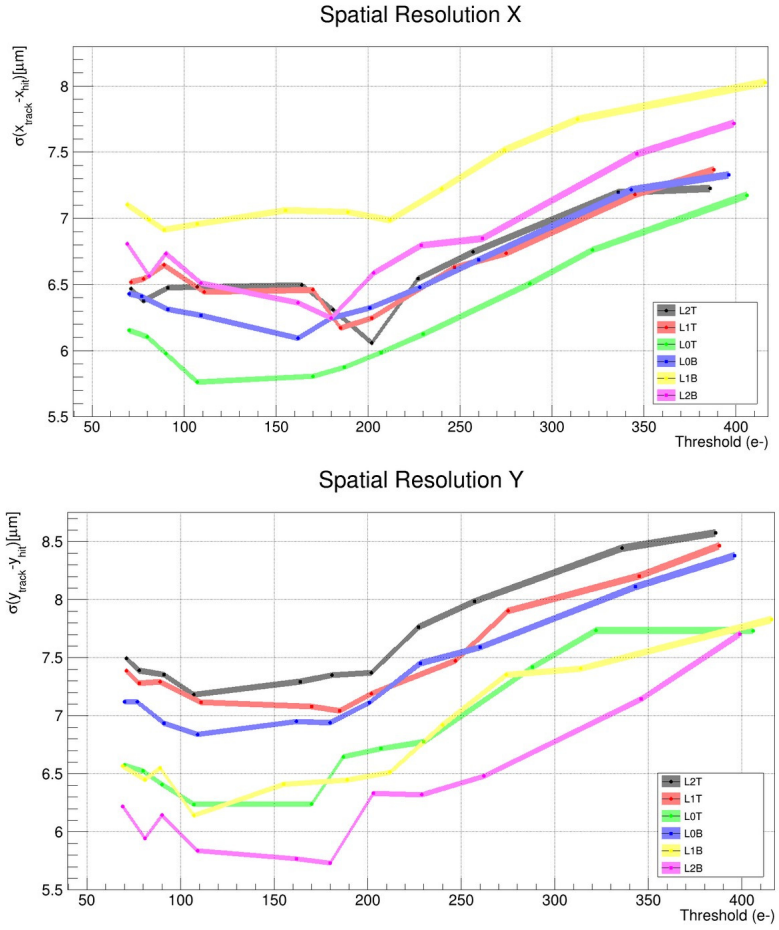


Figure 3.36: Width of the residual distribution in x (top) and y (bottom) vs runs at different thresholds calculated for all the DUTs in the first setup used during the July test-beam.

Telescope with micro-ITS3 and Cu target

The data collected with the telescope featuring the micro-ITS3 and the Cu target were foremost analyzed using the Corryvreckan reconstruction chain. The

principal interest was to make use of the data processing and alignment functionalities of the software, in order to perform a study of the tracks produced in the forward layers[5] by the particles of the beam, which undergo an inelastic interaction in the Cu target. The aim of this study is to validate the operation and test the tracking capability of a detector system with a configuration analogue to the ITS3 system.

Among the products of the inelastic interaction in the target there can be particles of lower momenta for which the multiple scattering interactions in the material is not negligible. Energetic secondary electrons ejected in ionization processes are generally referred to as δ-rays. Collisions with electrons dominate the energy loss of charged hadrons up to energies of hundreds of GeV. High-energy hadrons are also subject to inelastic nuclear collisions, which play a major role in the energy dissipation in a material if the hadron energy is of the order of GeV or higher. However, nuclear collisions are much less frequent than Coulomb interactions and high-energy hadrons still lose some energy in Coulomb interactions before being subject to a nuclear collision.

On the other hand, Coulomb interactions with nuclei dominate the angular deflection of charged particles in a material. The specific (i.e. per cm) energy loss in collisions with nuclei is generally much smaller than the electronic energy loss, except for low-energy heavy projectiles (typically ions in the keV/u energy range). The energy loss is called non-ionizing energy loss and can lead to the displacement of atoms from the lattice if the energy transfer is higher than a certain material-dependent damage threshold, which can also depend on the particle direction in crystalline materials.

Nuclear collisions are considered inelastic if new particles are produced or if the internal structure of the target nucleus or the projectile is changed. Some general features of nuclear collisions can be inferred from nucleon–nucleon cross-sections [91]. The total cross-sections are the sum of the elastic and inelastic ones. Above a few GeV, total cross-sections are relatively energy independent up to the highest energies reached in today's accelerators and over a wide energy range, inelastic cross-sections are much larger than elastic ones.

In the initial fast stage of a hadron–nucleus interaction ($10^{-22}\ s$), the incident hadron interacts with the nucleons and new particles can be produced (in particular, pions) if the projectile energy is high enough (in particular for energies $\geq\ GeV$).

The hadrons (protons, neutrons, pions) can form intranuclear cascades and the more energetic particles can leave the nucleus. These fast particles are mainly forwards-directed and their multiplicity scales approximately with the logarithm of the incident particle energy. The other particles can deposit their energy in the nucleus and leave it in an excited state.

After going through a pre-compound phase, the nucleus is subject to a slow de-excitation process ($10^{-16}\ s$). In this evaporation phase, lower energy (MeV) neutrons, protons, or light fragments (alpha particles) are emitted isotropically.

During the execution of Corryvreckan the module named "JSONWriter"

[5]Layers after the target following the beam direction.

was used. This reads the objects containing information of pixels, clusters and tracks from the clipboard and writes the details, event by event, of the objects selected by the user to a file of format .json. Consequently, for this analysis the global positions of the cluster centers were retrieved from the Json file. The figure 3.37 shows a sketch of the test-beam geometry (not in scale), where the green planes represent the ideal positions of the ALPIDEs composing the telescope and the micro-ITS3, the yellow volume illustrates the target location and the black points are the positions of cluster centers for all the events after the alignment, extracted from the Json file produced by the software.

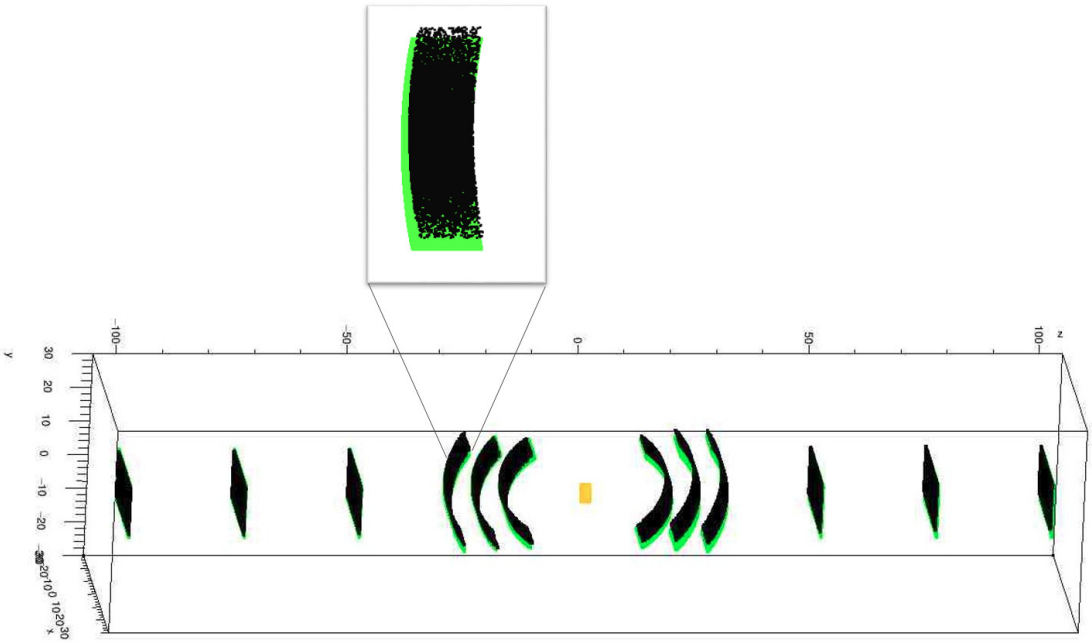


Figure 3.37: Schematic of the July 2021 test-beam setup geometry (not in scale). The green planes represent the ideal positions of the ALPIDEs composing the telescope and the micro-ITS3, the yellow volume illustrates the target location and the black points are the positions of cluster centers for all the events after the alignment, read from the Json file produced by the Corryvreckan module named "JSONWriter".

For the scope of this study a precise selection of data was applied to assure the purity of the samples. In order to classify the events featuring an interaction of the beam on the target with production of secondary particles, several cuts were imposed to the data, following the criteria described in the next. As necessary conditions for a "good" event, it should contain:

1. At least one cluster per layer.
2. No more than one cluster on each of the six layers before the target (upstream). A single point on each upstream ALPIDE would describe the crossing of a beam particle through the chips before a potential interaction with copper, thereby defining a "beam" track.
3. At least two clusters on the six layers after the target (downstream). This would indicate the eventual production of secondary particles.
4. Finally, the positions of the clusters on the upstream layers have to be inside a rectangular area defined in xy-coordinates for each detector plane. The region is designated to be safely compatible with the dimensions and the position of the target in the $x - y$ plane, so that the events featuring interactions of the beam with materials different from the copper, e.g. screws and holders, are rejected.

A total of 28 runs, with approximately 300k events each, were analyzed. After applying the selection criteria already described to the corresponding Json files, a total of 9957 events were classified as "good" over all the data sets. In figure 3.38 (top) one of the resulting events is displayed, while the figure 3.38 (bottom) shows a plot of all the events selected in a representative run, where it can be observed that the points on the upstream layers, i.e. from ALPIDE_0 to ALPIDE_5, draw a "shadow" of the target, given the superposition of the imposed conditions, including the geometrical cut.

After the pre-selection a track finding algorithm was performed to identify the track of the beam and the potential tracks that are coming out from the target after the beam interaction. All the potential tracks identified are then fitted to define their best track parameter by running a track fitting algorithm.

Track finding, also known as pattern recognition, refers to finding the hits that belong to all potential track candidates. Since there is no magnetic field the tracks follow straight line paths, apart the deviation due to the multiple scattering (MS) in the material crossed by particles, a straight line track finding algorithm has been developed [92].

The track candidates are fitted with a straight line track model by using the Global Chi-square (χ^2) fitting method [93]. Charged particles travelling inside a material, specifically the silicon of the downstream sensor layers in the test-beam setup, are subject to sizeable Coulomb interactions with electrons and nuclei. As a consequence, the particles slow down and are deflected with respect to their initial direction [94]. So that, despite the low material budget of the detector planes (0.05 % X_0), due to the tracks topology, the multiple scattering contribution can be not negligible for all tracks. To take into account the multiple scattering effect, a Global χ^2 that includes a proper treatment of the detector multiple scattering has been applied to the tracks produced by the particles in the downstream layers, i.e. from ALPIDE 6 to ALPIDE 11. Therefore, the χ^2 of the downstream tracks is minimized considering the covariance matrix of the detector spatial resolution (V_{SR}) and that of the multiple scattering (V_{MS}) contribution by using the following expression:

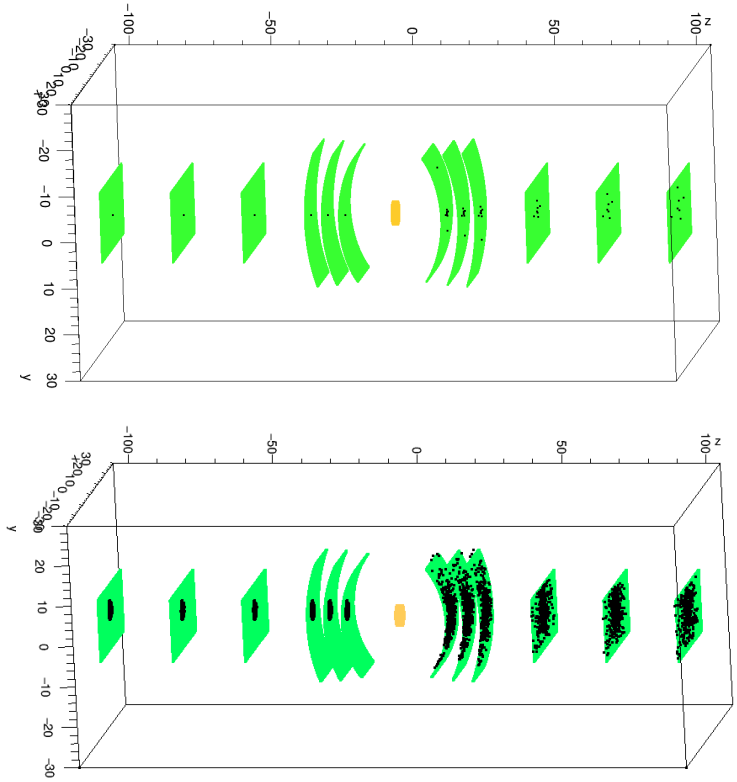


Figure 3.38: (top) Example of one event featuring an interaction of the beam on the target with production of secondary particles. (bottom) Display of all the events following the same selection criteria in one run.

$$\chi^2 = R^T (V_{SR} + V_{MS})^{-1} R \tag{3.8}$$

where R is the residual matrix for all the detector planes.

For the high-momentum track (beam), the multiple scattering is negligible and the tracking uncertainty is determined mainly by the intrinsic resolution of the layers (5 μm). So that χ^2 can be approximated by the following expression:

$$\chi^2 = R^T V_{SR}^{-1} R = \sum_i \frac{d_{x_i}^2}{\sigma_{x_i}^2} + \frac{d_{y_i}^2}{\sigma_{y_i}^2} \tag{3.9}$$

Global χ^2 fit using the equation 3.8 requires two covariance n×n matrices for n hit points: (V_{SR}) is a diagonal matrix with the entries representing the square of detector spatial resolutions at each detector plane and (V_{MS}) is constructed for each track based on momentum hypothesis using a fast simulation method. In this analysis, the equation 3.9 is used to fit the *beam* (they have

fixed momentum of about 120 GeV/c) while for the produced *tracks*, the equation 3.8 is adopted. To construct the covariance for the multiple scattering, the knowledge of the momentum of the produced tracks is required. Therefore, a PYTHIA8 simulation [95], in which the collision of a proton (with momentum of 120 GeV/c) with a Cu nucleus is considered at the parton level by using the Angantyr model [96], was performed to estimate the average momentum of a track as a function of pseudo-rapidity (η) (See figure 3.39). Then, for the track fitting, the η of the track candidate is evaluate by using the first and last hit information then the average momentum hypothesis, as estimated with the PYTHIA8 simulation is assumed for the track.

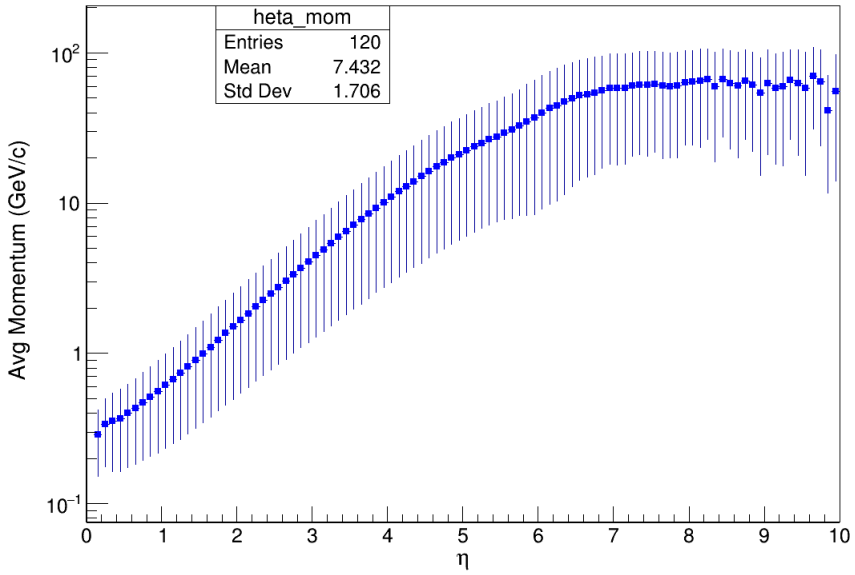


Figure 3.39: Track average momentum as a function of pseudo-rapidity (η) [92].

In figure 3.40 the distribution of χ^2/ndf for the reconstructed beam and tracks are shown. The distribution for the tracks is expected to be broader than that of the beam since about half of the tracks have less than 6 hits and they have a wide moment range.

Figure 3.41 shows a display of one of the selected events with reconstructed and fitted tracks.

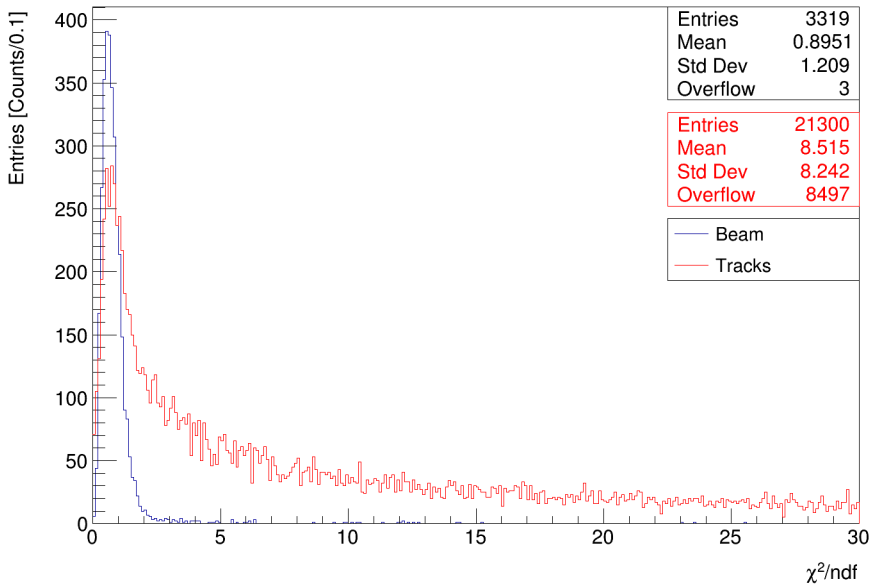


Figure 3.40: χ^2/ndf distribution for the reconstructed beam and tracks using the Global fit method [92].

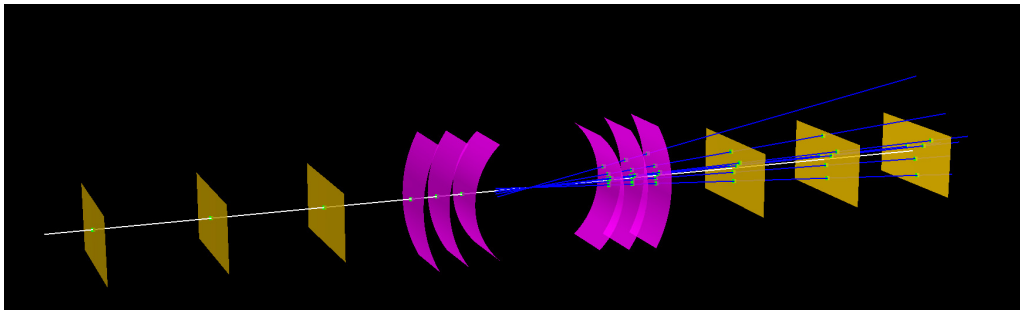


Figure 3.41: Display of one event with interaction of the beam in the target, where the tracks have been reconstructed and fitted. The white line represents the beam track and the blue ones are the tracks produced by secondary particles after the interaction with copper.

To check the reconstruction strategies and the final results a simulation of the full setup, using a Toy Monte Carlo simulation tool (Toy MC), was developed. It uses the generated interaction events, the output of the PYTHIA8 simulation and adds the main experimental features [92]. The same reconstruction procedure and code employed for the data were applied to the simulated events.

After that all the tracks in an event are reconstructed, their distance of closest approach (DCA) are evaluated. The DCA is the minimum distance in space between the beam and a downstream track straight lines. In particular, the average resolution of the distance of closest approach (DCA) can be measured, which plays an important role in the reconstruction of secondaries decaying close to the interaction vertex.

Figure 3.42 shows the distribution of DCA between the reconstructed beam and downstream tracks for the measured (top) and the simulated (bottom) data. The distributions are fitted with a two Gaussian distributions function to evaluate the resolution of the central part and in the tails individually. Comparing with simulations, the average mean of the DCA distribution for the measured data is higher by one order of magnitude, but still close to zero. The standard deviation is more than two times higher, given the group of tracks with higher measured values of DCA, which contribute to the tail on the distribution.

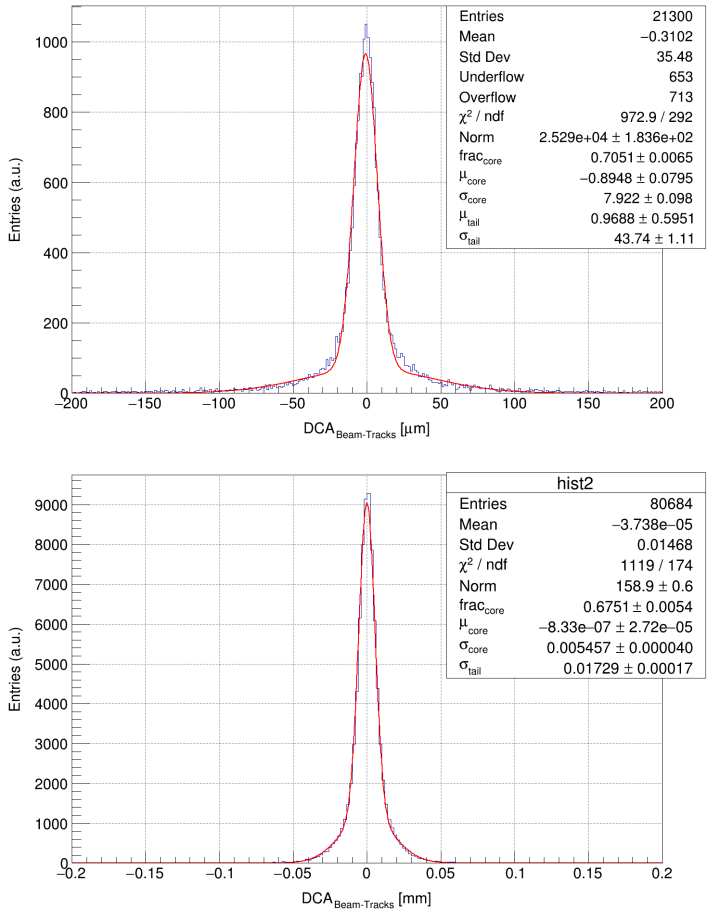


Figure 3.42: Distribuons of the DCA between the reconstructed beam and the tracks for the measured (top) and the simulated (bottom) data. The distributions are fitted with the two-component Gaussian distribution [92]

Chapter 4

Wafer-scale sensor. The "super-ALPIDE" project

As introduced previously in section 1.3, the ITS project is developing a new vertex tracker for the ALICE experiment, to be installed before the LHC Run 4. It will be based on thin wafer-scale Monolithic Active Pixel Sensors bent in cylindrical shape around the beam pipe. The innovative idea of the ITS3 upgrade is to make use of the stitching technology to produce a single sensor with such dimensions, where the power distribution is managed internally, confining the interconnections to the outside world to the wafer-scale sensor edge.

The largest field of view that is used in the photolithographic steps in CMOS manufacturing, defines the reticle size, normally of a few centimeters in both directions. Stitching is a technology that allows the fabrication of a sensor that is larger than the field of view of the lithographic equipment. In this technology, reticles which fit into the field of view are placed on the wafer with high precision, achieving a tiny but well defined overlap. In this way, wafer-scale sensors can be manufactured.

There are multiple possible approaches in developing the wafer-scale chip and one possibility is schematically represented in figure 4.1. The sensor size is about $15\,mm \times 280\,mm$, which corresponds to the full length of a layer (1.2). The columns run along the short side of the circuit, and have approximately the same length as in the ALPIDE chip. The hits are read out from the matrix as in the ALPIDE chip, that is with hit-driven circuitry based on a priority-encoder addressing scheme and are transmitted from the bottom of the columns to the periphery via the data bus running along the long side of the sensor (Figure 1.9). The control logic, digital-to-analogue converters, and data buffers and links are located at the short edge of the chip. Two types of reticles would suffice to produce a wafer-scale sensor in the proposed design; one for the chip periphery and one for the pixel matrix and data bus, with the latter repeated multiple times in the data bus direction. By reproducing the whole structure multiple times along the short side of the circuit (in the column direction), the

entire half-layers can be produced from a single wafer, as shown in figure 4.1 The wafer map would be optimised such that different layers can be diced; e.g. Layers 0, 1 and 2 would contain 3, 4 and 5 instances of the circuit, respectively, assuming a circuit width of about 18 mm.

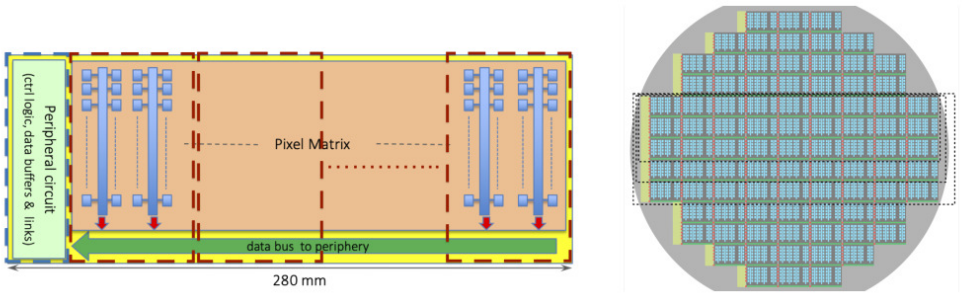


Figure 4.1: (left) The schematic representation of the proposed circuit as the main building block of the wafer-scale chip. The dashed lines represent the two reticle types: periphery (blue) and matrix (red). (right) The wafer map optimised for dicing different layers (dashed black lines) containing multiple instances of the circuit on the left. The rest of the wafer is filled with chips with shorter matrices.

As mentioned in section 1.3.3, the first choice of the technology for the implementation of the wafer-scale sensor is the Tower-Jazz 65 *nm* Imaging Sensor Chip process. TowerJazz has stitching experience with this process and offers production on 300 *mm* wafers. Furthermore, the 65 *nm* process, w.r.t. the 180 *nm* process used in ITS2, i.e. ALPIDE, benefits of an improved radiation tolerance and smaller feature size thus allowing the implementation of smaller pixels. The first submission containing test structures for process exploration have been finalized in summer 2021. Some prototypes have been fully characterized in laboratory and under beams [97], while others are currently under study, as will be explained in chapter 5. Next iteration in the sensor design will include first implementation of the stitching [98]. In parallel with sensor design and characterization, development of support structures, flex for powering and communication, procedures and tools for the assembly and integration of the layers must be carried out. This is mainly done by using blank silicon, sized to the dimensions of the final detector, as dummy sensor. The usage of such bare object does not allow to verify consequences, e.g. of the bending procedure, on the performance and electrical behavior of the sensor. The "super-ALPIDE" project has the goal to assembly a detector with geometry similar to one half-layer of the ITS3, using a working sensor with large dimensions made of a matrix of not diced ALPIDE chips. This program allows to explore the design of a flex and investigate its interconnection to the sensor. In the following section, the main idea of the super-ALPIDE design, including the mechanical structures for readout and support will be presented. Additionally, techniques

and tools employed to perform sensor bending, components assembly and wired interconnections will be described in detail.

4.1 The super-ALPIDE design

The principal components of the super-ALPIDE detector are sketched in figure 4.2. The sensor bent as a half-cylinder is represented at the bottom right (yellow), two flexes needed to access to each single sensor in the matrix can be seen in orange and green and the main support structures are drawn in yellow and black.

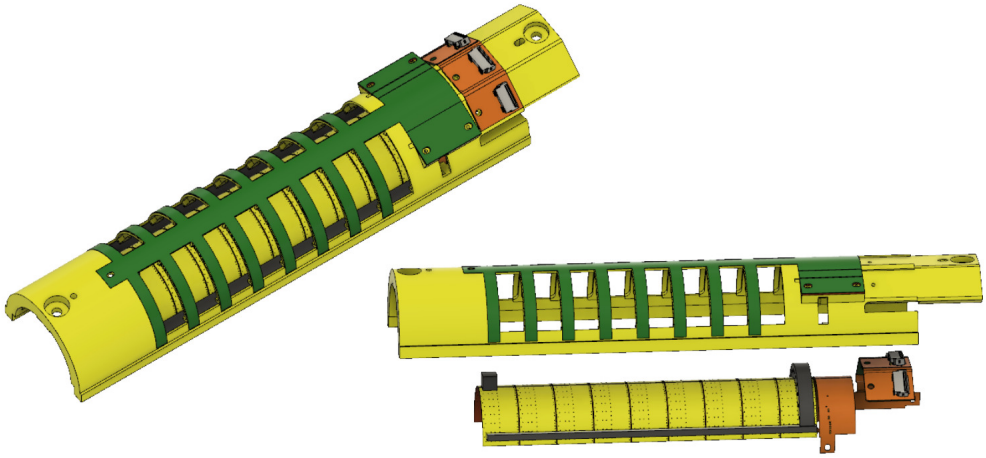


Figure 4.2: Design of the super-ALPIDE detector. Visible components are: sensor (in yellow), support structures (in yellow and black), edge-FPC (in orange) and exo-FPC (in green).

A sensor for the innermost layer of the ITS3 is expected to be $280 \times 56.5\ mm^2$. The super-ALPIDE is a matrix of 9×2 ALPIDE sensors[1] for a total area of $138 \times 60.1\ mm^2$ as shown in figure 4.3 (left). The object is a bit larger on the shorter direction and more than one half smaller in the longest direction in comparison with the foreseen dimensions. At a curvature radius of 18 mm, it is large enough to cover the 180^o of the half-layer. Interfaces to the external world of the ALPIDE sensor include: analog, digital and back biases, control, clock and data transmitters. Two types of pads with differing geometries are employed in the ALPIDE chip: Type A, with opening size of 88 μm, are distributed along one of the longest chip edge and Type B, with opening size of 290 μm, are distributed over the sensitive matrix. The nets accessible through pads of type B are also found in the pads of type A, i.e. both groups are equivalent and include all the interfaces. The choice of the pads set to be used depends on the

[1] ALPIDE chip measures $15 \times 30\ mm^2$ and contains 512×1024 sensitive pixels.

application of the sensor. As it will be described in sections 4.2.2 and 4.2.4, in the super-ALPIDE project type A pads are used for wire-bonding.

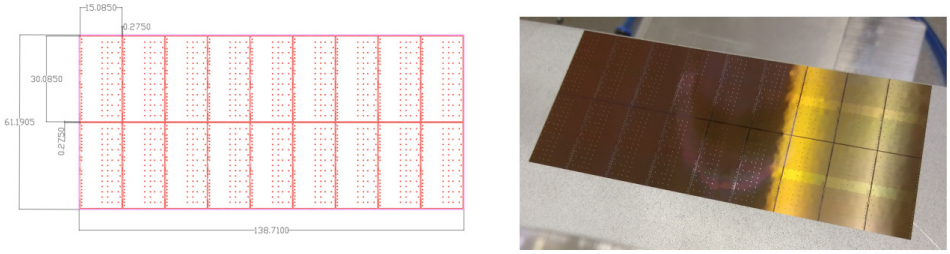


Figure 4.3: (left) Matrix of 9 × 2 ALPIDE sensors in one continuous silicon piece. (right) Picture of a functional super-ALPIDE sensor.

4.1.1 Support structures

Two groups of support structures must be highlighted: the carbon foam and the exoskeleton, as drawn in figure 4.2. Carbon foam support structures are meant to be very similar to the pieces that will be used to connect and keep in place two subsequent half-layers in the final detector. The material has been chosen because of the low material budget density (ERG $\rho = 0.06\ kg/dm^3$, Allcomp $\rho = 0.2 - 0.26\ kg/dm^3$) and, in some version, very low thermal conductivity (ERG $K \geq 17\ W/m \cdot K$, Allcomp $K = 0.033\ W/m \cdot K$). It will be used, not only as support structures but as part of the cooling system, through thermal dissipation. The three main carbon foam components are shown in figure 4.4 (top). The half-ring and the wedge are used to constrain the sensor in cylindrical shape along the curved direction, while the long-bars act in the longitudinal direction (Figure 4.4, bottom). In the design of the final detector, the wedge has been recently replaced by a second half-ring to reduce curvature deviations. As in the case of the final detector, the thickness of these objects is 6 mm, distance between two consecutive layers. The length of the long-bars has been adapted to the shorter length of the sensor. In the case of the super-ALPIDE, all the structures have been produced in ERG carbon foam, due to the material availability and the limited requirements on thermal dissipation by conductivity. A carbon fleece is adopted in order to smooth the irregular surface of the carbon foam, going to be glued to the sensor. The carbon veil material has an area weight of 8 g/m^2 and can be easily pressed, reaching 120 μm at 10 kPa of pressure. This allows to better control the amount of glue, reducing in turn the material budget. CT scan measurements estimated a total thickness of the adhesive layer of the order of 300 μm. Araldite2011 is used for gluing the fleece to the carbon foam and to the silicon.

Interconnections of the sensors in the super-ALPIDE matrix to the external world are made by using flexible printed circuits (FPC). The first two sensors

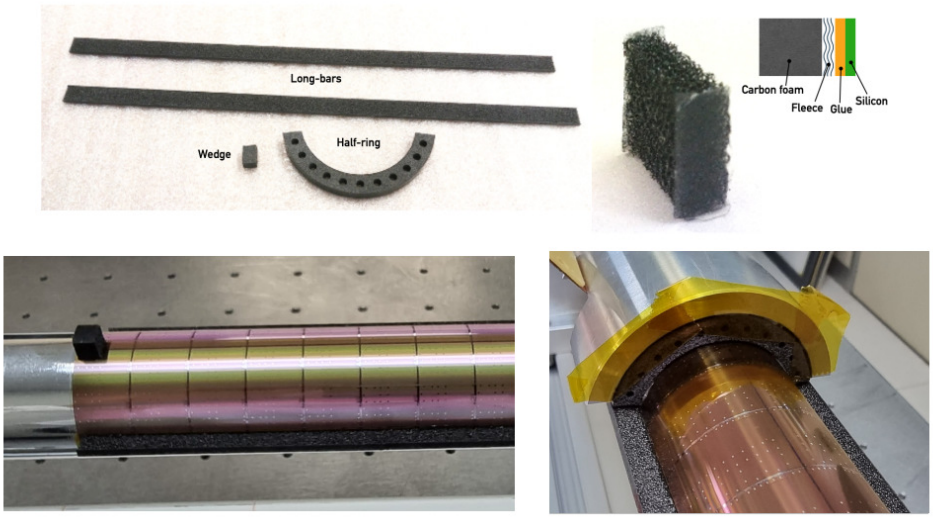


Figure 4.4: (top) Support structures in carbon foam and gluing interface to the silicon. (bottom) Pictures of the carbon foam pieces (wedge, half-ring and long-bars) glued to the surface of a bent super-ALPIDE sensor.

in the matrix (having the type A pads line on the external side) are connected to one FPC, called edge-FPC, mimicking the configuration that will be implemented in the final detector (Figure 4.5, left). The rest of the sensors in the matrix require a second FPC to be connected to, that allows wire bonding to be as short as possible. The solution is to place the FPC over the sensor matrix, supported by a structure of cylindrical shape and having windows to access the sensor below (Figure 4.5, right). This second support structure, called exoskeleton, is 3D printed in Accura25 SLA material, fabricated with a dimensional precision of the order of 100 μm. The second FPC, called exo-FPC, is glued on-top of this structure as can be observed in figure 4.5 (right), whose thickness is 1 mm, ending up to a total vertical distance between the sensor surface and the FPC surface of 7 mm. The structure is rigid enough to grant support during the bonding operation. This structure rests over the carbon foam support structures and is glued (Araldite2011) to them, acting as an external shell. Similarly in the final detector there will be an external Cylindrical Structural Shell, made in carbon fiber, that will encapsulate the full detector. However, the exoskeleton will not be part of the final detector, i.e. is a temporary solution to support the connectors for the exo-FPC and the edge-FPC. It is used in this initial verification stage to make easy the access for testing independently all the sensors in the matrix. Two narrow windows are available on the sides, in correspondence to the region where the super-ALPIDE faces the edge-FPC, allowing visual inspection of the region, otherwise covered by the exoskeleton.

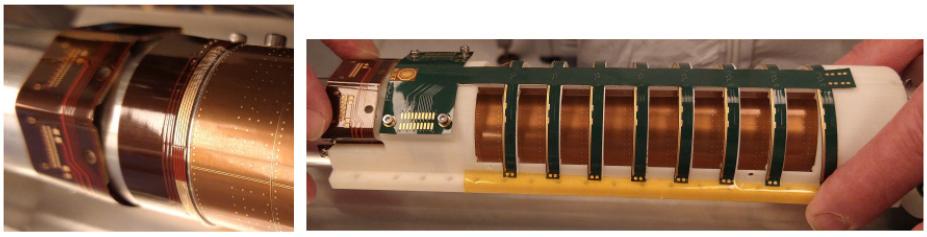


Figure 4.5: (left) Edge-FPC wired bonded to the type-A pads ring on the external side of the first two sensors in the matrix of a bent super-ALPIDE. (right) Exo-FPC (green) glued on top of the exoskeleton (white), which is placed over the sensor matrix covering its wire bond connections to the edge-FPC (brown).

4.1.2 Flex Printed Circuits

As mentioned above, two FPCs are connected to the sensor matrix in order to provide an homogeneous power distribution and allow configuration and readout. The edge-FPC, whose schematic is shown in figure 4.6 (left), is a precursor prototype of the FPC that will serve the sensor in the final detector. The main design features explored in this prototype are: (i) presence of a region in the FPC bendable at the same curvature of the sensor, combined with a flat region housing the connectors as shown in figure 4.6 (right); (ii) possibility to explore wire-bonding between two bent objects. As can be seen in figure 4.7, the bendable region of this FPC will be aligned to the sensor edge at a distance of the order of 500 μm. Soldering surfaces on the flex are strips and pads lying along the bending direction, close to the edge facing the sensor as can be observed in figure 4.5 (left). There are five strips, common to the two sensors, for powering (analog, digital and back-bias). Two common double differential lines used for control and clock and two independent double differential lines designed for data transmission expose small pads in correspondence to the equivalent pads in the sensor. Powering lines are equipped with decoupling capacitors and double differential lines for clock/control are equipped with terminations, to avoid signal reflection.

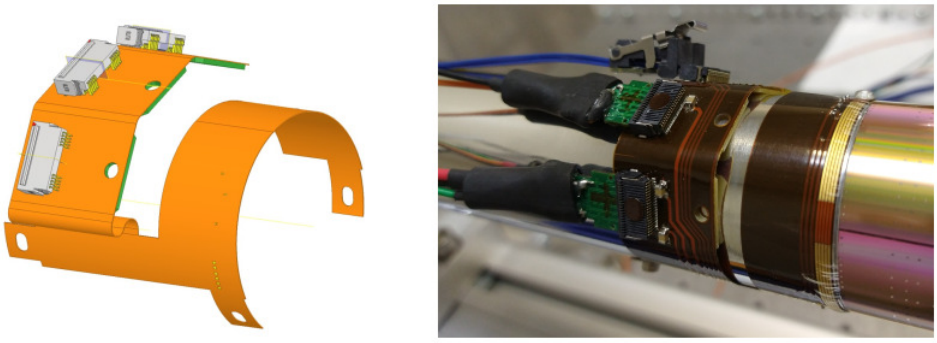


Figure 4.6: (left) Design of the edge-FPC. (right) Picture of the edge-FPC connected through the soldering surfaces along the bending direction to the sensor edge.

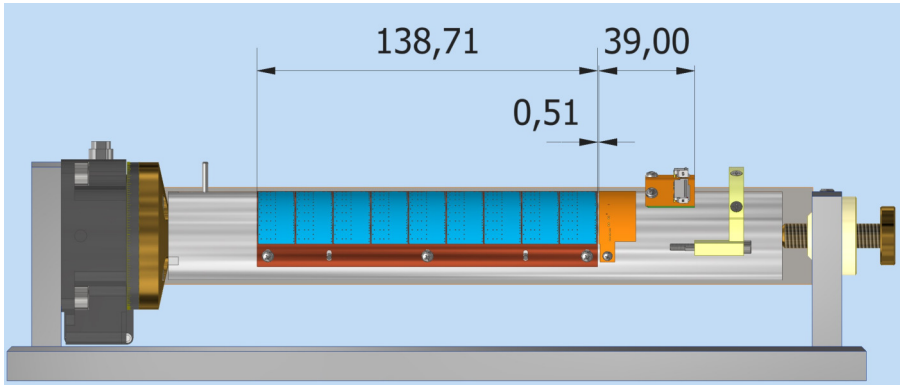


Figure 4.7: Schematic (dimensions are in mm, not in scale) of the positioning of the sensor matrix (blue) and the edge-FPC (orange) bent at the same curvature. They are aligned at a distance of the order of 500 μm.

The exo-FPC, even though is not foreseen for the final design, is needed in the test phase to access the remaining sensors in the matrix (Figure 4.8, left). Through a back-bone all the lines are routed from the connector to eight rectangular branches where soldering strips and pads are located. Lines pattern is symmetric on the two sides of the backbone. Powering lines are independent for each sensor, while ground is common. Control and clock lines are shared between all sensors on the same side. Decoupling capacitors are placed at the end of each branch, while termination resistance are placed at the end of the back-bone for the control and clock lines. The bonding strips, placed few hundreds μm from one edge of each branch, are as long as the below opening in the exoskeleton is large (Figure 4.8, right). Both FPCs are double layer printed

circuits, with copper metal lines and a total thickness of 170 μm.

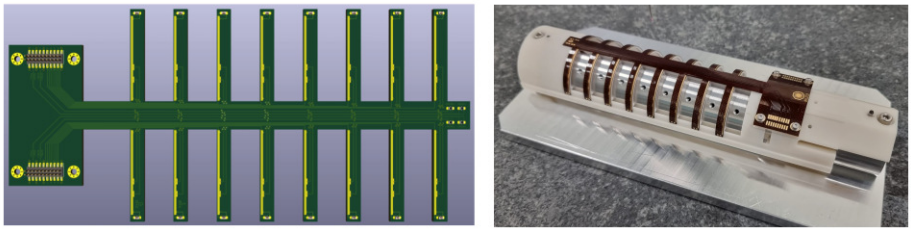


Figure 4.8: (left) Design of the exo-FPC. (right) Picture of the exo-FPC (brown) attached to the exoskeleton (white).

4.2 Detector assembly techniques

In order to explore as much as possible the techniques to be used for the assembly of the final detector, the envisaged tools (just scaled in longitudinal dimension) and procedures are used for the realization of this project.

4.2.1 Silicon bending technique and tools

Bending technique foresees the usage of a rotating cylindrical mandrel around which the large silicon sensor is wrapped, guided by a transparent Mylar foil, kept in tension parallel to the mandrel longitudinal axis. Once bent, as observed in figure 4.9, the sensor is held in this position by two strips of adhesive tape placed along its long edges, i.e. the two sides parallel to the mandrel rotating axis. The tools needed for handling the large dimensions sensor, alignment procedures and placing on the mandrel are shown in figure 4.10.

The main steps adopted for the bending procedure are the following: (a) positioning of one strip of adhesive tape on the mandrel, along longitudinal direction, only partially attached to the surface, leaving the remaining part free to be later attached to the sensor; (b) placement of one strip of adhesive tape on one of the long side of the sensor; (c) alignment of the sensor to the mandrel, from the side without adhesive tape, in correspondence to the already positioned adhesive tape strip; (d) gluing of the adhesive tape to the naked side of the sensor, leaving it anchored on one side to the mandrel and laying tangent to the cylinder surface; (e) actual bending of the sensor, rotating the mandrel in a controlled way through the step motor, and pressing the sensor to the mandrel using a Mylar foil kept in tension by a weight attached through a set of two pulleys, as can be observed in the right picture of figure 4.10; (f) when reaching the bending of the second side of the sensor the Mylar foil press the adhesive tape and completely constrains the sensor to the mandrel in cylindrical shape. Handling of the sensor in all the steps described is done using the tool shown in figure 4.10 (left). This consists in two vacuum plates made of a porous

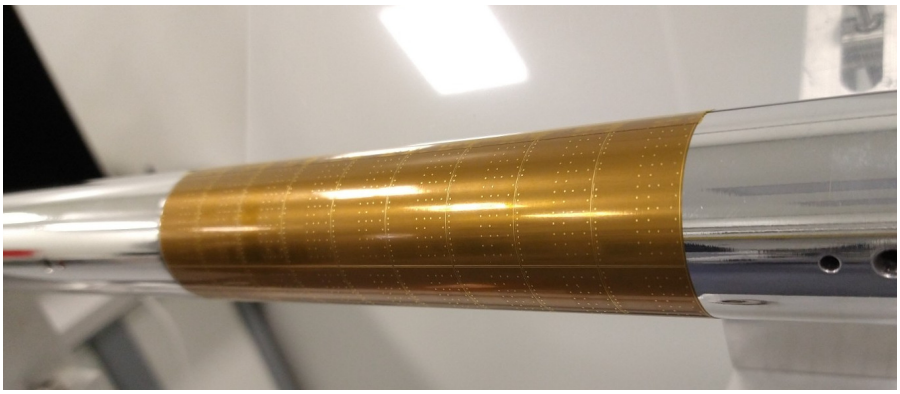


Figure 4.9: Super-ALPIDE sensor bent around the mandrel of 18 mm radius. It is held in position by two strips of adhesive tape placed along its lateral long edges.

material, one fixed and one rotating and equipped with handles of micrometric precision. The mandrel was fabricated with reference grooves in longitudinal and transverse directions, which facilitates the verification of the alignment, done by means of a digital microscope.

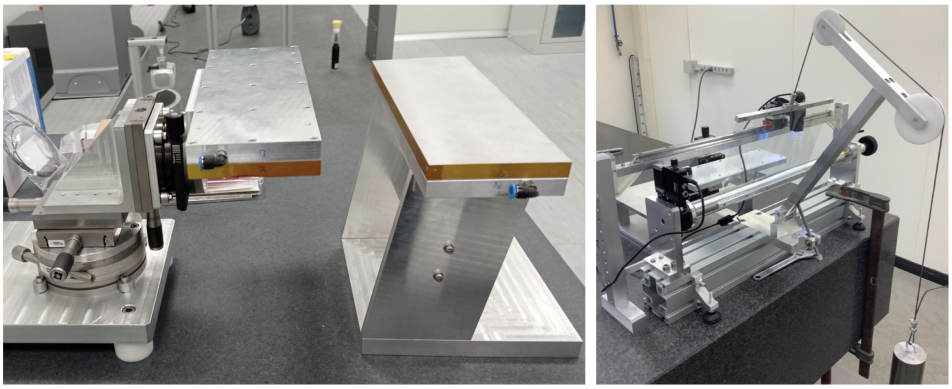


Figure 4.10: (Left) Vacuum plates for handling large dimensions sensors. (Right) Set of tools dedicated to bending. The cylindrical mandrel is attached to a step motor for rotating; Mylar foil is kept in tension parallel to the mandrel longitudinal axis by a weight attached through pulleys; a small camera is positioned on top of the setup to control the sensor alignment.

To test the bending procedure, five attempts using leftover pieces from silicon wafer cut with irregular shape, so called "half-moon", have been realized. Three of those trials were successful, i.e. the pieces of silicon were bent without breaking (See figure 4.11). From the other two, one failed because a too long

piece of adhesive tape was placed on the curved side and in the other case a defect was found on the mandrel surface as showed in figure 4.12.

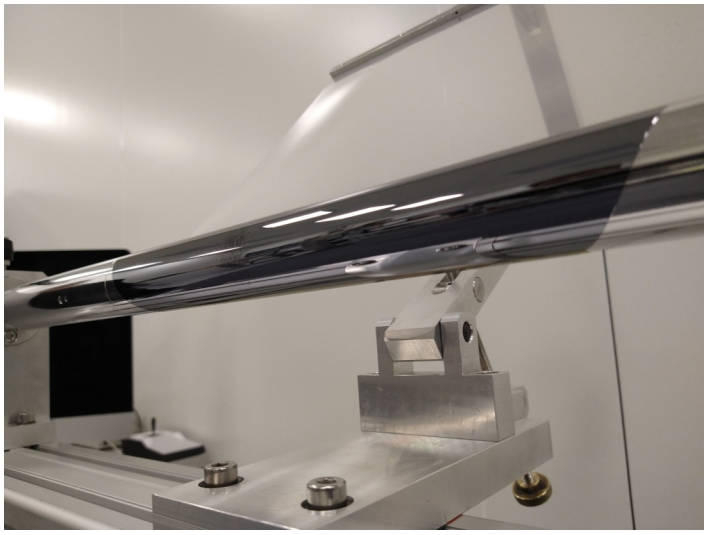


Figure 4.11: A piece from silicon wafer cut with irregular shape, so called "half-moon", bent around the cylindrical mandrel.

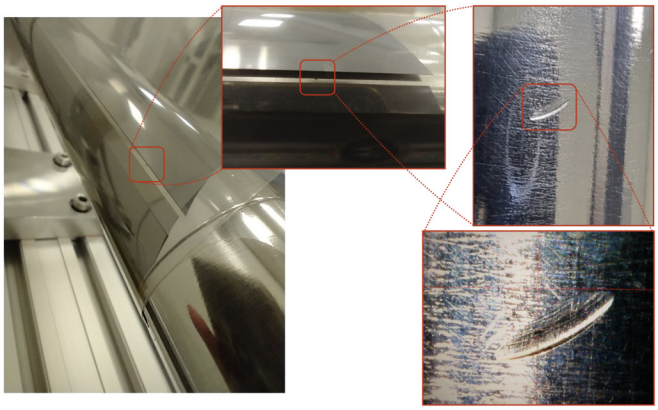


Figure 4.12: Surface deformity found on the mandrel surface during an attempt to bend a "half-moon" piece.

4.2.2 Edge-FPC to sensor interconnections

Next step in the assembly procedure is the interconnection of the edge-FPC to the sensor. Edge-FPC is kept in position around the mandrel by three screws. As anticipated, interconnection is done by wire-bonding. The region with sol-

dering strips on the FPC is large 5 mm, while the bonding pads on the ALPIDE sensor are few hundreds microns away from the border. FPC is positioned about 500 μm away from the border of the sensor. So, the maximum wire length is 6 mm. Wire-bonding is performed once the two objects are bent around the mandrel. The support structures for the mandrel used in the bending setup, have been designed also to allow its placement under the bonding machine as illustrated by the picture in figure 4.13. Despite the two objects are bent, the wire-bonding is actually performed between two points that seat in the same motion direction of the bonding machine head (longitudinal direction of the mandrel). This imply that the bonding machine is, to all effects, operating in standard condition. The following wire-bonding position is reached rotating the mandrel, below the machine head, by means of the step motor. As mentioned, two ALPIDE sensors are interconnected to this FPC, for a total of almost hundred wire-bonds.

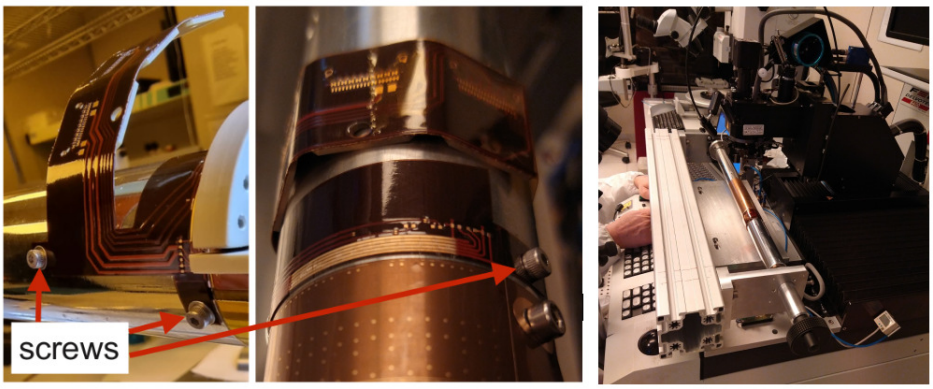


Figure 4.13: (left) Edge-FPC fixed to the mandrel by screws. (right) Mandrel and its support components placed under the bonding machine.

To reduce dangerous stress on the long wires, sensor and FPC are both glued to the exoskeleton, before the removal from the mandrel. In order to further protect the wires, encapsulation using a silicone elastomer material (Sylgard 186) is explored. This is a technique used in the assembly of other silicon tracker, like the CMS forward pixel detector [99]. Results of the bonding procedure between the two external sensor in the super-ALPIDE matrix and the edge-FPC are presented in the left picture of figure 4.5.

4.2.3 Support structures and exo-FPC gluing

At this point in the assembly, carbon foam must be placed over the sensor to provide rigidity to the structure. Two independent tools, made in aluminum, have been designed to perform the precise alignment of the carbon foam support structures and allow the curing of the glue in position; one is used for the long bars and the wedge (Figure 4.14, left) and the second for the half-ring (Figure

4.14, right). Precise positioning of the tools over the mandrel is granted with pins and screws. Araldite 2011 glue is used and its curing time is 24 hours.

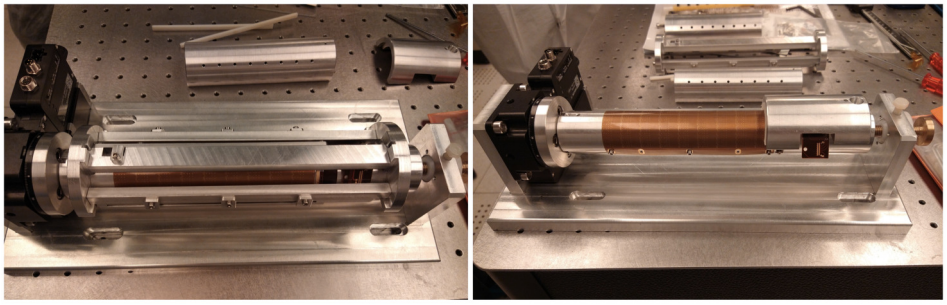


Figure 4.14: (Left) Tools designed to facilitate the correct alignment, during gluing and curing processes, of carbon foam long bars, wedge (left) and half-ring (right).

Three assembly attempts using blank silicon wafer of super-ALPIDE dimensions with metallic pads at the positions of those on the real detector, have been realized. Since one of them failed during the half-ring gluing, using the tool of figure 4.15 (left), the design of this element has been modified by including three screws to fix the half-ring to the silicon with higher precision as showed in the right picture of figure 4.15. The other two assemblies were completed successfully (See figure 4.21), using the new version of the tool; the first one used 3d printed support structures, while second one used carbon foam support structures, as will be done in the final object.

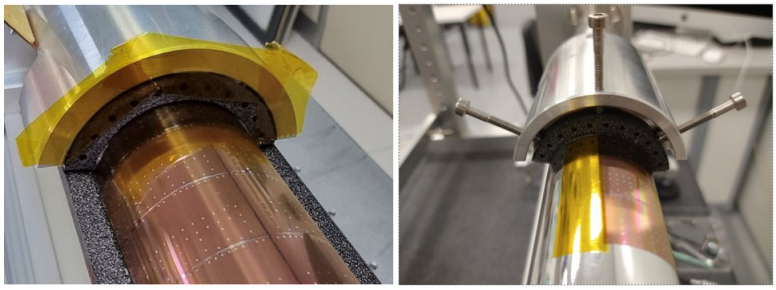


Figure 4.15: (left) First and (right) modified versions of the tool designed to align the half ring during the gluing and curing processes.

Parallel procedure is the gluing of the exo-FPC over the exoskeleton. A dedicated tool has been developed to grant needed alignment precision in this step. As depicted in figure 4.16 (left) and can be visualized in figure 4.16 (right), an aluminum half-cylinder is the basis where the exoskeleton is placed, avoiding any deformation of the structure during the procedure. A transparent cylindrical cap, made in plexiglass and fixed to the beneath cylinder with screws, is used

to press the FPC to the exoskeleton during the curing.

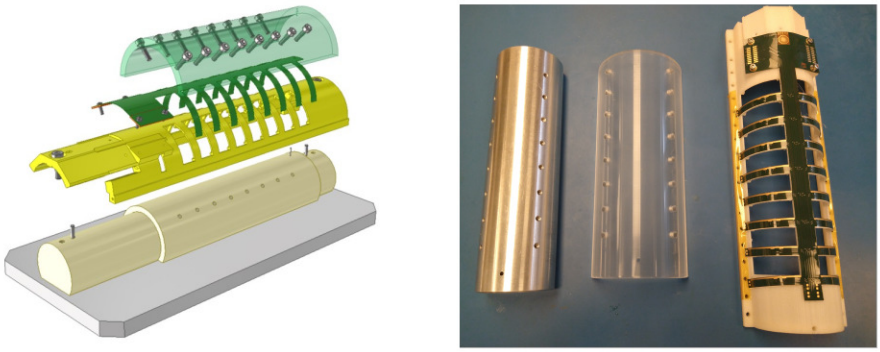


Figure 4.16: (left) Positioning of the elements employed for exoskeleton and exo-FPC alignment in the gluing procedure. (right) At the left, aluminum half-cylinder used as a base to support the exoskeleton; at the middle, transparent cylindrical cap used to press the FPC to the exoskeleton and at the right, exoskeleton + exo-FPC.

It is particularly important that the glue (Araldite 2011) is uniformly distributed under the region where bonding strips are, in order to ensure stable support during the ultrasonic bonding process. A second aspect to be avoided is the possibility that exceeding glue spread over the bonding region, compromising the surface. A proper technique has been developed also to control the amount of distributed glue.

Finally, the exoskeleton, with the exo-FPC glued on it, must be glued also to the carbon foam support structures. Longitudinal position of the exoskeleton along the mandrel is defined by pins and it must fit with the underneath support structures. Once aligned, the exoskeleton is fixed to the mandrel with two screws.

4.2.4 Exo-FPC to sensor interconnections

Next operation, before removal from the mandrel, is the interconnection between all the remaining sensors in the super-ALPIDE matrix and the exo-FPC. The geometry of the system and the dimensions of the bonding machine head, as sketched in figure 4.17 (left), put a lot of constrains to the exercise. The openings in the exoskeleton to access the bonding pads on the sensor are large 10 mm in the longitudinal direction. The first bond is done at the position of the sensor pad, then the machine head lifts up above the height of the FPC and finally moves towards the soldering strip on it, where the second bond is performed. In order to reduce the length of the wires, the height of the wire loop above the level of the FPC surface is kept minimal. They have a total extension of 12 mm approximately. The distance between the two soldering points, in the plane of the sensor surface, is 10 mm as represented in figure 4.17 (left); this separation is

needed because the head of the bonding machine has the wire dispenser inclined at an angle of 45^o and risks to touch the edge of exoskeleton while moving. Bonding pads on the sensor are 200 μm distant from each other, for a total of 45 interconnections, making the region really dense of wires. To reduce the possibility of contact, the direction of the bonding foot has been rotated by few degrees with respect to the line between bonding points, for the soldering at the sensor. As for the wire-bonding of the edge-FPC, the connections are performed between two points seating in the same motion direction of the machine head; so, the bonding machine do not require unusual actions to operate. Wire-bonding between the following two adjacent points is performed simply by rotating the mandrel with the help of the step motor. Results of the bonding procedure between the sensors matrix and the exo-FPC are shown in the right picture of figure 4.17.

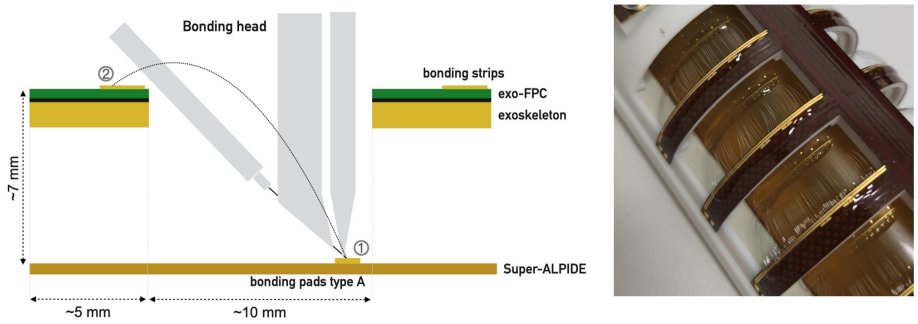


Figure 4.17: (Left) Sketch of the bonding procedure through an opening in the exoskeleton. (Right) Results of an actual attempt of bonding between sensors in the super-ALPIDE matrix and the exo-FPC.

Four assemblies, e.g. in figure 4.18 using kapton foil pieces of real super-ALPIDE dimensions with metallic pads placed at the same positions of those in the functional detector, were performed to study the bonding procedure.

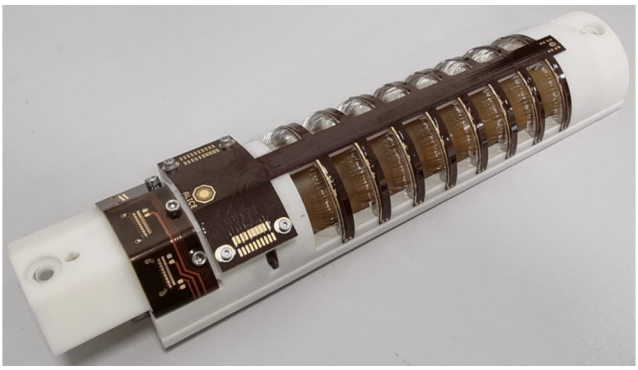


Figure 4.18: Assembly with a piece of kapton foil of real super-ALPIDE dimensions and metallic pads at the same positions of those in the functional detector, oriented to the bonding study.

4.2.5 Removal from the mandrel and transport box

Very last step in the assembly procedure is the removal of the detector from the mandrel. At this stage it is anchored to the mandrel by: two screws through the exoskeleton, three small screws that keep in position the edge-FPC and the adhesive tape that keep the sensor wrapped around the mandrel. First operation is the removal of the adhesive tape and it is a pretty delicate activity. The tool shown in figure 4.19 has been designed to cut, in a controlled way, the adhesive tape, few hundreds μm away from the sensor edge. It consists of a 3d printed plastic support, equipped with a sharp cutter, that could be approached close to the edge of the sensor by means of micrometric handles, and can slide along a track in the longitudinal direction. Same tool can be used to cut two of the three extensions of the edge-FPC used to fix it to the mandrel by screws. Third one can be simply cut with a scissor. At this point the last two screws, keeping the exoskeleton hooked to the mandrel con be removed and the detector is free to be lift up.

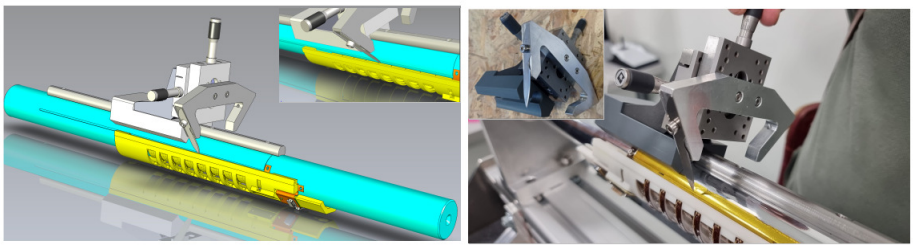


Figure 4.19: Sketch of the design (left) and picture (right) of the tool designed to cut, in a precise way, the adhesive tape that maintains the sensor attached to the mandrel.

On the other hand, a transport box, as the one shown in figure 4.20, made in plexiglass, has been designed. Special care has been put in reducing possible vibration in the detector, that could cause sensor and wire-bonds breaking. Two pins keep the object in position and soft foam is used to support it.

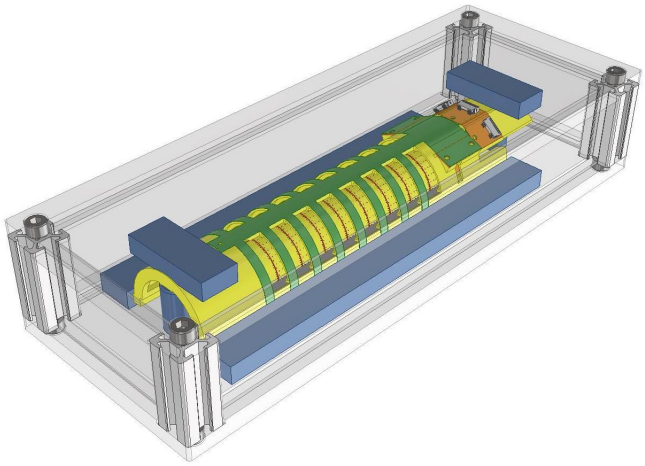


Figure 4.20: Design of the super-ALPIDE transport box.

4.3 Current status and plans

At the moment two assemblies, shown in figure 4.21, have been fully completed using a dummy version of the super-ALPIDE sensor. As introduced in section 4.2.3, this variant is made of blank-wafer of exactly the same size of the super-ALPIDE, equipped with metallic pads equal to the ones on the real detector. The purpose of using dummy silicon instead of the actual sensor is to test and verify the suitability of the tools and procedures developed to bend the super-ALPIDE and to assemble the components in a stable way. Having acquired such technical experience, the next step for the realization of the project will be to perform the assembly using a working super-ALPIDE detector. This activity is planned to be completed by the end of 2023.

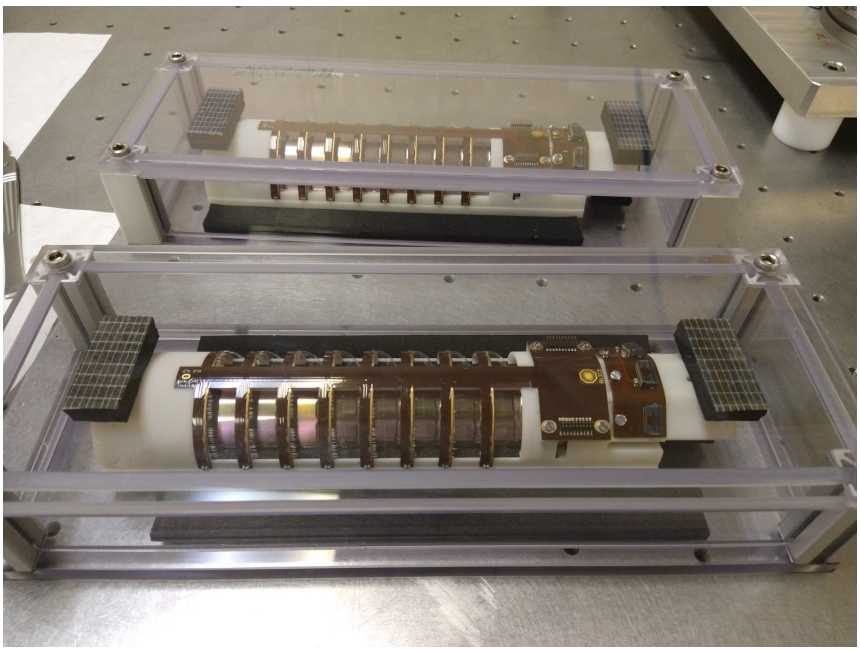


Figure 4.21: Two finalized assemblies with a dummy version of the super-ALPIDE sensor.

Chapter 5

Performance of first MAPS structures implemented in a 65 nm CMOS process

The Inner Tracking System of ALICE, built with chips developed with the image sensor foundry TowerJazz (TJ), has managed enhancing the MAPS technology to achieve advanced timing, fast readout and resistance to radiation needed in LHC experiments. The shift to an increased resistivity of the active sensitive volume (ASV) and introduction of additional doping steps to form nested wells (see section 2.3.1), led the TJ 180 nm process technology to became available for MAPS detector designers. In this process, implantations and distributions of resulting electric fields were optimized [100, 101], making possible depletion of ASV and the use of both transistor types for in-pixel circuit networks.

Significant experience has been obtained in the 180 nm TowerJazz CMOS imaging technology with large scale production of ALPIDE [102] for the ALICE ITS2 upgrade ($> 10\ m^2$). For the development of the upgrade for the ITS3 [103], efforts have been made to port the features of the enhanced TJ 180 nm process to the TPSCo 65 nm ISC technology, the 65 nm CMOS imaging flavor of TPSCo, a joint venture between TowerJazz and Panasonic. The new process is becoming available for the developers of a new generation of MAPS detectors. The TJ 65 nm process allows more than four-fold increase of the number of transistors per pixel compared to the older 180 nm. The key benefit of MAPS detectors to be developed in the TJ 65 nm, will be higher spatial resolution and embedment of processing, such as clustering, time-of-arrival or possibly even amplitude measurements in pixels. After the first Multi-Layer-Reticle submission in December 2020 (See section 5.2) and subsequent measurements it has been definitely chosen for the ITS3.

In the following sections, a description of different sensor prototypes based on the 65 nm technology developed for the ITS3 upgrade will be provided, putting special attention on the Analogue Pixel Test Structures (APTS). Results of

the electrical characterization of this devices, tested in laboratory, will be also presented.

5.1 65 nm process technology

A step towards smaller-feature-size technologies is necessary to improve the performances of pixel detectors in terms of size and/or functionality embedded per pixel. In general, any front-end circuit can benefit from a more advanced technology node in terms of density, especially for digital blocks. Analogue parts of pixels can also benefit from the reduced capacitances associated with the smaller sizes, resulting in a lower noise equivalent charge.

The 65 nm process is an advanced lithographic node used in CMOS (MOSFET) semiconductor fabrication. The particular manufacturing method used to make silicon chips is measured by the size of the transistor's elements. A 65 nm process technology refers to features 65 nm in size, also called "technology node" and "process node". The feature size is defined as the minimum length of the MOS transistor channel between the drain and the source. Figure 5.1 shows the typical structure of a MOSFET. The area between source and drain beneath the gate, channel, has length of L and width of W; it is where the inversion layer is formed when sufficient voltage is biased to the gate and drain which then turn on the MOSFETs. The oxide layer with certain thickness, t_{ox}, separates the gate and the channel.

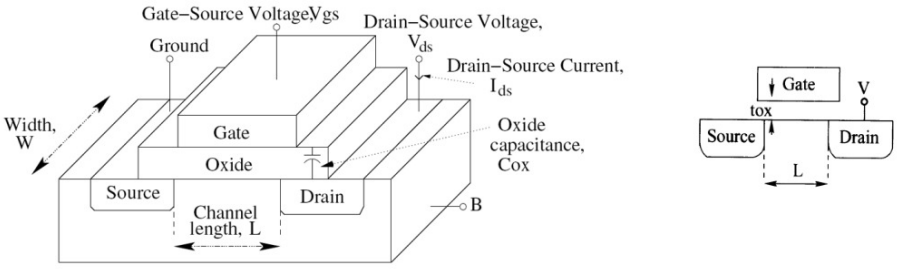


Figure 5.1: (left) Typical structure of a MOSFET with important parameters determining its characteristics. (right) CMOS transistor's physical feature. L represents the feature size and t_{ox} the oxide layer thickness [104].

In early 1970s, Gordon Moore of Intel Corp. predicted that the number of transistors in an integrated circuit would double itself in approximately 18-24 months. This prediction, popularly known as Moore's Law, has set out a challenge and a roadmap for designers to keep the scaling going. If double number of transistors can be incorporated on the same area, it means that double (roughly) functionality can be obtained for the same cost. Alternatively, the same functionality will be available at roughly half the cost. Moreover, the smaller the channel length, the faster would be the transient response of the transistors which would translate into better performance. The feature size

from 180 nm to 130 nm to 90 nm scale down by a factor of $S \approx 0.7$ [105]. If the feature size of the transistor is scaled by 0.7, the area would be scaled by a factor of $0.7^2 = 0.49 = \sim 0.5$. That means if feature sizes are scaled by a factor of about 0.7, the number of transistors could be packed twice on the same area as the previous technology.

Technology scaling down to 180 nm was based on the following reasoning. Reducing the feature size in the technology front-end (i.e. of the devices) and in the technology back-end (i.e. of the interconnect) combined with the addition of even more interconnect layers[1], drastically increased the density of the digital circuits while reducing the intrinsic gate switching delay. The main disadvantage of this is that the power supply voltage must be decreased too, potentially leading to an increase in gate switching delay. However, the threshold voltage (V_T) of the MOS transistors is lowered to compensate for this. In a full scaling scenario, where the scaling factor ($1/S$) is the same for all geometry and voltage parameters, one would have, for standard CMOS gates, a density increase of S^2, an intrinsic delay decrease of $1/S$ and a power consumption decrease of $1/S^2$, at a constant power density [106].

However, as technology scaling goes on to 90 nm and below (65 nm, 45 nm), physical and quantum-mechanical effects that were previously not relevant become influential or even dominant. An example of this is the advent of no longer negligible leakage currents. On the other hand, as all geometrical dimensions become smaller, variability of technological parameters also gains in importance. The same absolute tolerance becomes relatively more important when the absolute value of a parameter decreases. This holds for V_T, doping levels, widths and lengths. A V_T variation of 50 mV has more effect on the circuits when V_T is 200 mV compared to a V_T of 700 mV for instance.

5.1.1 General consequences of 65 nm CMOS at the circuit level

Scaling technology to nanometer sizes has brought effects to the surface that were negligible before. The most important ones are: subthreshold leakage current, gate leakage current, interconnect delay and increased process variability.

[1] In general, a sensor front end may consist of the following interconnect layers: 1) Substrate Layer. The bottommost layer that provides structural support for the sensor and houses the electronic components. 2) Signal Layer. This layer carries the electrical signals from the sensor elements to the subsequent layers or circuitry. It includes interconnects, traces, and signal routing. 3) Power Layer. The power layer is responsible for distributing power to the various components of the sensor front end. It includes power supply traces and distribution lines. 4) Ground Layer. This layer is dedicated to providing a low-resistance path for electrical grounding and serves as a reference point for the electrical signals. 5) Shielding Layer. Sometimes, a sensor front end may have a dedicated shielding layer to reduce interference from external electromagnetic sources. It helps protect the sensitive sensor components from noise and electromagnetic interference.

These layers are typically stacked on top of each other and connected using vias or interlayer connections to establish the required electrical connectivity and signal flow within the sensor front end. Generally, as technology nodes progress to smaller sizes, the number of available metal layers increases.

Their effect on circuit behavior will be described in the next subsections.

Subthreshold leakage current

The current conducted by a MOS transistor when it is off ($V_{GS} = 0$) is given by:

$$I_{subthreshold} = I_0 \exp\left(\frac{-V_T}{nkT/q}\right) \tag{5.1}$$

The subthreshold current thus increases exponentially with decreasing V_T. Moreover I_0 in this equation is inversely proportional to L. Nanometer scaling thus leads to a dramatic increase in leakage current up to the point where it can no longer be ignored in the power consumption nor in the functionality of the circuits. For very short transistors the subthreshold leakage not only increases because L scaling implies V_T scaling. When the drain region of a MOS transistor is close to the source region, the depletion region around the source and drain regions interact with each other. This lowers the potential barrier at the source side.

Gate leakage current

Before the nanometer era, gate currents were considered as pure dynamic. MOS gates were capacitors and did not conduct DC currents. However, since the advent of oxide thicknesses of only a few nm, current can tunnel through the gate and have a static "leakage" component. The gate leakage current is described by :

$$I_{gate\_leakage} = KW\left(\frac{V_{gb}}{t_{ox}}\right)^2 \exp\left(-\alpha t_{ox}/V_{gb}\right) \tag{5.2}$$

where K and α are fit factors. It is worth noticing that gate leakage is only present when there is a voltage across the gate, i.e. when the transistor is on. Subthreshold leakage on the contrary occurs when a transistor is off. Gate leakage currents affect the power consumption and functionality at the circuit level.

Interconnect delay and scaling

Interconnect delay in a MOS transistor refers to the time delay that occurs as a result of the resistance and capacitance of the metal interconnects that connect various components of the transistor. A simplified, first-order approximation of interconnect delay is given by:

$$t_{wire} = \rho\kappa\left(\frac{L}{\lambda}\right)^2 \tag{5.3}$$

where ρ and κ are the resistance and capacitance per unit area, and λ is the technology-related wire pitch. This shows that the interconnect delay is inversely proportional to the square of the wire pitch λ, that scales down with node technology. This means that if the interconnect length (L) and interconnect pitch (λ) scale identically, the interconnect delay will remain constant with technology scaling. As stated above, the intrinsic gate delay of a technology decreases with technology scaling. This means that the interconnect delay will gain in relative importance when the technologies are scaled down.

Process variability

The tolerance on some technological parameters will increase when the technology scales. One example is the discrete nature of doping levels. The statistical variation of the number of dopants N varies with $N^{1/2}$. This increases the uncertainty on V_T for small N. For decreasing channel dimensions, the number of dopants that realizes a given doping level also decreases (see figure 5.2). The random placement of the dopants creating the source and drain also causes an uncertainty on the effective channel length. This effect is also enforced as the number of dopants goes down.

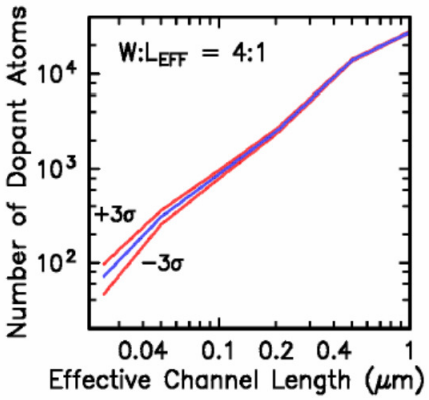


Figure 5.2: Number of dopant atoms vs. channel length [106].

5.1.2 Technology specifications

With the 65 *nm* technology, to reduce chip power, both the active and leakage power components are reduced. For active power it is important to reduce the voltage of operation, this has traditionally been possible by using thinner gate oxide following constant field scaling [107]. However, as introduced before, reduction in oxide thickness leads to significantly higher gate leakage, reduced reliability, and degraded channel mobility. These factors tend to offsets performance gain from thinning oxide at a given power, limiting the gate oxide scaling

to around 1.8 nm electrical oxide thickness, with the transistors operating at 1.1 V. Front-end technology features include shallow trench isolation, retrograde wells, with enhanced strained silicon channels, shallow abrupt extensions, halo implants, deep source/drain, and nickel salicidation. In order to support the small gate lengths following constant field scaling it is important to scale the gate oxide, as mentioned earlier.

Increased leakage and degraded channel mobility limits the gate oxide physical thickness to be scaled down to around 1.2 nm. Figure 5.3 shows a cross sectional TEM (Transmission Electron Microscope) of the gate oxide.

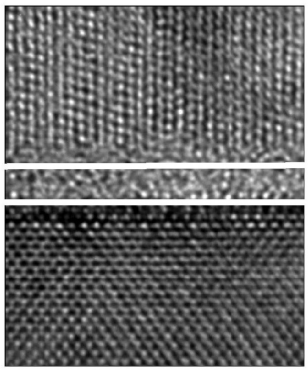


Figure 5.3: Cross-section TEM of 65 nm gate oxide [108].

Transistors

To get high transistor performance the transistor gate length is scaled down more aggressively. For this 65 nm technology, the gate length is scaled down to 35 nm, as shown in figure 5.4 for NMOS and PMOS devices. To improve the transistor performance their short channel effects (SCE), series resistance (Rext) and mobility were improved for NMOS by optimizing the thermal anneals and using advanced implants and coimplants; while to further improve the mobility second generation strain enhancement techniques were optimized. Similarly for PMOS SCE and Rext, the source and drain regions, the tips (S/D extensions) and the halos were improved by optimizing anneals and using advanced implants, the strain in the channel was optimized to get higher mobility using higher Ge fraction in stressor regions. Without gate oxide scaling, the junction depths are scaled aggressively for improving SCE. Ultra shallow junctions are formed in the transistor by controlling diffusion using appropriate thermal cycles and coimplants. These features enable MOSFETs with well controlled short channel characteristics, the threshold voltage reduces by only 20 mV for PMOS devices and by less than 120 mV for 35 nm NMOS gates. Transistor gate length scaling results in higher drive currents and this increases the relative importance of series resistance. Due to solid solubility and activation

limits, shallow junctions invariably lead to an increase in sheet resistance. However, using coimplants and optimized anneals, transistors with improved series resistance and shallow junctions have been fabricated.

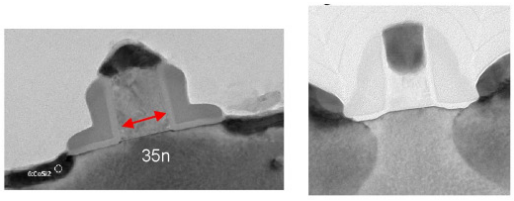


Figure 5.4: TEM cross-section of 35 nm NMOS (left) and PMOS (right) transistors [108].

Mobility enhancement with strain

Uni-axial channel strain has been shown to be an efficient way to enhance mobility by band repopulation and phonon suppression for electrons and due to band structure warping and carrier redistribution for holes [109]. The NMOS channel strain was optimized by improved thermo-mechanical design of the transistor's spacers and stress inducing film, this creates stress in both in plane and out of plane directions. As a result the mobility increases by more than 200 compared to ref. [110] and [111]. Similarly the hole Mobility in PMOS is increased by 3000 by optimizing stressor trench shape and using higher Ge content films. This is because band structure warping leads to reduction of effective mass due to band repopulation as well as mobility improves due to scattering reduction.

5.2 Multi-Layer-Reticle 1

A first submission in TPSCo 65 nm was done as a Multi Layer Reticle (MLR1), where several masks are implemented on the same reticle to reduce the number of reticles and the cost. It consists on 55 small prototype structures in different processing flavors as a test of the technology. Wafers were available in summer 2021, thinned and diced chips became available for testing early September. Very intense test setup developments and improvement are still in progress at the moment of writing this thesis. Figure 5.5 shows the layout of the submitted designs and in figure 5.6 pictures of a diced quarter wafer and diced chips can be observed. They include from single transistor to complex pixel structures and circuits for the evaluation of the technology and its detection performance (Multiple pixel test matrices, LVDS/CML receiver/driver, bandgaps, temperature sensors, Voltage Controlled Oscillators, amplifiers, ring-oscillators, transistor test structures and an interconnect design).

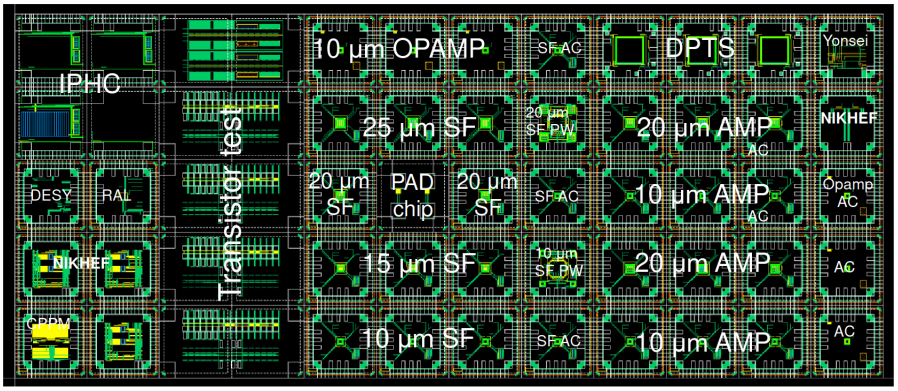


Figure 5.5: Overview of the MLR1 submission in the TPSCo 65 *nm* ISC technology.

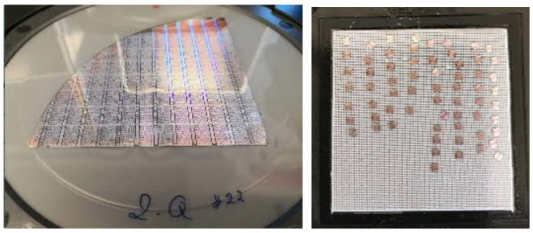


Figure 5.6: (left) A diced quarter wafer on tape, and (right) hand picked diced chips belonging to the MLR1 submission in the TPSCo 65 *nm* ISC technology.

5.2.1 Test Structures of MLR1

Three different pixel sensor designs corresponding to a given sensing node geometry, first validated in 180 *nm*, were prototyped in the first MLR run in the 65 *nm* ISC technology, as shown in figure 5.7.

Table 5.1 summarizes the three pixel test structures of particular significance for testing different design options in the ITS3 context. Analogue Pixel Test Structure (APTS) is a 6×6 pixel test structure, featuring analogue front-end with parallel analogue output from the central 4×4 pixels. It is available in four variants of pixel pitches (10, 15, 20, 25 μm), two versions of output buffer (source follower and op-amp), and the three design variants mentioned above. Digital Pixel Test Structure (DPTS) is a 32×32 pixel test structure with asynchronous digital readout and 15 μm pixel pitch. CE-65 is a pixel matrix with rolling shutter readout with single analogue output. It is available in two different sizes (64×32 and 48×32 pixels, 15 and 25 μm pixel pitch, respectively) and each matrix is divided into three sub-matrices with different in-pixel electronics (AC op-amp, DC op-amp, source follower). A more detailed description of these chip variants will be provided in the following.

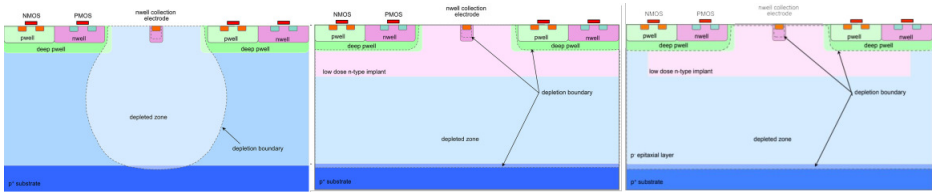
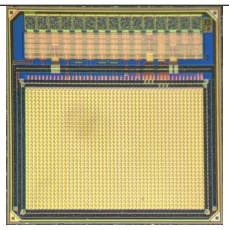
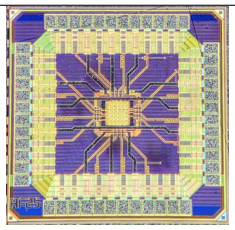
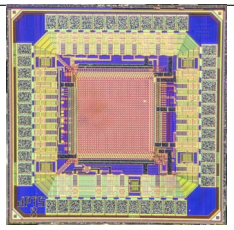


Figure 5.7: Various sensor designs prototyped in the first MLR1 run in the 65 nm ISC technology: (left) conventional, (middle) modified with blanket deep low dose implant over the full pixel area and (right) modified with gap in the same implant over the pixel boundary.

Table 5.1: Pixel chip prototypes. Die size is $1.5 \times 1.5\, mm^2$ [112]

	CE-65	APTS	DPTS
			
Matrix	64×32, 48×32	6×6	32×32
Pitch (μm)	15, 25	10, 15, 20, 25	15
Readout	Rolling shutter, analogue	Direct analogue (central 4×4)	Asynchronous digital with ToT

5.3 Digital Pixel Test Structure (DPTS)

The Digital Pixel Test Structure (DPTS) consists of a 32×32 matrix of $15 \times 15\, \mu m^2$ pixels (+ 1 monitor pixel). It aims at testing the reliability and performances of a distributed, asynchronous, time-based encoding, readout scheme. DPTS chips contain pixels that have fully integrated in-pixel front-ends with amplification and discrimination. They are read out in an asynchronous fashion and provide time-over-threshold (ToT) information. Figure 5.8 shows the high level functional diagram for DPTS. The analogue front-end comprising amplifier and discriminator triggers a digital encoder. The hits from all the pixels are digitized and conveyed by the digital readout into a single digital output. Defective or noisy pixels can be masked through the slow control registers. Selection for pulsing can be done again through the slow control registers or, alternatively through a decoder that directly accesses a subset of the pixels in the matrix. A single extra pixel feeds the buffered front-end output directly to an analogue output for front-end monitoring purposes. The readout scheme features a single digital output that is brought off-chip via a CML (Continuous Mode Logic)

buffer[2] that can be independently powered and biased. The 32×32 matrix represents what in reality would be a single column of pixels (1×1024) sharing a single digital output. To fit the column in a square format, it has been broken into groups, folded, and connected to a common line. A particular pixel addressing scheme is implemented to maintain a single output for all the columns, so that there is no arbitration in case of charge sharing. The approach was to handle all the activated pixels and delay the transmissions to avoid collision on the single output [97].

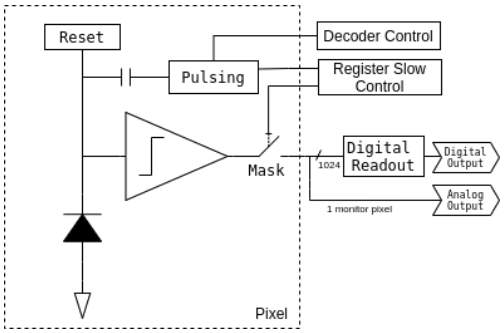


Figure 5.8: Functional diagram of Digital Pixel Test Structure.

The currents in the front-end are tunable across a very wide range of values. The main current of the front-end is IBIAS, whose value can be as low as 10 nA, whilst IBIASN should be nominally 1/10 of the main current. The analogue power of each pixel in this case would be 12 nW corresponding to an integration time of around 1us. The highest settable value for IBIAS is 5 μA and 500 nA for IBIASN; beyond these values, it is difficult to guarantee a good operating point for all the transistors in the circuit. In this mode, the power consumption is 6 μW with an integration time of tens of ns.

DPTS chips were taken to numerous test beams and showed excellent performance as represented in figure 5.9 (top). They were also tested after NIEL irradiation to the levels expected at ALICE ($\approx 10^{13} 1MeVn_{eq}/cm^2$), where they keep their performances. Even tests with doses 100 times larger ($10^{15} 1MeVn_{eq}/cm^2$) showed that the chips can still be operated at above 99% detection efficiency (5.9, bottom). These results validated the process modification, in sensors with gap, for better charge collection and full functionality of the front end and of the digital logic.

[2] A buffer is an electronic device that is used to isolate one circuit from another, by providing a high impedance input and a low impedance output. In other words, a buffer can receive a signal from one circuit and send it to another circuit without altering the signal in any significant way. A CML buffer is designed to operate in continuous mode, which means that it is always active and can handle a continuous stream of data without interruption. CML buffers are commonly used in high-speed digital systems because they can operate at very high frequencies while consuming relatively low amounts of power.

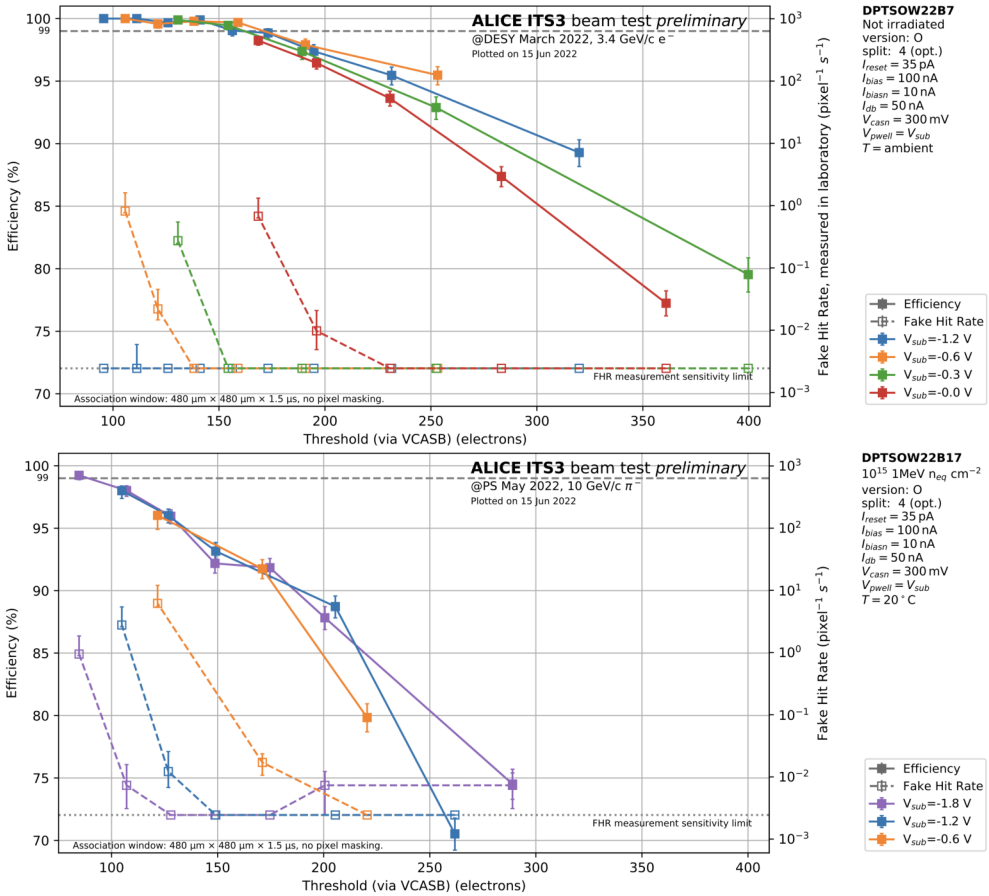


Figure 5.9: Efficiency and fake-hit rate vs threshold (top) before irradiation and (bottom) after NIEL irradiation to 10^{15} $1MeVn_{eq}/cm^2$ of DPTS [112].

5.4 CE-65

CE-65 structure is intended to obtain preliminary information on the tracking capabilities of the 65 nm technology node with a matrix large enough to be tested on a beam. This sensor family consist of four different chip variants featuring diverse sensing node geometries with similar readout electronics and matrix sizes of 64×32 and 48×32 with pixel pitches of 15 and 25 μm respectively, as reported in table 5.2. Each matrix is further divided into three smaller sub-matrices that incorporate different in-pixel electronics architectures. These are: AC coupled pre-amplifier, DC coupled pre-amplifier and DC coupled source-follower input stages. The size of each sub-matrix depends on the chip variant. A high level functional block diagram of the CE-65 pixel matrix is sketched in figure 5.10. The whole design can be divided into several major building blocks:

pixel matrix, column and row selectors, column buffers and an output buffer.

Table 5.2: Chip variants of the CE-65 structure

Chip variant	Pixel pitch (μm)	Sensor geometry	Matrix size (pixels)
A	15	Standard	64×32
B	15	Modified with gap	64×32
C	15	Modified	64×32
D	25	Standard	48×32

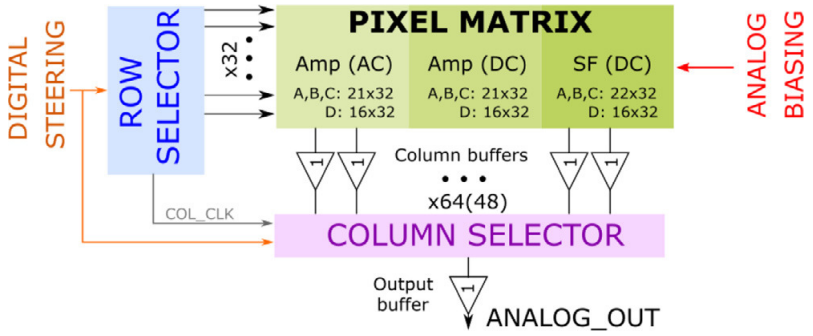


Figure 5.10: High level functional block diagram of a CE 65 matrix.

In the AC-AMP architecture (figure 5.11, left), the sensing node (reversely biased diode to substrate) is DC separated from the input stage of the readout electronics by a capacitance C_{SEP} ($\sim$ 10 fF). This AC coupling disentangles the diode reset voltage level[3] from the electronic readout voltage limitations, allowing for more flexibility in the reverse biasing of the sensor, as the readout circuit is not exposed to the DC component of the sensor bias. Accordingly, the voltage applied to the resetting diode is not limited by the electronics supply voltage. The input stage is made by an amplifier operating in 3.3 V power domain, followed by a PMOS source follower that acts as a level shifter to 1.2 V domain. The DC-AMP sub-matrix exploits the same in-pixel electronics (figure 5.11, left), but in this variant the sensing node is directly connected to the gate of input transistor. Thus, the reset diode is no longer present and the input node voltage is determined by the pre-amplifier operating point. In the third approach a simple 3T-like architecture has been utilized (figure 5.11, right). The sensing node is directly connected to the source-follower made on NMOS transistors operating in 3.3 V power domain, that at the same time also act as

[3]The reset voltage connected to the sensing node on a circuit is a voltage level that is used to reset or discharge the sensing node. Sensing nodes typically output a voltage that is proportional to the measured parameter. However, over time, the voltage on the sensing node can drift due to leakage or other effects. To ensure accurate measurements, it is often necessary to periodically reset or discharge the sensing node to a known voltage level. This is done by connecting a reset circuit to the sensing node.

a level shifter to 1.2 V domain. The reset diode has been implemented similarly as for AC-AMP, but the reset voltage is directly transferred to the gate of the input stage, so VRESET cannot exceed the supply limits in order not to damage the transistor.

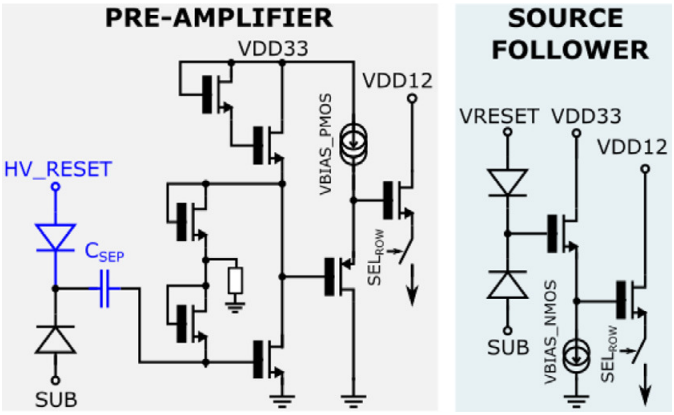


Figure 5.11: In pixel circuitry of (right) DC source follower, (left) DC pre-amplifier and AC pre-amplifier, whose scheme is analogous with AC coupling added (blue) [113].

The matrix is read out in the rolling shutter mode, where analogue outputs from each pixel are serialized to the one common output one after another. The integration time can be configured down to 50 μs, translating to readout speed up to 40 MHz. The signal is digitized outside the chip by a fast 16-bit Analogue-to-Digital Converter, defining the signal arbitrary digital unit (ADU).

Initial performance studies of two CE-65 variants, featuring basic and optimized diode (modified with gap), both with 15 μm pitch, showed low noise operation and uniform pixel baseline among submatrices, as well as excellent charge collection efficiency. As an example of measurement results, figure 5.12 shows the ^{55}Fe spectrum for one pixel of a single pixel cluster, where the seed pixel carries almost whole cluster energy, selected from the DC-SF submatrix with optimized diode. This approach allows to get rid of the pixel-to-pixel gain variation at the cost of reduced statistics. A clear separation between the two characteristic lines $Fe - K_\alpha$ (5.90 keV) and $Fe - K_\beta$ (6.49 keV) is observed and confirms the low pixel noise as well as a promising energy resolution.

On the other hand, while significant charge sharing has been measured with a basic collecting diode, an optimized structure was demonstrated to strongly mitigate this effect, for the AC-AMP pixels with a bias voltage of 10 V. Charges are more focused on a single collection node for the optimized structure, which is beneficial for the time resolution and tolerance against radiation generating bulk damage. A comparison of the charge sharing between basic and optimized sensor diodes can be observed in the maps of the average pixel contribution to the total cluster signal, presented in figure 5.13. The significant charge sharing for the

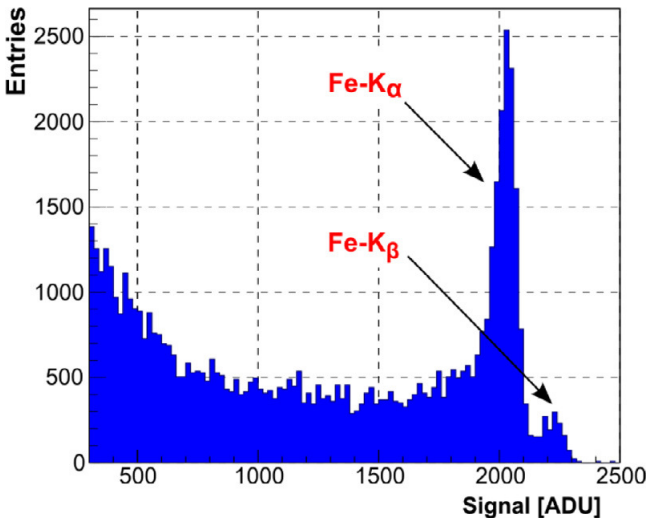


Figure 5.12: Iron source spectrum of single pixel clusters from one source follower pixel optimized diode [113].

basic diode structure is confirmed, since the seed pixel collects in average slightly less than half of the total cluster signal. The electric field in the optimized diode is designed to accelerate the charge from the pixel edge to the collection electrode, and strongly suppresses charge sharing.

5.5 Analogue pixel test structures (APTS)

Analogue Pixel Test Structure (APTS) aims at sensor characterization, so it has the purpose to finely test the parameters of the technology itself. It consists of a 6x6 pixel matrix with 16 buffered analogue outputs for parallel readout of the central 4x4 pixels. It is available in the three design forms illustrated in figure 5.7, four variants of pixel pitches (10, 15, 20, 25 μm), DC and AC coupled in-pixel architecture and two versions of output buffer: source follower and op-amp. The APTS family consists of 36 variants of single matrix chips and 2 variants of 4x multiplexed matrix, 44 pixel variants in total. Figure 5.14 shows a high level functional block diagram of the in-pixel architecture of the central 4x4 pixels, where both DC (top) and AC (bottom) coupled versions are represented.

5.5.1 Principle of operation

The front-end circuit is similar for all chip flavors, as can be noticed from figures 5.15 and 5.16. It operates between AVDD = 1.2 V and AVSS (ground) and is controlled via three currents (IBIASP, IBIASN and IRESET) and one voltage (VRESET). The reverse bias (Vbb) can be supplied separately to the substrate

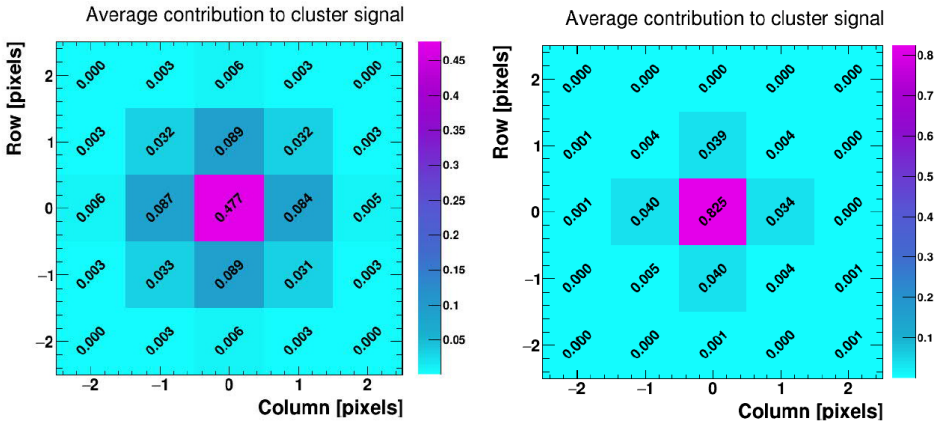


Figure 5.13: Average contribution to cluster signal of individual pixels exposed to a ^{55}Fe source, for basic (left) and optimized (right) diode geometries from CE-65 AC-AMP pixels with a bias voltage of 10 V. [113].

(V_{sub}) and the deep p-wells (V_{pwell}) hosting the circuitry, however, in this work they were kept at the same potential. The maximum applicable reverse bias voltage is -5 V, but to ensure a safe operation, the largest Vbb applied was -4.8 V. All the parameter values are externally generated and supplied to the chip via the interface pads.

Initially, the collection diode is set to a potential V_1. Thus the input capacitance C_{in}, defined by the sum of the junction capacitance of the collection diode C_j ($\sim$ 1 fF) and the parasitic capacitance of the in-pixel readout circuit C_p, is charged to $Q_1 = C_{in}V_1$. A particle passing through the sensor deposits charge in the epitaxial layer and the diode collects a certain amount Q_{coll}. This charge is subtracted from the input capacitance which is discharged to $Q_2 = Q_1 - Q_{coll}$, resulting in a potential drop on the source follower (M2) transistor gate $V_2 = Q_2/C_{in}$. As such, the input voltage signal is buffered by the first stage PMOS source follower consisting of input transistor M2 and M1 to set the bias current IBIASP. The source follower transistor (M2) acts as a buffer which isolates the sensing diode from the readout circuit. The source follower gain g_s is $\lessapprox 1$ and the output voltage is proportional to the input voltage ($V_{out} = g_s V_{in}$) (see figure 5.15). The output voltages from this stage before and after the charge collection are, respectively, $V_{out1} = g_S V_1$ and $V_{out2} = g_S V_2$, where g_S is the source follower gain, and the potential drop is: $\Delta V_{out} = g_s V_1 - g_s V_2 = g_s(Q_1 - Q_2)/C_{in} = g_s Q_{coll}/C_{in}$

The output voltage difference is ideally proportional to the collected charge. The output line is routed to the second stage NMOS source follower consisting of input transistor M3 and M4 to set the bias current IBIASN. The NMOS transistors are subject to the reverse bias applied to the sensor (V_{pwell}). As the reverse bias increases, the threshold voltage (V_T) of the NMOS transistors also increases if their substrate is not connected to their source, which is the

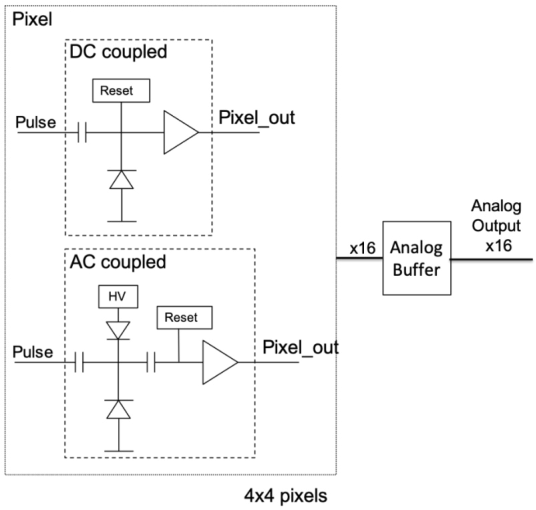


Figure 5.14: High level functional block diagram of the in-pixel architecture of the central 4x4 pixels in the APTS.

case for M3. The increase in V_T of M3 due to the reverse bias will cause the output of the NMOS follower (the source of M3) to approach more the ground, so an increase in reverse bias translates into a decrease in margin of the circuit. This would probably require a decrease of the current IBIASN to allow for more operation margin (more room for an increased amplitude with increased bias).

Two processes are responsible for the discharge of the sensing diode: leakage current and collection of charge (deposited by a particle). The former is a small effect that can be considered constant per unit of time while the latter is a short pulse, large respect to the leakage current. Since both mechanisms discharge the collection diode, from time to time a potential must be applied to the sensing node to restore the original conditions. This is done by the reset transistor M0, which continuously charges the input capacitance, maintaining the input node at a potential VRESET.

A test circuitry, which can inject charge in the collection electrode through a capacitance of $C_{inj} = 242\ aF$ (Figure 5.14), is also integrated in the pixel. The amount of injected charge can be regulated by an external voltage reference, V_h (0 - 1.2 V), and the injection is triggered by asserting the TRG signal via an interface pad. In the case of chips with AC coupled pixels, the collection electrode to front-end coupling capacitance is equal to 10 fF. In the source follower version, the output of the second in-pixel stage is connected to two additional source followers[4], biased with currents IBIAS3 and IBIAS4 as shown

[4]A source follower is also known as a common drain amplifier, as the output voltage is taken from the source terminal, which is connected to the input signal source. In a source follower circuit, the input signal is applied to the gate of a field-effect transistor (FET), and the output signal is taken from the source of the transistor.

in figure 5.15, operated at typical values of +800 μA and -6 mA respectively. At this stage the signal is buffered to the output BOARD pads and read with a high impedance. A signal fall time of 140 ps and about 50% amplitude loss are measured on board for the injection of 100 e^- at typical value of biases. The source follower structures have a power consumption of 108 mW.

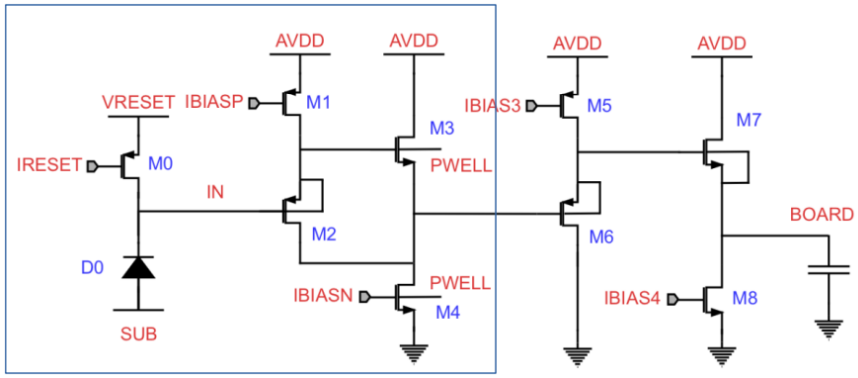


Figure 5.15: Schematic of APTS source follower version.

For the op-amp version (Figure 5.16) the OUT2 is connected to the high impedance non-inverting input of the op-amp circuit[5], which is configured as unity gain buffer with the common source transistor M5, terminated at 50 Ohm on board. The amplifier is biased with two currents IBIAS3 and IBIAS4, operated at typical values of +800 μA and +2.5 mA respectively and two voltages VCASP and VCASN, operated at 300 mV and 750 mV respectively. The op-amp structures present a power consumption of 116.5 mW. A signal fall time of $\sim 204ps$ and 31% amplitude loss is measured on board for an injection of 300 e^-, achieving better timing performance than the basic source follower structures.

A source follower and an operational amplifier (op-amp) connected in unity gain both have the same basic function of providing a high input impedance and low output impedance buffer stage. However, there are some differences in

The primary advantage of a source follower is that it provides a high input impedance, which means that it draws very little current from the input source, and it does not load the source down. One disadvantage of a source follower is that the output voltage is reduced by the voltage drop across the FET. Therefore, it cannot provide voltage gain, but it can provide current gain.

[5] An Operational Amplifier, or op-amp for short, is fundamentally a voltage amplifying device designed to be used with external feedback components such as resistors and capacitors between its output and input terminals. These feedback components determine the resulting function or "operation" of the amplifier and by virtue of the different feedback configurations whether resistive, capacitive or both, the amplifier can perform a variety of different operations. By connecting the output directly back to the negative input terminal, 100% feedback is achieved resulting in a Voltage Follower (buffer) circuit with a constant gain of 1 (Unity). In this non-inverting circuit configuration, the input impedance R_{in} has increased to infinity and the feedback impedance R_f reduced to zero. The output is connected directly back to the negative inverting input so the feedback is 100% and Vin is exactly equal to Vout giving it a fixed gain of 1 or unity.

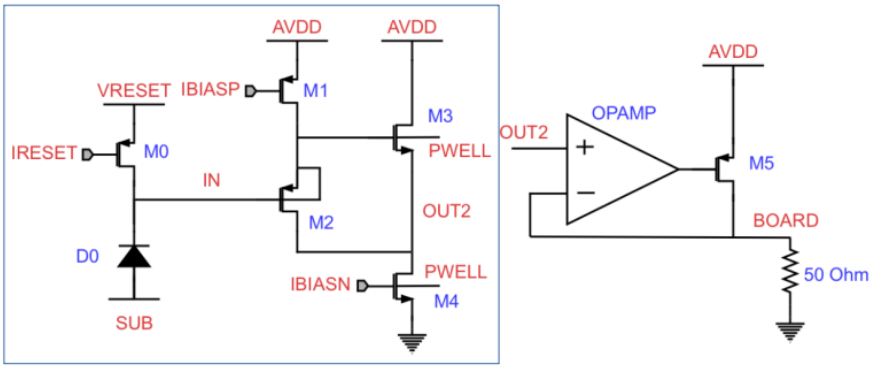


Figure 5.16: Schematic of APTS op-amp version.

their performance and behavior. A source follower is a simple amplifier circuit that uses a single transistor to provide a buffer between the input and output signals. The input signal is applied to the transistor's gate, and the output signal is taken from the transistor's source. The output signal is a replica of the input signal, but with a lower output impedance. The voltage gain of a source follower is slightly less than unity, and it can introduce some distortion into the signal due to the voltage drop across the transistor. An op-amp connected in unity gain is a more sophisticated amplifier circuit that uses multiple transistors and feedback to provide a high input impedance and low output impedance buffer stage. The input signal is applied to the op-amp's non-inverting input, and the output signal is taken from the op-amp's output. The op-amp is configured to provide a gain of exactly one, so the output voltage is an exact replica of the input voltage, but with a lower output impedance. The voltage gain of an op-amp connected in unity gain is precisely unity, and it can provide a very low distortion signal.

Table 5.3 summarizes the electrical specifications for pixel circuits of the Analogue Pixel Test structures, operated at $T_A = 27^o$ and AVDD =1.2 V.

Table 5.3: Electrical specifications for pixel circuits of the APTS.

Parameter		Typ.	Unit
Power consumption	SF structures	108	mW
	op-amp structures	116.5	mW
ENC	DC coupled	27	e
	AC coupled	54	e
Pulse-IN coupling capacitance	DC coupled	242	aF
	AC coupled	242	aF
Collection electrode-input of frontend coupling capacitance in AC chips		10	fF

5.5.2 Studies on the performance of APTS source follower version.

In order to validate the sensors from the point of view of charge collection efficiency, APTS source follower sensors have been studied in a setup that includes a radioactive source ^{55}Fe.

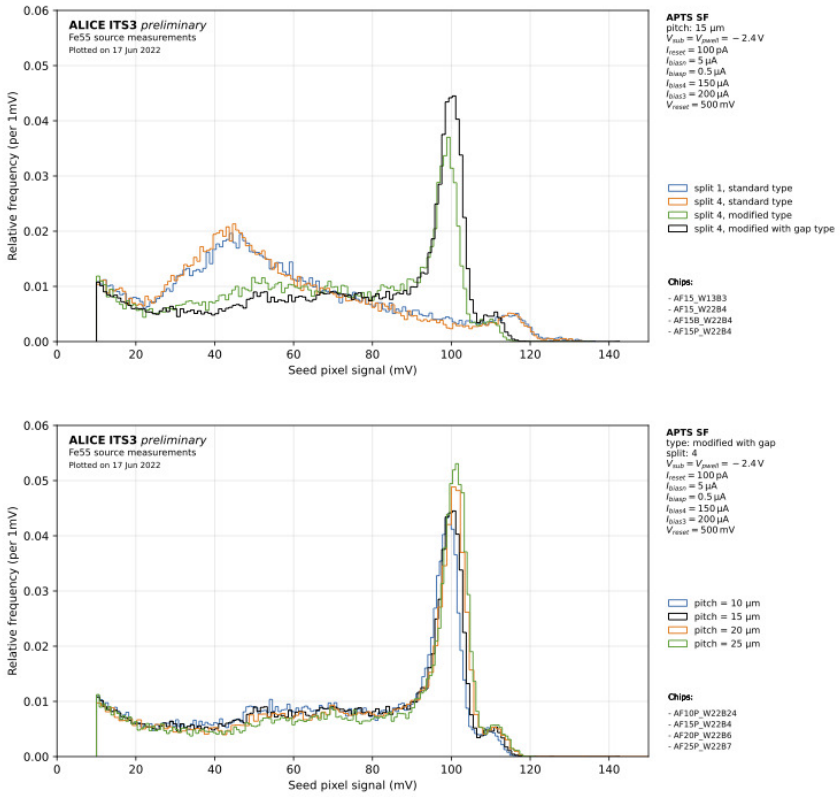


Figure 5.17: APTS ^{55}Fe spectra of seed pixel for different sensor types with pitch of 15 μm (top) and for different pitches for sensors modified with gap (bottom).

Figure 5.17 shows the ^{55}Fe spectra for different APTS sensor types with same pitch of 15 μm and for different pitches of sensors modified with gap. On the bottom plot, the peaks related to K_α (5.9 KeV, calibration peak) and K_β (6.5 keV) can be well distinguished. In the plot at the top a similar trend is observed for the modified process, with the calibration peak slightly more pronounced for the modified with gap type, while for the standard types a dominant peak much more pronounced than the calibration peak is observed at lower energy. This corresponds to what is usually named charge sharing peak,

i.e. when the signal is shared by multiple pixels. The process modified with gap shows a more efficient charge collection compared with the other variants. In the modified with gap version, for all the pitches, the calibration peak is by far the dominant contribution to the distribution, much more pronounced than the sharing peak. This is in line with high charge collection efficiency and would lead to a potentially better signal-to-noise ratio. Moreover it can be noticed that, comparing different pitches, a similar result is observed; this would have the benefit of choosing the pitch for the final sensor independently of the charge collection.

5.5.3 Laboratory measurements of the APTS op-amp performance

An APTS op-amp, modified with gap version with DC coupled in-pixel architecture and 10 μm pitch, has been tested in laboratory in order to characterize this chip flavor. The pixels front-end response have been monitored by the measurement of the output at different chip bias settings. This allowed to evaluate the influence of the main parameters of the in-pixel circuit on the signal processing. Moreover, thanks to these results it has been possible to determine an optimal working point of the chip, defined by the best trade between a high output signal amplitude, a low value of signal fall time[6] and a high gain of the circuit at the output on-board. On the other hand, measurements with X-rays emitted by an ^{55}Fe source were also performed in order to study the detector charge collection properties.

5.5.3.1 Test system

Since no commercial readout system is available for the studied devices, a portable readout system (figure 5.18) was developed at INFN Cagliari in collaboration with CERN to characterize and test some of the chips of MLR1. The ALPIDE data acquisition (DAQ) board was used as a starting point and revised to allow it to also operate MLR1 chips. A PCIe connector provides 28 digital I/Os of adjustable logic voltage (1.5 to 3.3 V) to operate a proximity card and it hosts the additional components that were not part of the original ALPIDE DAQ board design. The actual chips are bonded to a custom carrier card (APTS OA version 2) connected to the proximity card via a PCIe connector that in turn carries power and digital or analogue signals.

The DAQ board measures 108 $mm \times 100$ mm and has at its heart an Altera Cyclone IV FPGA configurable via a JTAG interface or by an EEPROM (Electrically Erasable Programmable Read-Only Memory). Communication with a PC via the USB 3.0 interface is mediated by Infineon's EZ-USB FX3, a peripheral controller that provides a configurable, parallel General Programmable Interface (GPIF II) to the FPGA. In order to program the FPGA via USB, a patch cable must be added to the DAQ board to bridge the J8 ("UART") and

[6]The signals have a negative polarity because of negative charge collection on the n-well electrode

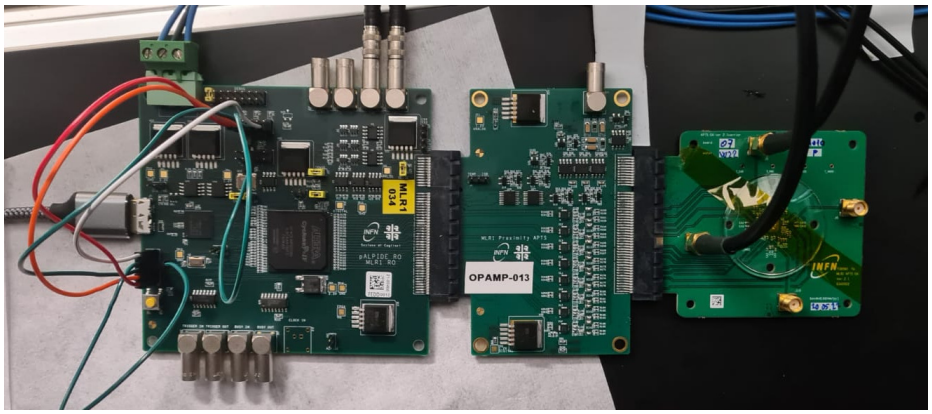


Figure 5.18: APTS op-amp test system: FPGA-based DAQ board + chip-specific proximity card + chip carrier card.

J3 ("JTAG") connectors. The board hosts a 40 MHz oscillator that drives the PLLs[7] (phase-locked loop) in the FPGA, where four clock domains are then derived, among which a 160 MHz clock is fed to the module that operates the ADCs for the APTS readout. The board can be powered either over the USB bus or through its 5V power connector by an external power supply. The FPGA reads out the chip data, formats it to events consisting of 32-bit words that include custom headers and trailers, e.g., timestamps, word count, event id. It then ships it via USB over the FX3 interface. I/O LEMO connectors are used to apply back-bias voltage and operate the system on external trigger and busy signals.

A specific proximity card (OPAMP-013) provides voltage/current biasing and ADC conversion, as the DAQ board only supports sensors with digital readout. Two on-board 8-channel 16-bit DACs provide direct voltage biases or voltage inputs to current sources and sinks that provide current biases to the chip. The rest of the proximity card provides chip specific functionality. For APTS, readout capabilities are provided by eight dual channel 16-bit ADCs for parallel readout of all the 16 analogue pixels up to 4 MSa/s (Megasamples per second). The readout is controlled by the FPGA and can be triggered either via software, externally via the appropriate LEMO input, or when the signal exceeds a threshold set by the user. The chip can also be pulsed, i.e. the FPGA can trigger a charge injection in one or more pixels for all sorts of functional tests.

For the op-amp version, the readout requires a bandwidth in the GHz range which would have required drastic modifications to the design of the ALPIDE DAQ board. It is then performed externally via an oscilloscope plugged into

[7]The phase-locked loop (PLL) block is a feedback control system that automatically adjusts the phase of a locally generated signal to match the phase of an input signal. PLLs operate by producing an oscillator frequency to match the frequency of an input signal.

four SMA outputs located on the carrier card, used to read the outputs from the four internal pixels of the chip pixel matrix. A scheme of the 4 x 4 pixel matrix with the identifying number of each pixel is shown in figure 5.19, the four inner pixels (highlighted) correspond to numbers 5, 6, 9 and 10. For the experimental setup dedicated to the measurements presented in this work, a Rohde & Schwarz RTO 1044 oscilloscope was employed, featuring a bandwidth of 4 GHz and a sampling rate of 20 GSa/s. The SMA outputs on the carrier where connected to the four BNC adaptors on the oscilloscope to monitor simultaneously the signals from the inner chip pixels with the four scope channels.

An HMP4040 Rohde & Schwarz power supply was used to provide the external power to the DAQ board (5 V) and the back-bias voltage to the sensor substrate and deep p-wells (kept at the same potential in this work).

0	1	2	3
4	5	6	7
8	9	10	11
12	13	14	15

Figure 5.19: Schematic view of the APTS op-amp 4 x 4 pixel matrix with the identifying number used for each pixel. The highlighted correspond to the four inner pixels connected to the oscilloscope.

5.5.3.2 Chip response tests

The chip qualification starts with a test pulsing to verify the chip response to an inner pulsing signal. As mentioned above, charge is injected in the collection electrode through the capacitance $C_{inj} = 242\,aF$ using the test circuitry integrated on each pixel. A voltage difference (V_h) = -1.2 V (value recommended by circuit designers) was sent to the hole matrix (represented in figure 5.20, left) via software through the DAC registers and the outputs from the external pixels, read through the ADC channels, were monitored. The outputs from the external 12 pixels submatrix are shown in figure 5.20 (right). The back bias voltage was set to 0 V. As observed in the figure, the signals have amplitudes of $\sim 23\ mV$ and to calculate the values in mV it was considered that 1 ADC unit is equal to 38 μV. The straight lines at 0 V in the plots represent the absence of output from the four internal pixels (figure 5.20, left), which can be read only externally with the oscilloscope.

The output signals from the pixels of the inner submatrix were monitored using the oscilloscope. The figure 5.21 shows the waveforms displayed on the oscilloscope channels connected to the SMA outputs on the carrier board. The sampling period (time window on the screen) was set to 10 μs and the trigger was fixed 10 mV below the signal baseline. As observed, the signal in the four channels has an amplitude of $\sim 24\,mV$ at $V_{bb} = 0\,V$.

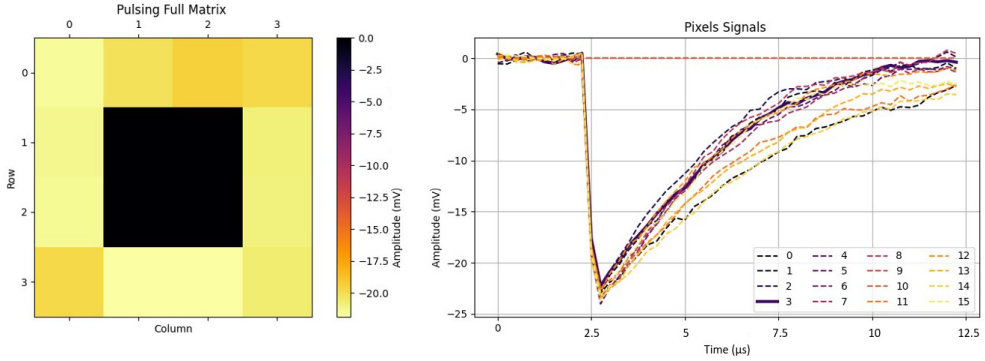


Figure 5.20: (Left) Scheme of the APTS op-amp pixel matrix and amplitude values measured after a charge injection. The 2 x 2 internal pixels (colored in black) can be read only externally with an oscilloscope and the rest of the matrix is read through the ADC channels in the proximity board. (Right) Output from the external 12 pixels, read by the ADC channels, after injecting charge in their collection electrodes through the capacitance present in their test circuits.

Gain calibration An important characteristic on the evaluation of the chip performance is the behavior of the output signal with respect to the reset voltage (VRESET), which is used to adjust the level of the signal baseline. The rate of change of the output signal with respect to VRESET depends on the gain of the system, it is a measure of how much the output changes for a given change in the input, that is, it tells us how sensitive the output signal is to changes in VRESET. The gain of a system is typically expressed as a ratio of the output to the input signal, so it will determine how much the output changes in response to changes in input parameters. The amount by which the output signal changes in response to a small change in the VRESET voltage is mathematically equivalent to taking the derivative of the signal at the output with respect to VRESET. If the baseline is represented by the function $y = g(VRESET)$, where VRESET is the variable that controls its value, then the rate of change of the baseline value with respect to VRESET is given by the derivative $\frac{\partial g}{\partial VRESET}$. As such, this quantity has been established by convention as a description of the gain of the circuit under study.

Figure 5.22 (left) shows the dependence between the value of the baseline measured at the output of the 16 pixels in the matrix of the APTS op-amp chip and the value of reset voltage applied at the input of the front-end circuit. To perform this measurement the VRESET was varied on a range from 20 to 900 mV with steps of 10 mV. At each VRESET value, the signal waveforms were read from the 12 external pixels using the data provided by the ADC channels on the proximity board, one ADC unit is equivalent to 0.0381 mV. The waveform of the 4 internal pixels were extracted from the information given by the oscilloscope channels (in mV units) connected to the SMA outputs on

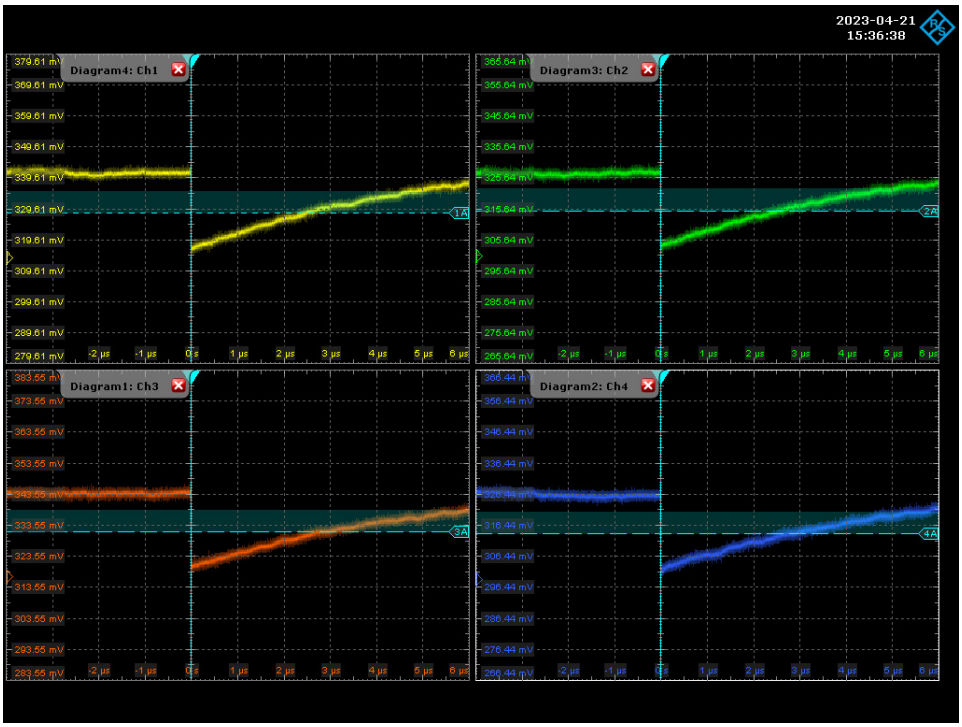


Figure 5.21: Signal waveforms displayed on the oscilloscope channels connected to the SMA outputs on the carrier board for the four innermost pixels.

the carrier board as descrived previously. After each setting of VRESET, 100 readout events were sent via software to collect good statistics. The average baseline value was calculated as the mean of the points in the 100 waveforms collected and the associated error was taken as the standard deviation of such distribution [$\sim 2.5\,mV$]. As observed in the plot, the baseline tends to increase almost linearly in the central region of the VRESET range for all pixels; for values lower than $\sim$ 100 mV and higher than $\sim$ 600 mV the increment is smoother. Setting VRESET to higher or lower values will cause ultimately the output to saturate either because the PMOS follower M2 output goes against the voltage AVDD or the NMOS follower M3 output approaches the ground, creating a minimum and a maximum voltage for VRESET in between which the circuit can be operated normally without saturation effects.

The derivative of the baseline with respect to VRESET, represented in figure 5.22 (right), was computed for each pixel as follows: the first nine points on the curve are fitted to a straight line, whose slope corresponds to the derivative at the central selected point (5th point). Then, starting from the second point on the curve, the next nine points are selected and fitted to calculate the derivative at the second central point (6th point) and so on. This procedure is repeated until the end of the curve. It can be noticed that using this method the derivative

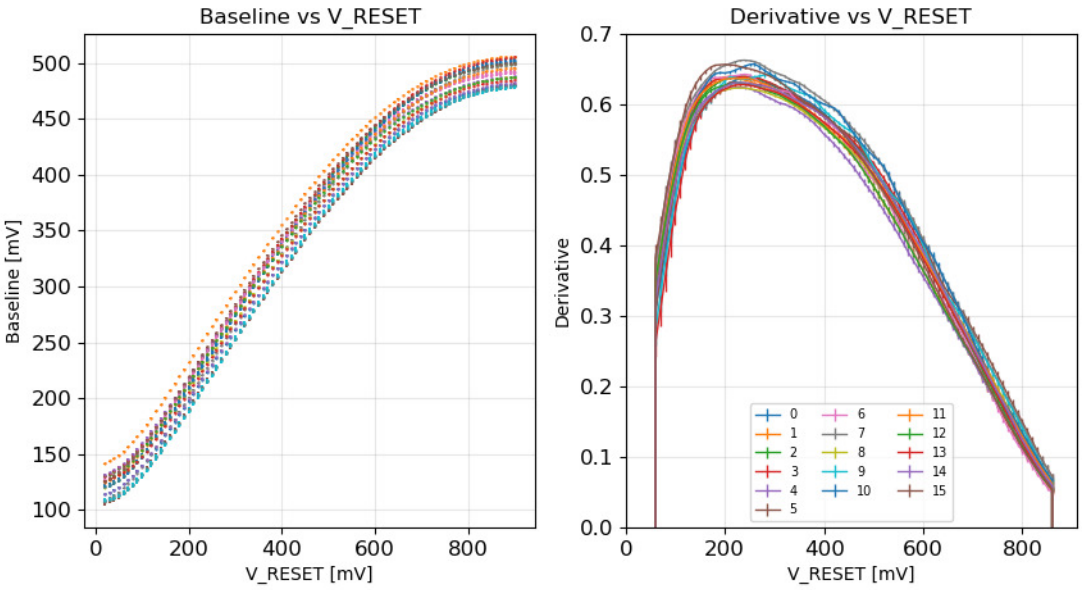


Figure 5.22: (Left) Baseline values of the signals measured at the output of the 16 pixels in the matrix of the APTS op-amp chip vs value of reset voltage (VRESET) applied at the input of the front-end circuit. (Right) Derivative of the baseline values with respect to VRESET.

is not calculated for the first and the last four points, but this does not represent a problem since in these regions the dependency of baseline and VRESET is not linear, i.e. the circuit does not have a stable behavior. As shown in the plot, the derivative values for the 4 inner pixels (see figure 5.19) present fluctuations between 0.6 and 0.7, which could be related to additional noise introduced to the signal through the components associated with the oscilloscope. In general, the curves reach their highest stability in a range of VRESET around 260 $\pm$ 100 mV, which corresponds to derivatives higher than 0.6. This would represent the range of VRESET where the circuit can be operated optimally for what concerns the gain.

Pulsing calibration After finding the range of highest stability of the pixels response, the circuit calibration goes on. A further test is realized in order to look for an optimal chip configuration which assures the best functioning for each pixel, i.e. an optimal balance between signal amplitude and fast response. For these measurements, the range of VRESET values found for each pixel in the gain calibration test was considered. At each VRESET, 100 internal pulses are sent to the matrix and the waveforms corresponding to the pixels response are read from the oscilloscope. The trigger was fixed 10 mV below the baselines.

The waveforms were analyzed externally to calculate the amplitude and the fall time of the signals. Figure 5.23 shows an example of a signal extracted from the oscilloscope. The amplitude (ΔV) is calculated as the difference between the mean of the points in the baseline and the mean of the points after the signal voltage drop (lowline). The fall time is defined as the time difference between the points at 10% and 50% of the voltage drop; this definition corresponds to a convention established by the different groups in charge of the APTS op-amp chip characterization. These calculations can be done accurately only for the pixels read by the oscilloscope, given its high sampling rate (20 GSa/s) which is able to record precisely the fast pixels response. As observed in figure 5.23, for a stricter computation of the variables, the time scale of the oscilloscope was increased, i.e. the sampling period was reduced to 50 ns, allowing to visualize in detail only the signal falling edge. The points in figures 5.24 and 5.25 represent the mean of the amplitude and fall time values for the 100 waveforms recorded at each VRESET for different back-bias voltages (V_{bb}). In figure 5.24 amplitudes variate from 20 mV, at $V_{bb} = 0\,V$ to $\sim 70\,mV$ at $V_{bb} = -4.8\,V$.

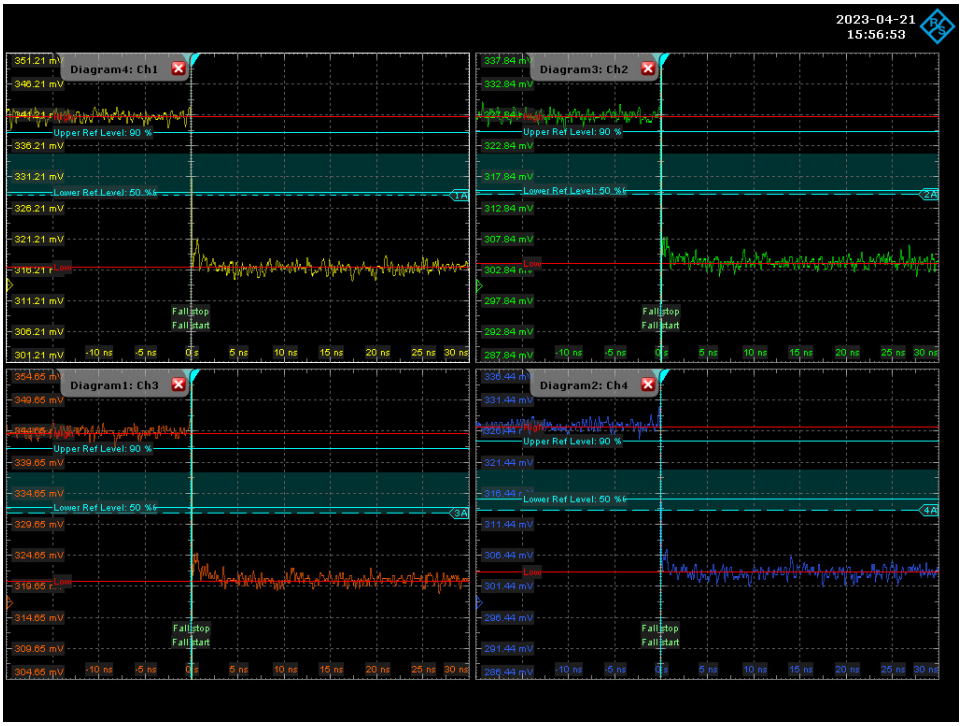


Figure 5.23: Signal waveforms of the four inner pixels read by the oscilloscope. The lines displayed are the reference levels for the calculation of the amplitude and fall time.

In figure 5.25 an increase of the fall times is observed at higher V_{bb}, ranging from $\sim 70\,ps$ at $V_{bb} = 0\,V$ to $\sim 130\,ps$ at $V_{bb} = -4.8\,V$. At 0 V the points

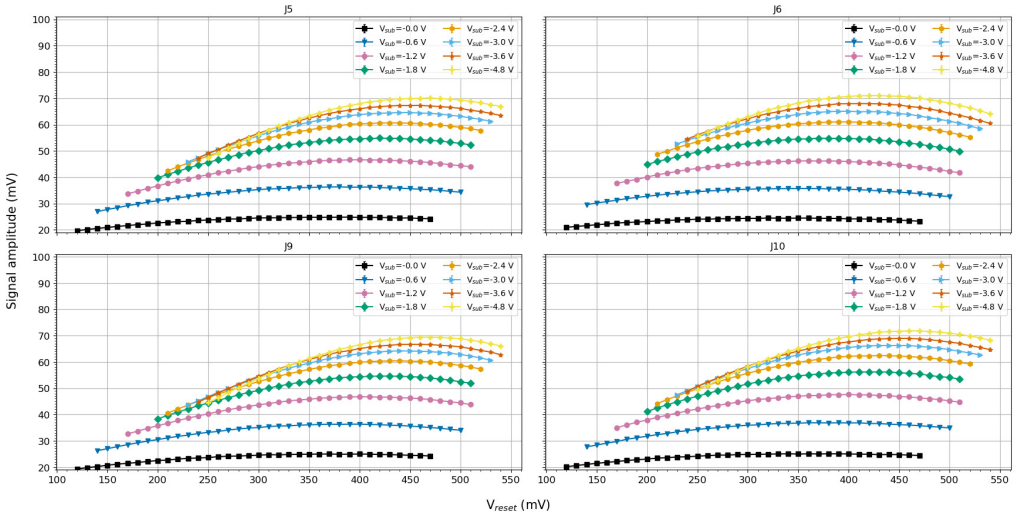


Figure 5.24: Signal amplitude as function of VRESET at different V_{bb} for the four inner pixels (read by the oscilloscope) of the APTS op-amp chip.

have a stable behavior through the VRESET range, while from -0.6 V on, the curves show a decreasing trend and seem to converge, at the last VRESETs, to fall times in the range between 80 and 110 ps.

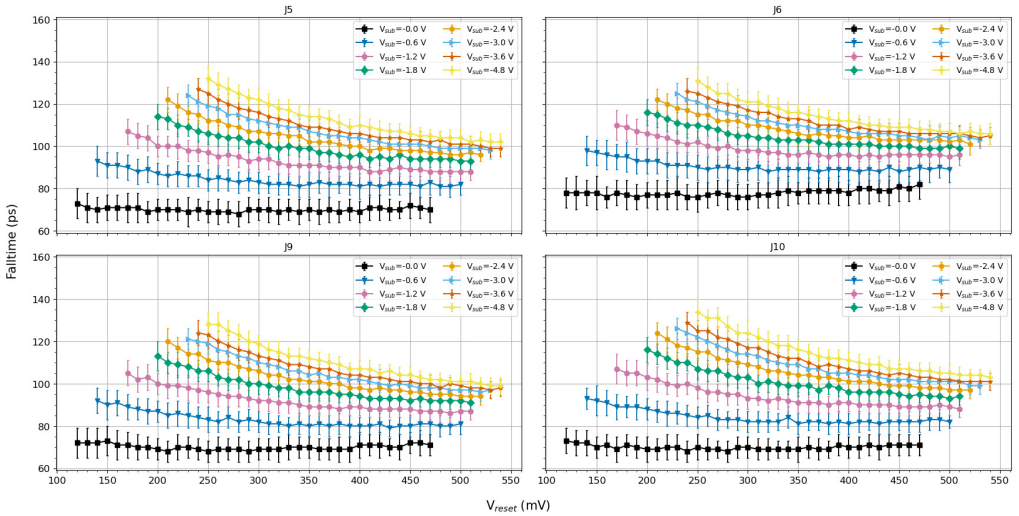


Figure 5.25: Signal fall time as function of VRESET at different V_{bb} for the four inner pixels (read by the oscilloscope) of the APTS op-amp chip.

Having at hand the results reported in figures 5.22, 5.24 and 5.25, the ratio

between the amplitudes and fall times is calculated as $\frac{dV}{dt}$ and shown in figure 5.26. Since an optimal response is expected as a trade off between highest amplitude and lowest fall time, the best VRESET is selected as the one with the maximum $\frac{dV}{dt}$ on each curve. Additionally, the VRESETs corresponding to the maximum gain values are taken, which means that, in the end, two VRESET values are chosen for each inner pixel and each V_{bb}. Since the pixels contain their own front end and their output are read independently, the responses observed are slightly different from pixel to pixel; consequently a different set of VRESETs is found for each pixel at the same V_{bb}. In order to uniform the results, i.e. to determine a common optimal working point for all the matrix, at every V_{bb} the best VRESET values of the four inner pixels are averaged. In this way an unique value of VRESET is established as optimal for the entire pixel matrix at different settings of V_{bb}. Table 5.4 resumes the results obtained from the described calibration procedure performed to the APTS op-amp chip under study.

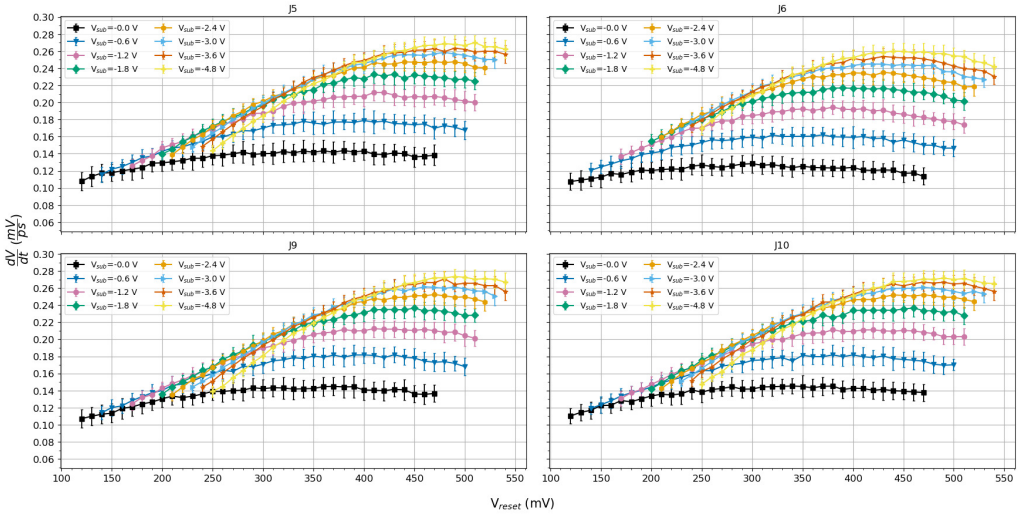


Figure 5.26: Derivative of amplitude with respect to fall time.

Vh scan Investigation into the front-end response of the chip using the in-pixel pulsing circuitry involved also the study of the signal amplitude and fall time through the variation of Vh, which controls the amount of injected charge. For this measurement, the settings of VRESET, found to be optimal from gain and pulsing calibrations (see table 5.4), were applied to the matrix. Vh was varied between 200 and 1200 mV in steps of 100 mV. An increment of the injected charge is expected to increase the amplitude and the length of the signal produced. In figure 5.27 an increasing linear trend of the signal amplitude with Vh is observed, which indicates the linearity of the pixels response and its consistency with the expected behavior. This is confirmed also in figure

Table 5.4: Best working point determined from the gain and pulsing calibration of the APTS op-amp chip under study.

V_{bb} (V)	VRESET (mV)
0.0	290
-0.6	330
-1.2	370
-1.8	390
-2.4	410
-3.0	420
-3.6	430
-4.8	440

5.28, where the dependence of fall time with respect to the variation of Vh is presented. The increment of the charge and consequently of the amplitude, leads to a longer duration of the pulse, as can be noticed in the fall time trends.

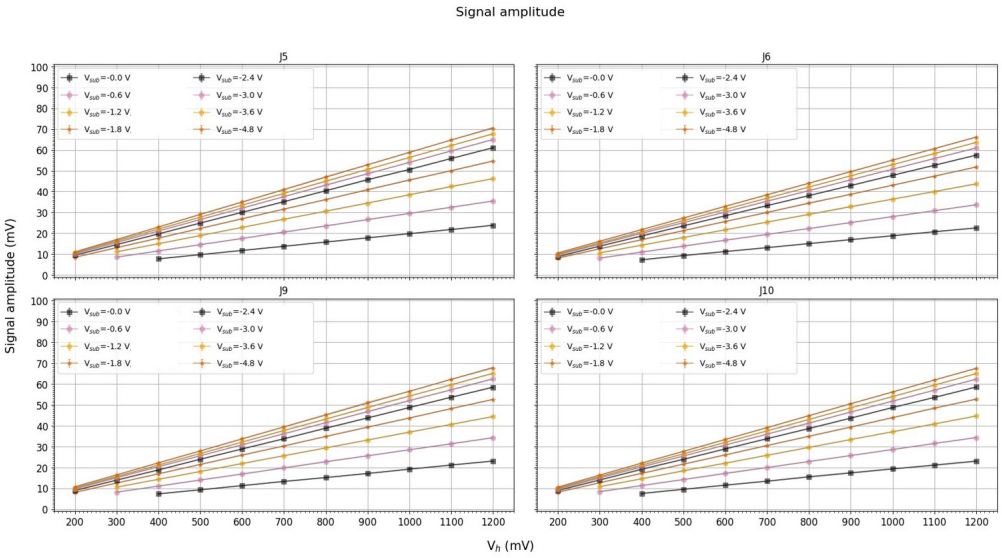


Figure 5.27: Signal amplitude as function of Vh (injected charge) at different V_{bb} for the four inner pixels (read by the oscilloscope) of the APTS op-amp chip.

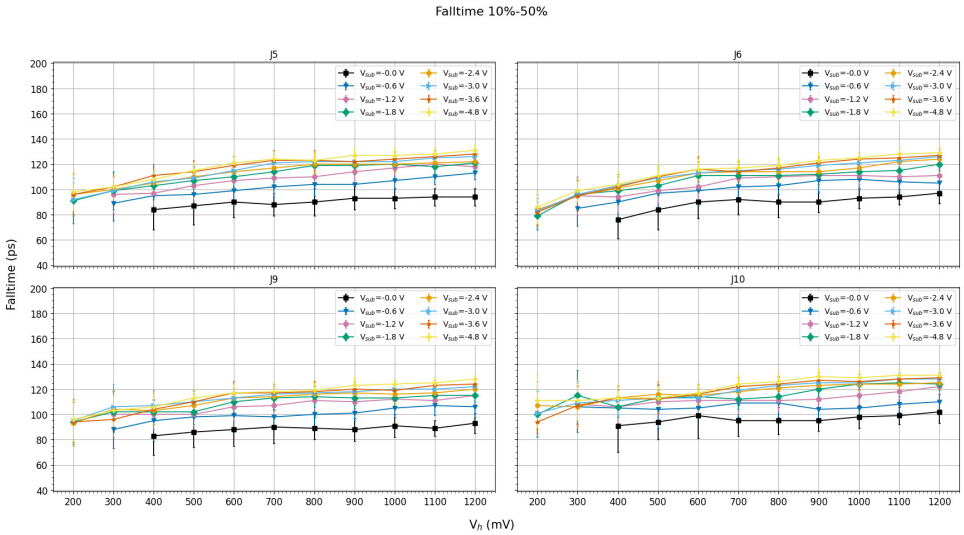


Figure 5.28: Signal fall time as function of Vh (injected charge) at different V_{bb} for the four inner pixels (read by the oscilloscope) of the APTS op-amp chip.

5.5.3.3 Measurements with X-rays emitted by an ^{55}Fe source

A ^{55}Fe source is commonly used in characterization of MAPS, since the emitted X-rays deposit an amount of charge similar to that of a minimum ionizing particle orthogonally incident to the sensitive volume. $^{55}_{26}Fe_{29}$ disintegrates by electron capture to ^{55}Mn with half life of $T_{1/2} = 2.737$ years [114]. After the vacancy in the K shell left by the nuclear-captured electron is filled by an electron from a higher shell. The difference in energy is released by emitting Auger electrons of 5.19 keV, with a probability of about 60%, $K_{\alpha-1}$ X-rays with energy of 5.89875 keV and a probability about 16.2%, $K_{\alpha-2}$ X-rays with energy of 5.88765 keV and a probability of about 8.2%, or K_{β} X-rays with nominal energy of 6.49045 keV and a probability about 2.85%. The energies of the $K_{\alpha-1}$ and $K_{\alpha-2}$ X-rays are so similar that they are often specified as mono-energetic radiation with 5.9 keV photon energy. Its probability is about 28% [115]. The remaining 12% is accounted for by lower-energy Auger electrons and a few photons from other minor transitions.

A 5-6 keV electron has a range of $\sim$ 2.9 g/cm^2 [116] in air, that is 2.4 mm. The distance between the ^{55}Fe source and the detector studied in this work was about 2 cm. For distances between the source and detector of about few centimeters, all emitted X-rays can reach the detector. The main interaction process for soft X-rays is the photoelectric absorption. The electron emitted by the photoelectric absorption event generates electron-hole (e-h) pairs along its path (the range of a 6 keV electron in silicon is $\lessapprox$ 1 μm [117]). An average energy of 3.6 eV is required to produce an e-h pair. Therefore 5.9 and 6.5 keV photoelectrons create on the average $\approx$ 1640 and $\approx$ 1800 e-h pairs, respectively.

The intrinsic energy resolution is given by $R = 2.35\sqrt{Fw/E}$ where F is the Fano factor in silicon ($F \approx 0.116$ for 5.895 keV [118]), w the energy required to excite an e-h pair and E the energy deposited by a photoelectron. Thus $R \approx 2\%$ ($\approx 120\,eV$) for energies between 5.9 and 6.5 keV and therefore 5.888 and 5.899 keV (6.490 and 6.535 keV) X-rays can't be resolved. Consequently, it is expected to observe only two peaks from the photoelectric absorption of 5.9 and 6.5 keV photons in silicon.

The primary regions for the photoelectric conversion on the detector geometry are:

1. Depleted region (high electric field): since the range of the photoelectron is $\lessapprox 1\,\mu m$, it is possible that all the e-h pairs are created inside the depleted region and therefore collected by a single pixel.

2. Undepleted epitaxial layer (low or no electric field): electrons remain trapped in the epitaxial layer and diffuse thermally until they reach a depleted region (or until they recombine) and are collected by the n-well diodes. Therefore the deposited charge is shared between several pixels.

3. Substrate: electrons diffuse thermally in the substrate. Because of the potential difference between p^{++} and p^{-} type silicon, those electrons which reach the epitaxial layer are trapped there and eventually collected.

In the first two cases the entire charge deposited by the photoelectron is expected to be collected. In the last case the amount of collected charge is generally smaller than the deposited charge, since it depends on the depth of the photoelectric conversion and the carrier lifetime in the substrate. The resulting spectrum would be a continuous distribution up to the 5.9 keV peak. Photoelectric conversions in other regions, e.g. circuitry, would also contribute to this spectrum.

The detector charge collection characteristics are reflected in the seed pixel[8] signal distribution. It is assumed that photoelectric conversions in the high electric field (depleted) region directly below collection diodes will result in clusters where only the seed pixel collects charge. Consequently, the seed pixel signal of these events corresponds to a known charge deposit.

Figure 5.29 shows the seed pixel ^{55}Fe spectra measured for the 2x2 inner pixels of two versions of 10 μm pitch APTS op-amp chips at $V_{bb} = -4V$. The blue line corresponds to the result from the standard type APTS op-amp, while the orange line represents the one from the variant with a low-dose n-type implant with a gap (see figure 5.7). In particular in the spectrum of the modified version, two peaks can be well recognized at higher amplitudes corresponding to the conversions of 5.9 and 6.5 keV X-rays in the depleted region. The first is often referred as calibration peak since it is clearly distinguished and its position can be used to calibrate the sensor response in electrons. In the modified with gap version (orange curve) the calibration peak is the most dominant

[8]The seed pixel is the pixel with the largest collected charge in a cluster (a set of adjacent pixels which collected an amount of charge higher than a certain threshold).

contribution to the distribution, since this type enables to collect almost the total charge produced by the incident particle in a single pixel, resulting in a highly efficient charge collection. Instead, the standard type (blue curve) shows a different distribution with a smaller calibration peak and more events in the amplitudes range below 80 mV; this indicates the presence of signals shared by multiple pixels, i.e. a superior amount of charge sharing in the pixel matrix.

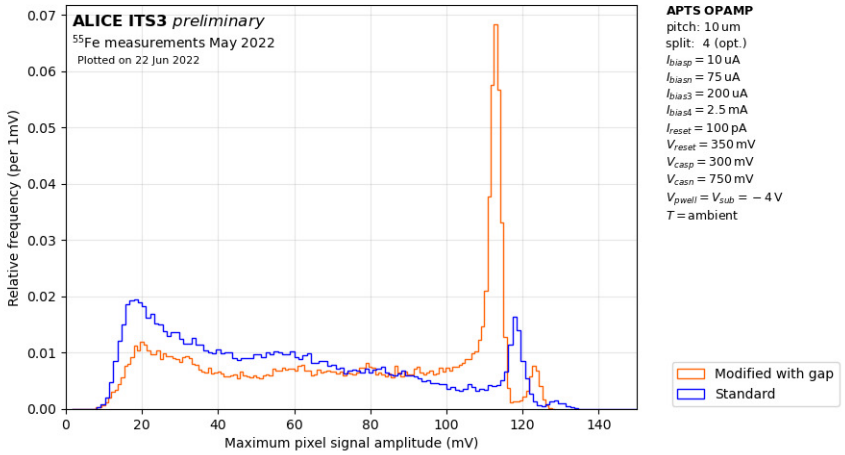


Figure 5.29: ^{55}Fe spectra from the seed pixel in the 2x2 inner matrix of an APTS op-amp standard type (blue) and a process with gap variant (orange) [119].

Conclusions

The next ALICE detector upgrade foresees the replacement of the three innermost layers of the current Inner Tracking System (ITS2) with an entirely new detector (ITS3) with increased vertexing and tracking performance, specially for particles at low transverse momentum ($p_T < 1\,GeV/c$). The new ITS inner barrel will thus consist on a vertex detector structured by three cylindrical concentric layers composed by curved wafer-scale silicon sensors, built on CMOS Monolithic Active Pixel Sensors (MAPS). The new detector requirements are focused on: 1) the demand for improved vertex resolution, which will be beneficial for the measurement at low transverse momentum of many of the observables studied by ALICE; 2) a refined impact parameter resolution and 3) a reduced material budget, meaning the enhancement of the physics capability for the study of heavy-flavour production, notably in the baryon sector, and for the study of low-mass dielectrons, two of the main objectives of the ALICE physics programme in the next decade.

The detection unity of the current ITS2 is the ALPIDE (ALice PIxel DEtector) chip, based on MAPS technology and fabricated with the commercial 180 nm CMOS imaging sensor process by TowerJazz. It has reached unprecedented performance in terms of signal/noise ratio, spatial resolution, material budget and readout speed. Single ALPIDE chips have been taken under study as reference to investigate the viability of bending a thin silicon device to be used for the scope of the next ITS3.

In order to verify if the electronic circuitry on the ALPIDE preserves its functionalities after the mechanical stress of bending it, I carried out a test campaign in laboratory dedicated to the electrical characterization and comparison of the results before and after bending a chip. The evolution of the threshold and fake-hit rate were measured over time, maintaining invariant the in-pixel front-end DACs. The threshold trend was more stable for the bent geometry and shifted to values 0.5 DAC ($5e^-$) higher with respect to the initial flat configuration. The values of analogue current consumed during the chip operation were also higher for the bent geometry and consistent with results reported previously. Additionally, the pixel matrix showed to be $\sim 0.1e^-$ less noisy when the chip was bent, product of the higher thresholds, which contributes to mask part of the noise. The noise occupancy (fake-hit rate) was in the order of $10^{-7} pixel^{-1} event^{-1}$ for all geometries, but slightly higher for the bent position, indicating the increment of noise on a small group of pixels. The

effect of masking the noisiest pixels was the reduction of fake-hit rate down to $\sim 10^{-10} pixel^{-1} event^{-1}$. For the bent configuration 5 pixels more were masked than the flat position, nevertheless, the fraction of masked pixels was of the order of 10^{-5} in both cases.

On the other hand, the observables were measured over the variation of the front-end parameters. The trends showed the expected increment for THR vs ITHR and decrease for THR vs VCASN, with threshold values moderately higher in the case of bent chip. The noise exhibited a rising trend with ITHR, whose increment leads to a higher level of current in the corresponding front-end transistor. The same behavior was observed with VCASN, whose augmentation results in lower threshold and consequently the measurement of higher noise. For what concerns the fake-hit rate, it decreased with ITHR, given that pixel signals have to cross a higher threshold level in order to record a hit; with VCASN increment threshold values drop, giving rise to higher fake-hit rate, as resulted from the measurements. In general, there were not remarkable differences between the values of threshold, noise and fake-hit rate measured with the ALPIDE chip when it was positioned in flat and bent postures.

An important component in the investigation of a detector performance is the realization of a so-called test-beam experiment. The first ever in-beam characterization of a bent MAPS was performed in a test-beam campaign at the DESY II test-beam facility in June 2020. As part of this activity I was dedicated to the analysis of data acquired from the measurements with a bent ALPIDE and their systematic comparison with data taken during the same campaign, but with a different telescope featuring a flat ALPIDE as DUT. The study had the main scope to evaluate the sensing performance of a bent chip through the variation of the working point defined by front-end parameters and contrast it with the results of an originally flat sensor. The study of cluster size revealed a strong dependence between this quantity and the sensor threshold, given that the charge generated by a particle traversing the sensor and shared to an adjacent pixel has to overcome the in-pixel threshold in order to be registered as a hit. As such, the decrease of the cluster sizes with the increment of the threshold values was observed for flat and bent sensors. In addition, the profiles of cluster sizes along the rows direction showed that, for the bent ALPIDE, clusters become larger with increasing the chip curvature, which confirms that particles featuring a larger incident angle produce larger clusters, given their increased path in the active volume of the detector and the generation of more free charge carriers along their trajectories. The measurements on efficiency showed a decreasing trend with respect to threshold, because at higher threshold, even though a particle passed through the DUT no hits are registered and no cluster can be associated to the corresponding accepted track, thus resulting in a decreased efficiency. Both flat and bent chips showed efficiencies higher than 99.7% over the full threshold range, which indicated that the bending had no substantial impact on the ALPIDE performance.

A further test-beam was performed in July 2021 at CERN SPS facility. I was dedicated to analyze a part of the data collected during this campaign, where a set of Devices Under Test was composed by six ALPIDE chips, bent

at the foreseen ITS3 layer radii, forming the so-called micro-ITS3. The objective was the simultaneous in-beam characterization of several concentrically arranged bent ALPIDE chips, thus modeling the tracking layers of the future ITS3. The residuals distributions of the bent ALPIDEs resulted to be mostly Gaussian and centered at zero, which confirmed the reaching of good alignment between layers, validating eventually the possibility to align the six bent ALPIDEs using the software Corryvreckan dedicated to the test-beam data analysis. The detection efficiency vs threshold showed the expected trend for all bent sensors, with values above 98% between $80e^-$ and $160e^-$, starting to decrease towards higher threshold values as observed also in the previous test-beam data. Around the nominal working point of $100e^-$ the efficiencies are approximately of 99%. This represents an evidence that a system of ALPIDE chips, bent in a configuration similar to that of the future ITS3 layers, maintain an optimal detection efficiency over a large sensor working point range. Furthermore, the presence of a minimum in the standard deviation of the residual distributions indicated a working point between $100e^-$ and $200e^-$ where the position resolution is optimal for all the bent ALPIDEs. As resulted from the electrical characterization, for a certain threshold value there is a minimum also in the noise level, this fact indicates that the optimal position resolution is directly related to the noise level of the respective DUT.

Additional measurements were carried out by introducing a copper target at the center of the detector barrel of the micro-ITS3 in order to mimic particles emerging from a real collision. My contribution to this activity was the preliminary analysis of data collected with the telescope using the data processing and alignment functionalities of Corryvreckan, as the previous step to perform a further study of tracks produced in the forward layers by the particles of the beam interacting with the target. The second part of the activity was carried out by colleagues of the group in Bari, Italy. From the experimental setup, the three downstream bent layers were used along with the additional flat ALPIDE chips placed at larger distance from the target to reconstruct the tracks produced in the interactions. The symmetric configuration, placed upstream of the target was used to reconstruct the tracks of the beam particles. The distribution of DCA of the downstream tracks to the primary interaction vertex in the copper target was obtained and compared with a Monte Carlo simulation. The average pointing capability of the system was estimated, in particular, the average resolution of the DCA was measured. The aim of this effort is to validate the operation and test the tracking capability of a detector system with a configuration analogue to the ITS3 system.

The innovative idea of the ITS3 upgrade is to make use of the stitching technology to produce a single sensor based on thin wafer-scale Monolithic Active Pixel Sensors bent in cylindrical shape around the beam pipe, where the power distribution is managed internally, confining the interconnections to the outside world to the sensor edge. In parallel with the design and characterization of new sensor prototypes on which the stitching will be implemented, development of support structures, flex for powering and communication, procedures and tools for the assembly and integration of the layers are carried out. The

"super-ALPIDE" project has the goal to realize a demonstrator of a detector with geometry similar to one half-layer of the ITS3, using a working sensor with large dimensions made of a matrix of not diced ALPIDE chips. Interconnections of the sensors in the super-ALPIDE matrix to the external world are made by using flexible printed circuits (FPC) wire bonded to the pads on the sensor matrix in order to provide an homogeneous power distribution and allow configuration and readout. The super-ALPIDE necessitates of two FPCs. A first one, at the edge of the detector, mimics the configuration that will be implemented in the final detector. A second one is necessary only for this demonstrator to reach the ALPIDE pixels far from the edge, and it is held by an exoskeleton. The main steps adopted for the bending of the large dimension sensor have been exhaustively described. The developed technique foresees the usage of a rotating cylindrical mandrel around which the large silicon sensor is wrapped, guided by a transparent mylar foil, kept in tension parallel to the mandrel longitudinal axis. In the procedure, two vacuum plates made of a porous material are needed to perform a precise sensor alignment and placing on the mandrel. The realization of the super-ALPIDE project has been on charge of mechanical engineers, technical operators and other member collaborators of the Bari group. This work has allowed for the acquisition of significant experience on handling ultra thin large scale sensors, by the development of instruments and techniques for bending and mechanical mounting of these large dimensions structures.

For the development of the ALICE ITS3 upgrade, efforts have been made to port the features of the TJ 180 nm process to the 65 nm CMOS imaging flavor of TPSCo, the first choice of technology for the implementation of the wafer-scale sensor. The key benefit of MAPS detectors to be developed in the TPSCo 65 nm, will be higher spatial resolution and embedment of processing, such as clustering, time-of-arrival or possibly even amplitude measurements in pixels. A first submission in TPSCo 65 nm has been done as a Multi Layer Reticle (MLR1), consisting on 55 small prototype structures in different processing flavors as a test of the technology. A description on the structure and performance of the available sensor prototypes based on the 65 nm technology within the context of the ITS3 upgrade R&D (Digital Pixel Test Structure (DPTS), CE-65 and Analogue Pixel Test Structures (APTS), source follower and op-amp versions.) has been provided in the chapter 5 of this thesis.

My contribution to this activity was focused on the realization of laboratory measurements to test the performance of an APTS op-amp version. Initially, the behavior of the signal baseline at the output of the 16 pixels in the matrix with respect to the reset voltage (VRESET) applied at the input of the front-end circuit, was monitored. The baseline tends to increase almost linearly in a VRESET range between 100 mV and 600 mV at $V_{bb} = 0V$. The derivative of the baseline with respect to VRESET represents a measure of the gain of the circuit; at $V_{bb} = 0\,V$ it reached the highest stability for VRESET $\approx 260 \pm 100\ mV$, which is the range where the circuit can be operated optimally for what concerns the gain and considered for the rest of the calibration. The values of amplitudes and fall times were computed in the found VRESET arrays for the individual pixels at different V_{bb}. The amplitudes varied from 20 mV, at $V_{bb} = 0\,V$ to

$\sim 70\,mV$ at $V_{bb} = -4.8\,V$. An increase of the fall times was observed at higher V_{bb}, ranging from $\sim 70\,ps$ at $V_{bb} = 0\,V$ to $\sim 130\,ps$ at $V_{bb} = -4.8\,V$. In order to determine the optimal working point of the pixel matrix at each V_{bb}, the values of VRESET with the maximum ratio $\frac{\mathrm{d}V}{\mathrm{d}t}$ and the maximum gain were selected for all the pixels and averaged in order to establish an unique optimal setting for the entire sensor. Applying the found VRESETs, the signal amplitude and fall time were studied through the variation of Vh, which controls the amount of injected charge. An increasing linear trend of the signal amplitude with Vh was observed, indicating the linearity of the pixels response and its consistency with the expected behavior. The increment of the charge and consequently of the amplitude, leads to a longer duration of the pulse; this was confirmed by the dependence of fall time with respect to Vh. Finally, results of the detector charge collection properties, investigated through the seed pixel ^{55}Fe spectra measured for the 2x2 inner pixels of two versions of 10 μm pitch APTS op-amp chips (standard and modified with gap) at $V_{bb} = -4V$, have been presented. In the spectrum of the modified version, the peak at 5.9 keV (calibration peak) is dominant, since this type enables to collect almost the total charge produced by the incident particle in a single pixel, resulting in a highly efficient charge collection. For the standard type a smaller calibration peak and more events in the range of amplitudes below 80 mV are observed, indicating the presence of signals shared by multiple pixels, i.e. a superior amount of charge sharing in the standard type pixel matrix.

The installation of the new ITS3 is planned to take place during the next LHC Long Shutdown 3, between 2026 and 2028. In the context of the ALICE program, an intensive R&D effort is currently ongoing and some milestones have already been achieved. The Multi-Layer Reticle 1 (MLR1) was received in summer 2021 and as discussed previously, it has the scope to verify the TPSCo 65 nm technology both in terms of radiation hardness and pixel performance for high-energy physics applications, including a large number of dies in different processing flavors that provided crucial insight and input to design larger and more complex structures. The following Engineering Run 1 (ER1) submission was completed in late 2022 and first processed wafers have been delivered in May 2023; it includes two large stitched sensors named Monolithic Stitched Sensor (MOSS) and Monolithic Stitched Sensor with Timing (MOST). It is built by joining ten sub-units only limited in length by the size of the wafer, measuring a total of approximately 25 cm. Both have digital front-ends with a different approach in the readout design. Two additional submissions labeled Engineering Run 2 (ER2) and Engineering Run 3 (ER3) are in the planning stage and will include the first full-size sensors as actual prototypes of the final detector (including its readout) for ITS3.

Appendix A

Interaction of radiation with matter and Silicon properties

A.1 Energy loss for charged particles

A charged particle passing through matter undergoes soft electromagnetic processes with the electrons of the medium. Different types of collision can lead to atomic excitation, displacing an electron to a higher atomic bound state, or ionization, creating an electron-hole pair if the transferred energy is higher than the ionization potential of the atom. For silicon, the mean energy required to create a single e-h pair is about 3.6 eV [120], which is in fact more than three times larger than the band gap (1.12 eV). The difference goes into the generation of phonons and eventually will dissipate as thermal energy. Ionization represents the predominant way for charged particles to lose energy in matter. The principle of solid state detectors is based on the energy loss of ionizing particles in the sensor material. The motion of the electron-hole (e-h) pairs created induces a signal current on their respective collection electrodes.

For particles heavier than electrons, the mean rate of ionization energy loss for charged particles, also known as stopping power, is given with good approximation by the Bethe-Bloch formula [121]:

$$\langle \frac{\mathrm{d}E}{\mathrm{d}x} \rangle = 2\pi N_A r_e^2 m_e c^2 \frac{Z}{A} \frac{z^2}{\beta} \left[\ln \left(\frac{2m_e \gamma^2 \upsilon^2 W}{I^2} \right) - 2\beta^2 - \delta - 2\frac{C}{Z} \right] \qquad \text{(A.1)}$$

where E is the kinetic energy of the incident particle of charge z, moving with velocity β and Lorentz factor $\gamma = 1/\sqrt{1-\beta^2}$. $I \approx 10Z^{0.9}$ eV is the mean excitation energy of the material (137 eV for silicon), characterized by density ρ and by atomic and mass number Z (atomic number of the absorption medium,

14 for silicon) and A (atomic mass of the absorption medium, 28 g/mol for silicon), respectively. W is the largest kinetic energy that could be transferred to a free electron in a single collision. δ and C are correction factors accounting for the polarization of the medium. $N_A = 6.022 \cdot 1023\ mol^{-1}$ is Avogadro's number, r_e and m_e are classical electron radius and mass.

According to the Bethe-Bloch formula, the ionization loss is proportional to the electron density in the medium $\frac{\rho Z N_A}{A}$ and to the square of the projectile charge, and it strongly depends on the projectile velocity. For non relativistic particles the energy loss rate decreases with β proportionally to $1/\beta^2$.

Figure A.1 shows that in different materials the energy loss in $MeV \cdot cm^2/g$ for heavy charged particles has almost the same dependence on particle momentum. At $1 - 2\ MeV \cdot cm^2/g$ the energy loss reaches its minimum in all materials. A particle with energy at the minimum of this formula is called minimum ionizing particle (MIP): relativistic particles have mean energy loss rates close to the minimum. In silicon, minimum ionization occurs at $\beta\gamma = 3$, corresponding to a value of $\langle \mathrm{d}E/\mathrm{d}x \rangle_{min} \sim 1.66\ MeV \cdot cm^2/g$.

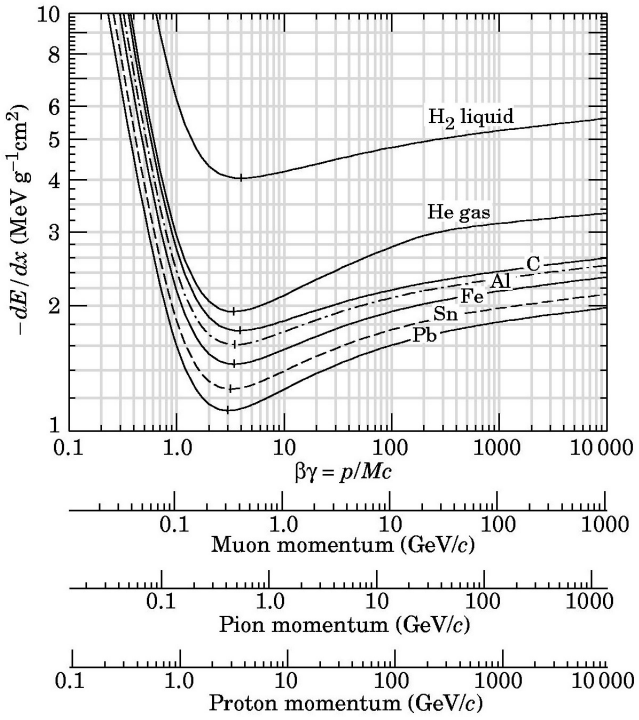


Figure A.1: Bethe-Bloch formula for different materials [122].

The Bethe-Bloch formula only provides the mean energy loss $\langle \Delta \rangle$ suffered by a charged particle passing through a layer of material. The amount of energy

lost Δ for a particle in a single crossing, however, is subject to statistical fluctuations because of the variations in the number of collisions and in the energy transferred in each collision. The probability density function of the energy loss (straggling function) depends on the thickness of the absorbing medium [123]: for sufficiently thick layers such that the number of collisions is large the central limit theorem is applicable and the distribution shows a Gaussian shape around the mean energy loss. In thin layers with less collisions occurring, the situation changes and the fluctuations become larger. Landau was the first to calculate this [124] and hence the probability density function $f(\Delta)$ of the energy loss in thin layers is often referred to as Landau distribution. The Landau distribution shows a long upper tail due to rare, but highly ionizing knock-on or δ-electrons, which obtain sufficient energy during the interaction to become ionizing particles themselves (Figure A.2). As a result of this tail, the mean energy loss $\langle\Delta\rangle$ is larger than the most probable value for the energy loss (MPV or ΔMPV). The scale factor between ΔMPV and the mean energy loss depends on the particle energy, but is typically around 1.3.

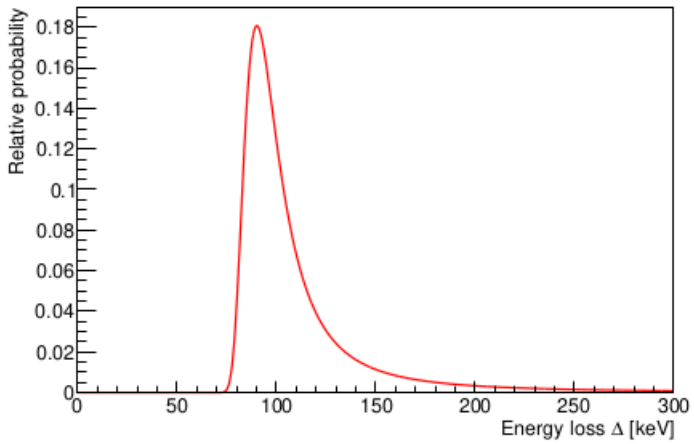


Figure A.2: Landau distribution of the energy loss of highly relativistic particles ($\beta\gamma \gg 100$) in a silicon layer of 320 μm thickness. [58].

Moreover, for very thin sensor layers ($\lesssim$ 300 μm for silicon), as used in MAPS, the Landau model fails to describe the energy loss accurately [125]. In such cases an improved model by Bichsel can be used, which provides good agreement with experimental data down to 10 μm [126]. For a silicon layer with a thickness of 20 μm the Bichsel model yields $\Delta_p = 4.12\ keV$ (most probable value), corresponding to 1140 e-h pairs or about 60 e-h pairs per μm, and $\omega = 3.34\ keV$ (full width at half maximum). The Landau model overestimates the MPV of the energy-loss distribution for thicknesses below about 160 μm, but significantly underestimates its width.

The way in which electrons interact with matter is mostly through Coulomb

force with the orbital electrons of the absorber atom; since their mass is equal to the orbital electrons mass, when they interact they can lose a greater amount of energy (if compared to heavy particles) and be deviated from their path in a more significant way. Moreover their small mass causes the emission of electromagnetic radiation (bremsstrahlung) to be a significant term in the energy loss. This radiation comes from the acceleration of the incoming electron when it is deflected via its electromagnetic interaction with the nucleus.

The total energy loss for electrons is then composed of two parts:

$$\frac{\mathrm{d}E}{\mathrm{d}x} = \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{coll} + \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{bremss} \tag{A.2}$$

The radiative loss for the electrons becomes significant when the energies increase over a certain critical energy E_c:

$$E_c \sim \frac{(610 - \ 700\ MeV)}{Z} \tag{A.3}$$

which is the critical value at which collision energy loss and radiative loss are the same. For electrons of a few MeV the radiative losses are a small fraction of the total energy losses; the radiative term is more significant in materials with high atomic number (it varies as Z^3). Introducing a quantity dependent on the the material, the radiation length X_0 (usually g/cm^2), the radiative term can be written as:

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{bremss} = \frac{E}{X_0} \tag{A.4}$$

with E energy of the incident particle. The radiation length is defined as the mean distance over which the electron energy is reduced by a factor $1/e$ by bremsstrahlung. The energy loss by bremsstrahlung, varies as the inverse of the particle squared mass:

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{bremss} \propto \frac{1}{m^2} \tag{A.5}$$

This explains the negligibility of radiative term as the mass increases (as for heavy charged particles); for muons, the radiative loss is $4 \cdot 10^4$ times smaller than that of electrons.

A.2 Interaction of photons with matter

Photons mainly interact with matter in three ways, depending on their energy: photoelectric effect, Compton scattering, and pair production. Photon interaction is deeply different from charged particles; they are much more penetrating

in the matter and a beam of photons, passing through a thickness, is just attenuated in intensity, keeping the same energy. The intensity of the beam I related to the initial intensity I_0, as a function of the crossed thickness of the material x is:

$$\frac{I}{I_0} = e^{-\alpha x} \tag{A.6}$$

with α absorption coefficient, which is characteristic of the absorbing material and can be defined as:

$$\alpha = \frac{\sigma N_A \rho}{A} = \frac{1}{\lambda} \tag{A.7}$$

with $\sigma = \sigma_{photoelectric} + \sigma_{compton} + \sigma_{pair}$ the total cross section, N_A Avogadro's number, ρ the density of the material and λ the mean free path, is defined as the average distance traveled in the absorber before an interaction takes place.

A.2.1 Photoelectric effect

The photoelectric effect arises in the absorption of a photon by an atom of the absorber material; the impinging photon disappears and is replaced by a photo-electron; the atomic nucleus recoils (due to momentum conservation). The binding energy of the electron is $E_{binding}$ (energy threshold) so that its final energy E will be:

$$E = h\nu - E_{binding} \tag{A.8}$$

with $h\nu$ photon energy.

The photoelectric cross section is dominant for low energy photons and is enhanced by high atomic number Z:

$$\sigma \propto \frac{Z^k}{(h\nu)^n} \tag{A.9}$$

with k between 4 and 5 and $n \sim 3.5$.

A.2.2 Compton scattering

The Compton scattering is instead characterised by the presence of the photon at the end of the interaction. The photon is scattered and deviated from its original direction of an angle θ, proportional to the amount of energy transferred to the electron. If $h\nu$ and $h\nu'$ are the photon energy before and after the collision:

$$\frac{h\nu'}{h\nu} = \frac{1}{1 + \frac{h\nu}{m_e c^2 (1 - \cos\theta)}} \tag{A.10}$$

with $m_e c^2$ the rest mass energy of the electron. The cross section is in general enhanced at lower energy and is proportional to Z:

$$\sigma_{compton} \propto Z \tag{A.11}$$

A.2.3 Pair production

The pair production process has an energy threshold equal to 1.022 MeV which is twice the mass of an electron; indeed this effect involves the transformation of a photon into an electron-positron pair:

$$h\nu > 2m_e c^2 \sim 1.022\ MeV \tag{A.12}$$

The pair production cross section is dominant at high energy and, above the energy threshold, is proportional to Z^2:

$$\sigma_{pair} \propto Z^2 \tag{A.13}$$

The positron emitted in this process quickly annihilates in two γ.

In figure A.3 a comprehensive illustration of the relative importance of the three processes described for different absorber materials and photon energies is shown.

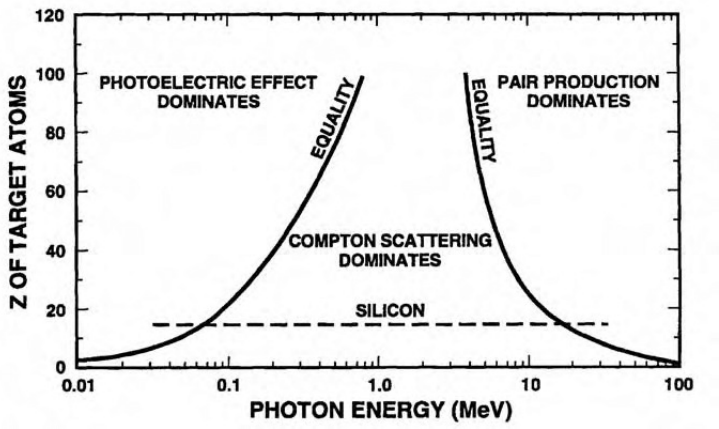


Figure A.3: Regions of predominance of the three main forms of photon interaction with matter. The curves indicate when the cross sections of the two main processes involved (Photoelectric effect and Compton Scattering on the left hand side curve and Compton Scattering and Pair production on the right hand side curve) are equal [127].

A.3 Silicon properties

Silicon is a semiconductor, namely a solid which can be roughly considered as an insulator at low temperatures and a conductor at higher temperatures. Its

conductivity assumes values of about $10^3\ \Omega \cdot m$ and it is intermediate between that of insulators and that of conductors. In a silicon crystal, at a temperature of $0\ K$ there are no free electrons, while with increasing temperature the number of free electrons increases and there is a finite probability that an electron in the lattice will be knocked loosing from its position and leaving behind an electron deficiency called a "hole". The energy distribution of these conduction electrons is governed by the Fermi-Dirac distribution and is given by:

$$f(E,T) = \frac{1}{1 + e^{\frac{E - E_f}{kT}}} \tag{A.14}$$

where E is the electron energy, E_F the Fermi energy, namely the energy of the electron in the highest occupied state at $0\ K$, T is the temperature of the system and k the Boltzmann constant. If an external voltage is applied, both the electron and the hole can contribute to a small current flow. Since for $T > 0\ K$ electrons can reach the conduction band and contribute to electrical conduction, at room temperature electrons occupy the conduction band and can recombine with holes. Electrons which have been freed from their lattice positions into the conduction band can move through the material. In addition, other electrons can hop between lattice positions to fill the vacancies left by the freed electrons. This additional mechanism is called hole conduction because it is as if the holes are migrating across the material in the direction opposite to the free electron movement. The current flowing in an intrinsic semiconductor is thus due to electrons and holes. The term "intrinsic" distinguishes between the properties of pure silicon and those of doped n-type or p-type material. The carrier density for intrinsic silicon is the same for electrons and holes and its dependence on temperature is given by the function:

$$n_i = T^{3/2} exp\left(- \frac{E_{gap}}{2kT} \right) \tag{A.15}$$

where $n_i = n = p = 1.45 \cdot 10^{10} cm^{-3}$ at $T = 300\ K$ is the carrier density [128] (n for electrons and p for holes) and E_{gap} is the energy of the forbidden gap separating valence and conduction bands ($1.12\ eV$ for silicon).

A.3.1 Doping silicon

By adding a small percentage of foreign atoms in the crystal lattice of silicon, it is possible to change its electrical properties, producing n-type and p-type semiconductors. The addition of these impurities to the semiconductor lattice is called doping. To obtain a n-type material, silicon is doped with pentavalent atoms (i.e. donors with one more valence electron) such as antimony, arsenic or phosphorous, which contribute extra electrons. These electrons form energy levels near the conduction band, greatly increasing the conductivity of the intrinsic semiconductor. A p-type material, instead, is obtained with the addition of trivalent impurities (i.e. acceptors with one less valence electron) such as

boron, aluminum or gallium, creates holes which form energy levels near the valence band, with an increase of the conductivity. A doped semiconductor is called "extrinsic". In thermal equilibrium, the concentration of positive and negative charge carriers is constant in time and follows the mass action law [128]

$$np = n_i^2 \tag{A.16}$$

which says that the product of the concentration of electrons and holes is equal to the square of the intrinsic carrier density and it is not altered by doping. In n-type silicon electrons are majority and holes minority carriers, while in p-type silicon electrons are minority and holes are majority carriers [129].

A.3.2 Drift

Charge transportation in semiconductors is mainly due to drift and diffusion mechanisms. Drift is achieved by applying an external electric field across the material, so that the charge carriers start to drift with a certain velocity which is proportional to the applied electric field. By introducing electrons and holes mobilities as:

$$\mu_n = \frac{e \cdot \tau_n}{m_n} \qquad \mu_p = \frac{e \cdot \tau_p}{m_p} \tag{A.17}$$

with e the electron charge, $\tau_{n,p}$ the free mean time between successive collisions and $m_{n,p}$ the effective mass, drift velocities are:

$$\overrightarrow{v_e} = -\mu_n \overrightarrow{E} \quad \text{for electrons,} \tag{A.18a}$$

$$\overrightarrow{v_h} = \mu_p \overrightarrow{E} \quad \text{for holes,} \tag{A.18b}$$

Mobility depends on dopant and charge carrier concentrations, temperature and electric field. In silicon, mobilities can be considered constant for electric field values $E < 1\ kV/cm$: at $300\ K$, $\mu_n = 1350\ cm^2V^{-1}s^{-1}$ and $\mu_p = 450\ cm^2V^{-1}s^{-1}$. As the electric field becomes higher, mobility is more and more field-dependent, while the drift velocity tends to become constant, reaching the saturation value. Note that holes have a lower mobility value compared to electrons, leading to a drift velocity reduced by about a factor 3, when not in saturation. The motion of electrons toward the positive terminal and holes to the negative one induces a current, the so called drift current, whose total value is:

$$J_{TOT} = \mu_n Enq + \mu_p Epq \tag{A.19}$$

The total induced drift current increases with the increasing of E until it reaches the saturation value. Drift velocity in silicon saturates at about $v_{sat} \sim 10^7\ cm/s$ for both electrons and holes (Figure A.4). To be more precise, holes velocity saturates later and the value is slightly lower ($\sim 9.5 \cdot 10^6\ cm/s$).

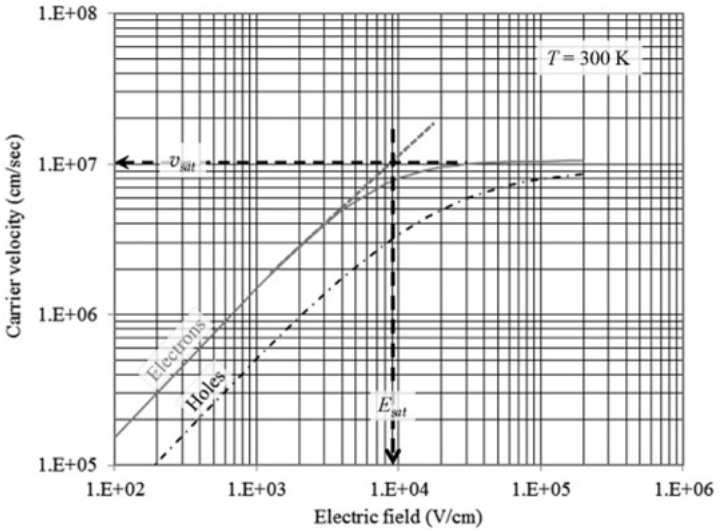


Figure A.4: Drift velocities of electrons and holes in silicon at room temperature as a function of applied electric field. [130].

A.3.3 Diffusion

Even without an external electric field applied, charge carriers can move inside the semiconductor following a concentration gradient. Diffusion consists in fact in the motion of charged particles from high to low concentration regions. The resulting diffusion currents are then [131]

$$J_n = qD_n \frac{\mathrm{d}n}{\mathrm{d}x} \quad \text{for electrons,} \tag{A.20a}$$

$$J_p = -qD_p \frac{\mathrm{d}p}{\mathrm{d}x} \quad \text{for holes.} \tag{A.20b}$$

where D_n and Dp are the diffusion constants, $\mathrm{d}n/\mathrm{d}x$ and $\mathrm{d}p/\mathrm{d}x$ are the concentration gradients for electrons and holes, respectively. Einstein's relation for diffusion relates diffusion constant and mobility to the absolute temperature value for each type of carriers:

$$\frac{D}{\mu} = \frac{kT}{q} \tag{A.21}$$

Since electrons and holes are free charge carriers, the currents induced by their drift and diffusion motions are affected by an intrinsic noise, whose magnitude is higher than the expected signal for a particle detection. Consequently, intrinsic semiconductors are not effective as particle detectors. For this purpose, they have to be doped and then depleted from free carriers: this can be achieved by a pn junction.

A.3.4 p-n junction

A p-n junction is a structure which shows diode characteristics. It consists in a p-type and a n-type silicon sections bonded together, with electrodes on each end.

A.3.4.1 Zero bias

When no voltage is applied to the diode, a diffusion motion establishes between the junction: the presence of a region with excess of electrons and lack of holes bonded to a region with excess of holes and lack of electrons causes a gradient of concentration. Hence electrons from the n-type material tend to diffuse into the p region filling the holes of the positive side of the junction and leaving fixed positive ions behind them. Likewise, holes from the p-type side diffuse toward the n side leaving a region with negative fixed ions in the p side: this creates in proximity of the p-n interface a region without free charge carriers called depletion zone or space charge zone. Since ions are fixed and do not move, in a depletion zone the semiconductor material is returned to its original insulating state and charge cannot flow. The depletion zone gives then rise to an electric field which opposes the diffusion process for both electrons and holes (Figure A.5 a). In equilibrium, the potential which establishes in the depletion zone is called built-in potential

$$V_{bi} = \frac{k_B T}{q} \ln \frac{N_A N_D}{n_i^2} \tag{A.22}$$

and depends on the dopant concentrations, N_A for acceptors and N_D for donors. There are two concurrent phenomena: the diffusion process that tends to generate more space charge, and the electric field generated by the space charge that tends to counteract the diffusion. The built-in potential in silicon V_{bi} takes a value of about 0.7 V.

A.3.4.2 Forward bias

To bias a diode, a potential difference has to be applied to the junction. The first way to bias a pn junction is directly (forward bias): the n side is connected to a potential lower than the one at the p side (Figure A.5 c). The n side, with respect to the p side, is then at a negative potential. When forward biased, the positive potential in the p-type region repels the holes and the negative potential in the n-type region repels the electrons: the width of the depletion zone decreases and the built-in potential value lowers. On the contrary, the number of minority carriers increases in both regions, leading to an increase of the diffusion currents. With increasing forward-bias voltage, the depletion zone eventually becomes thin enough that the built-in potential cannot counteract charge carrier motion across the pn junction, consequently reducing electrical resistance. The electrons which cross the pn junction into the p-type material (or holes which cross into the n-type material) will diffuse in the near-neutral

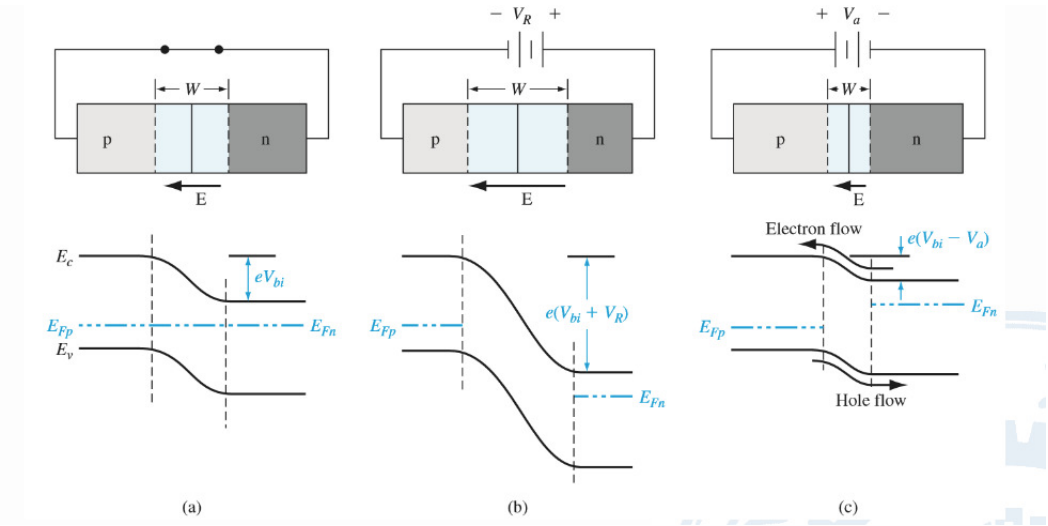


Figure A.5: pn junction and its associated energy band diagram for a) zero bias, b) reverse bias and c) forward bias [132].

region. Therefore, the amount of minority diffusion in the near-neutral zones determines the amount of current that may flow through the diode.

A.3.4.3 Reverse bias

Reverse bias is achieved by applying a negative potential to the p-type material and a positive potential to the n-type (A.5 b). This removes holes and electrons from the junction, therefore the width of the depletion zone increases. By increasing the reverse-bias voltage V_R, the depletion zone W widens further:

$$W = x_n + x_p = \sqrt{\frac{2\varepsilon_0\varepsilon_{Si}}{e}\left(\frac{1}{N_A} + \frac{1}{N_D}\right)(V_{bi} + V_R)} \tag{A.23}$$

where x_n and x_p are respectively the width of the depletion zone in the n and p sides, ε_0 is the absolute dielectric constant and ε_{Si} is the one relative to silicon. Since a silicon sensor is typically formed by a highly doped p+ implant ($N_A \sim 10^{18}\ cm^{-3}$) on a low doped n bulk ($N_D \sim 10^{12}\ cm^{-3}$), one finds that the width of the depletion region is deeply shifted to the less doped side (n, in this case), also leading to a negligible built-in voltage value:

$$W \approx\ x_n \approx \sqrt{\frac{2\varepsilon_0\varepsilon_{Si}}{eN_D}V_R} \tag{A.24}$$

By defining the resistivity as

$$\rho = \frac{1}{eN_D\mu} \tag{A.25}$$

the width of the depletion zone becomes:

$$W = \sqrt{2\varepsilon_0\varepsilon_{Si}\rho\mu V_R} \tag{A.26}$$

where μ is the mobility (for electrons it is $\mu_e = 1350\ cm^2/(Vs)$, while for holes it is $\mu_h = 450\ cm^2/(Vs)$. The whole diode is depleted at a voltage called full depletion voltage V_{FD}

$$V_{FD} = \frac{d^2}{2\varepsilon_0\varepsilon_{Si}\rho} \tag{A.27}$$

being d the total thickness of the sensor. When the electric field intensity increases beyond a critical level, the pn junction depletion zone breaks down and current begins to flow, usually by either the Zener or avalanche breakdown processes. Both of these breakdown processes are non-destructive and are reversible, as long as the amount of current flowing does not reach levels that cause the semiconductor material to overheat causing thermal damage. The $I(V)$ characteristic for an ideal silicon pn-diode is shown in figure A.6.

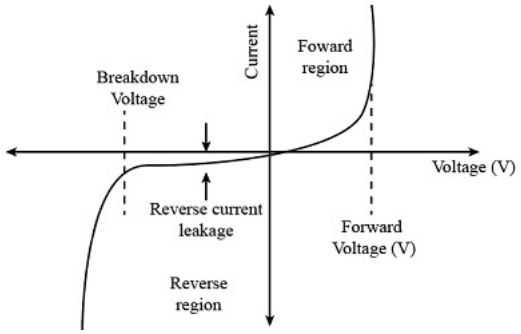


Figure A.6: Current vs voltage characteristic for a silicon diode in both forward and reverse bias directions [122].

A.3.5 Piezoresistivity

Piezoresistivity is a material property that couples bulk electrical resistivity to mechanical strain [133]. In other words, as a material is stretched, compressed, or distorted in any way, the electrical resistivity changes. Almost all materials have this property, even though in some it is so small that it is virtually undetectable. However, other materials show a large change in resistivity for a

relatively small strain. One of these materials is silicon. The change in resistance can be related to a physical phenomenon such as displacement, pressure, or force. The electric field vector of an anisotropic cubic crystal is related to the current vector by a resistivity tensor [134].

$$\begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix} = \begin{bmatrix} \rho_1 & \rho_6 & \rho_5 \\ \rho_6 & \rho_2 & \rho_4 \\ \rho_5 & \rho_4 & \rho_3 \end{bmatrix} \begin{bmatrix} J_1 \\ J_2 \\ J_3 \end{bmatrix} \tag{A.28}$$

The resistivity tensor is symmetric and is therefore made up of only six components. When the crystal is in an unstressed state, the resistivity components along the diagonal of the matrix, which represent the resistivity along the $\langle 100 \rangle$ axes of the cubic, are equal to each other. The other three components are also equal to each other and are null.

$$\rho_1 = \rho_2 = \rho_3 = \rho \qquad \rho_4 = \rho_5 = \rho_6 = \rho \tag{A.29}$$

When the cubic is stressed, each component of resistivity changes. The new resistivity of each component in the resistivity tensor is:

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \\ \rho_5 \\ \rho_6 \end{bmatrix} = \begin{bmatrix} \rho \\ \rho \\ \rho \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} \Delta\rho_1 \\ \Delta\rho_2 \\ \Delta\rho_3 \\ \Delta\rho_4 \\ \Delta\rho_5 \\ \Delta\rho_6 \end{bmatrix} \tag{A.30}$$

The link relating change in resistance to stress is the Π matrix. The change in resistance is directly related to the stress on the object through piezoresistive coefficients by:

$$\frac{1}{\rho}\Delta\rho_i = [\pi_{ij}][\pi_{ij}] \tag{A.31}$$

where the stresses relate to those shown in figure A.7.

σ_1, σ_2, and σ_3 are along the axes of the cube and σ_4, σ_5, and σ_6 map to the shear stresses τ_1, τ_2, and τ_3. With six stress components and six resistivity components, the π matrix must be a six-by-six matrix. Therefore, 36 π coefficients are required to populate the matrix. However, for the cubic crystal structure of silicon, the matrix simplifies to

$$\begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{12} & 0 & 0 & 0 \\ \pi_{12} & \pi_{11} & \pi_{12} & 0 & 0 & 0 \\ \pi_{12} & \pi_{12} & \pi_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \pi_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \pi_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \pi_{44} \end{bmatrix} \tag{A.32}$$

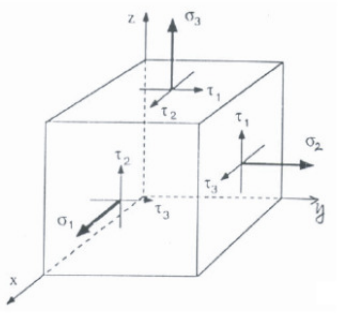


Figure A.7: Stress cube [135].

Only three π coefficients (π_{11}, π_{12}, and π_{44}) are needed. When all of the values of the components are filled in, equation A.31 becomes:

$$\frac{1}{\rho}\begin{bmatrix}\Delta\rho_1\\ \Delta\rho_2\\ \Delta\rho_3\\ \Delta\rho_4\\ \Delta\rho_5\\ \Delta\rho_6\end{bmatrix} = \begin{bmatrix}\pi_{11} & \pi_{12} & \pi_{12} & 0 & 0 & 0\\ \pi_{12} & \pi_{11} & \pi_{12} & 0 & 0 & 0\\ \pi_{12} & \pi_{12} & \pi_{11} & 0 & 0 & 0\\ 0 & 0 & 0 & \pi_{44} & 0 & 0\\ 0 & 0 & 0 & 0 & \pi_{44} & 0\\ 0 & 0 & 0 & 0 & 0 & \pi_{44}\end{bmatrix}\begin{bmatrix}\rho_1\\ \rho_2\\ \rho_3\\ \tau_1\\ \tau_2\\ \tau_3\end{bmatrix} \tag{A.33}$$

As shown in Refs. [136], [133], and [134], equations A.28, A.30, and A.33 can now be combined to obtain:

$$\begin{aligned} E_1 = &\ \rho J_1 + \rho\pi_{11}\sigma_1 J_1 + \rho\pi_{12}(\sigma_2+\sigma_3)J_1 + \rho\pi_{44}(J_2\tau_3 + J_3\tau_2)\\ E_2 = &\ \rho J_2 + \rho\pi_{11}\sigma_2 J_2 + \rho\pi_{12}(\sigma_1+\sigma_3)J_2 + \rho\pi_{44}(J_1\tau_3 + J_3\tau_1)\\ E_3 = &\ \rho J_3 + \rho\pi_{11}\sigma_3 J_3 + \rho\pi_{12}(\sigma_1+\sigma_2)J_3 + \rho\pi_{44}(J_1\tau_2 + J_2\tau_1) \end{aligned} \tag{A.34}$$

which describe the electric field potential along the crystal axes. The first term in each of the equations is the unstressed condition of the cubic. The next term takes in the effect of the stress along the axis of the respecting electric field component. The effect of the other two axial stresses are presented in the third term, and the last term models how the shear stresses effect the electrical potential.

The symmetry of the piezoresistance effects in silicon can be visualized readily in terms of the many-valley model of Herring [137]. In the $\langle 100\rangle$-valley model, which applies to electrons in silicon, the conduction electrons are imagined as lying in three equal groups, or valleys, aligned with the three (100) directions[1]. For any valley, the mobility is lowest in the valley direction (i.e.,$\langle 100\rangle$) and

[1]Actually there are six valleys, one for each positive and negative (100) direction, but for our purposes it is sufficient to imagine three.

highest normal to that direction, but the average mobility for all the valleys is isotropic. Tension in a valley direction removes electrons from that valley and transfers them to valleys lying normal to the tension. Hence, the average mobility becomes higher in the direction of tension (longitudinal effect) and lower in directions transverse to the axis of tension (transverse effect).

Many-valley energy surfaces for n-type silicon in k (wave vector) space are shown in figures A.8 and A.9 [138].

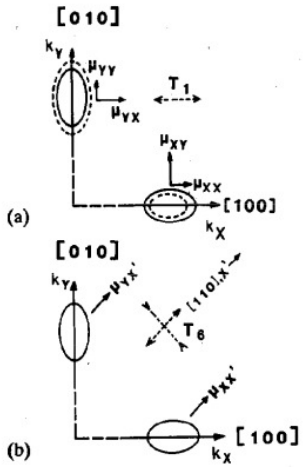


Figure A.8: Schematic diagram of the $\langle 100 \rangle$ and $\langle 010 \rangle$ valleys in k-space for n-Si. Dotted lines show the effect of stress. (a) corresponds to stress T_1 and (b) to T_6. [139].

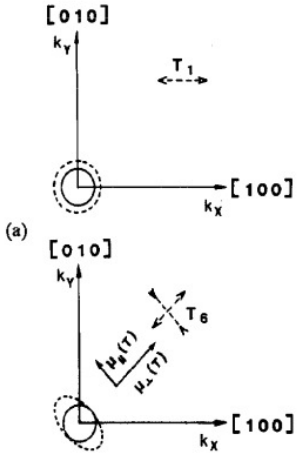


Figure A.9: Schematic diagram of the $\langle 001 \rangle$ valley in k-space for n-Si. Dotted lines show the effect of stress. (a) corresponds to stress T_1 and (b) to T_6. [139].

When a uniaxial tensile stress, T_1, is applied along the [100] axis, the band edges for the four valleys on the [010] and [001] axes go down and those on the [100] axes go up. The dotted lines in figures A.8 and A.9 show the positions and sizes of the surfaces for a certain energy. Electrons transfer from the high energy levels to the low energy levels. Thus, there are more electrons with the large mobility $\mu_\perp$ along the $\langle 100 \rangle$ direction. When an electric field is applied in this direction, it results in a decrease in resistivity or an increase in conductivity. Assuming that all the electrons of the donor levels are exhausted with and without stress, the total number of conduction electrons n_0 does not change. The longitudinal piezoresistance component π_{11} is given by [140, 141]

$$\pi_{11} = \frac{2\Xi_u}{3k_B\Theta} \frac{(s_{11} - s_{12})(1 - L)}{1 + 2L} \tag{A.35}$$

where Ξ_u is the shear deformation potential, k_B is Boltzmann's constant, Θ is the temperature in K, s_{ij} are the compliance constants, and $L = \mu_\perp/\mu_\parallel = m_\parallel \tau_\perp / m_\perp \tau_\parallel \equiv m_\parallel / m_\perp$, assuming the relaxation times $\tau_\perp = \tau_\parallel$. $\mu_\perp$ and $\mu_\parallel$ are the mobilities and $m_\parallel$ and $m_\perp$ are the effective masses of electrons in a single valley.

When an electric field is applied perpendicular to the direction of the applied stress, the transverse piezoresistance component π_{12} is given by:

$$\pi_{12} = -\pi_{11}/2 \tag{A.36}$$

When a shear stress, T_6, is applied in the plane normal to the (001), the band edges for the $\langle 100 \rangle$ and $\langle 010 \rangle$ valleys remain unchanged as shown by arrangement (b) in figure A.8, but the effective mass of the $\langle 001 \rangle$ valley changes as shown by

arrangement (b) in figure A.9 due to the special character of the conduction band edge of Si [142], i.e. under a shear strain the electron ellipsoid perpendicular to the shear plane will be distorted. The shear piezoresistance component π_{44} is given by [143]

$$\pi_{44} = \frac{-\, a_0 m_{\parallel} s_{44}}{1 + \, 2L} \tag{A.37}$$

where $a_0 = (86.8 \pm \ 5.0)/m$ and m is the free electron mass. Energy surfaces for p-type silicon in k space are very complex and different from those for n-Si, as shown in figure A.10 (a). The band edge, the upper $P_{3/2}$ state, consists of a pair of two-fold degenerate bands at $k = 0$ usually designated as the "light" and "heavy" hole bands. These energy surfaces are warped spheres. The lower $P_{1/2}$ state has a spherical energy surface. When a uniaxial tensile stress is applied parallel to the $\langle 111 \rangle$ direction, the degeneracy of the valence band is lifted and two bands of prolate and oblate ellipsoidal energy surfaces with anisotropic mass parameters are formed as shown in figure A.10 (b) [144, 145]. Consequently, the resistivity change comes from both the mass changes and hole transfer.

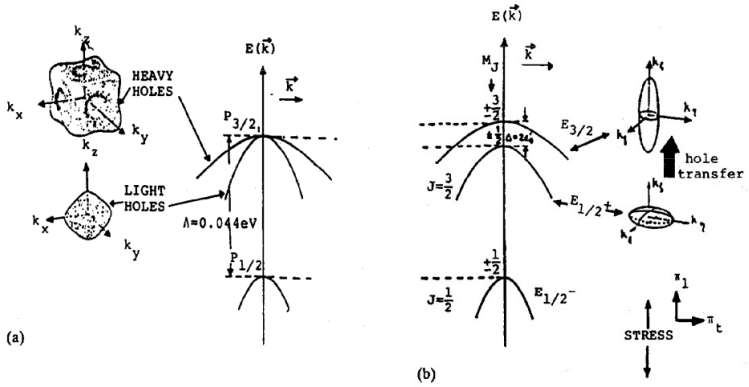


Figure A.10: Schematic diagram of the energy surfaces in k-space for p-Si (a) without stress and (b) under stress [139].

Even if the effective masses change, the density of state effective masses $m_i^{3/2}$ remain unchanged [144] and $m_1^{3/2}$ is approximately equal to $m_1^{3/2}$ [146]. Assuming that all holes of the acceptor levels are exhausted with and without stress, the total number of conduction holes p_0 does not change. The longitudinal piezoresistance coefficient is given by [146]

$$\pi_l = \frac{2\gamma_3 D'_u s_{44}}{3\gamma_1} \left(\frac{1}{k_B \Theta} + \frac{2}{\Lambda} \right) \tag{A.38}$$

where Λ is the spin-orbit splitting energy, D'_u is the deformation potential con-

stant and $\gamma_i s$ are the effective mass parameters. For the electric field perpendicular to $\langle 111 \rangle$, the transverse piezoresistance coefficient is also given by:

$$\pi_t = -\pi_l/2 \tag{A.39}$$

In silicon devices, piezoresistance can affect their performance and reliability in a number of ways. A relevant effect, as mentioned above, is the change in the material resistance, which can affect its ability to conduct electricity. Such response is dominated by changes in the charge carrier mobilities, which can lead to changes in the device's performance, such as sensitivity or accuracy. Overall, the impact of piezoresistance on silicon devices depends on the specific application and design. In some cases, the effect may be desirable, while in others it could be detrimental.

List of publications

- "First bent wafer-scale sensor in truly-cylindrical geometry for the ALICE ITS3 detector." (Proceeding published on EPJ Web of Conferences, volume 270, 2022)
- "Tracking performance of bent Monolithic Active Pixel Sensors mimicking a truly cylindrical barrel configuration." (status: under completion, submission end of 2023)
- "Technical Design Report of the Inner Tracking System 3." (status: under preparation, submission beginning 2024)
- "Time performance of Analogue Pixel Test Structures with in-pixel operational amplifier implemented in 65 nm CMOS imaging process." (status: under preparation)
- List of ALICE publications on peer reviewed journal.

Bibliography

[1] ALICE Collaboration. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002, 2008.

[2] ALICE Collaboration. Technical design report for the upgrade of the ALICE inner tracking system. *J. Phys. G: Nucl. Part. Phys.*, 41(8):087002, 2014.

[3] Musa Luciano. Letter of Intent for an ALICE ITS Upgrade in LS3. Technical report, CERN, Geneva, 2019.

[4] David J. Gross. The discovery of asymptotic freedom and the emergence of QCD. *Proc. Nat. Acad. Sci.*, 102:9099–9108, 2005.

[5] H. D. Politzer. Reliable Perturbative Results for Strong Interactions? *Phys. Rev. Lett.*, 30(26):1346–1349, 1973.

[6] B. L. Ioffe, V. S. Fadin, and L. N. Lipatov. *Quantum chromodynamics: Perturbative and nonperturbative aspects*. Cambridge University Press, 2010.

[7] J. Greensite. *An introduction to the confinement problem*, volume 821. Springer, 2011.

[8] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. Broken Symmetries. *Phys. Rev.*, 127:965–970, Aug 1962.

[9] Volker Koch. Aspects of Chiral Symmetry. *International Journal of Modern Physics E*, 06(02):203–249, jun 1997.

[10] H. Fukaya, S. Aoki, and S. Hashimoto et. al. Determination of the Chiral Condensate from (2+1)-Flavor Lattice QCD. *Phys. Rev. Lett.*, 104:122002, Mar 2010.

[11] N. Cabibbo and G. Parisi. Exponential hadronic spectrum and quark liberation. *Physics Letters B*, 59(1):67–69, 1975.

[12] Shuryak and Edward V. Theory of hadronic plasma. *Sov. Phys. JETP*, 47:212–219, 1978.

[13] A. Bazavov, Tanmoy Bhattacharya, and C. DeTar et. al. The equation of state in $(2+1)$-flavor QCD. *Physical Review D*, 90(9), nov 2014.

[14] Peter F. Kolb and Ulrich Heinz. Hydrodynamic description of ultrarelativistic heavy-ion collisions, 2003.

[15] Edward V. Shuryak and Ismail Zahed. Toward a theory of binary bound states in the quark-gluon plasma. *Physical Review D*, 70(5), sep 2004.

[16] Wit Busza, Krishna Rajagopal, and Wilke van der Schee. Heavy Ion Collisions: The Big Picture and the Big Questions. *Annual Review of Nuclear and Particle Science*, 68(1):339–376, oct 2018.

[17] S. Voloshin and Y. Zhang. Flow study in relativistic nuclear collisions by Fourier expansion of azimuthal particle distributions. *Zeitschrift für Physik C Particles and Fields*, 70(4):665–671, may 1996.

[18] Yu.L Dokshitzer and D.E Kharzeev. Heavy-quark colorimetry of QCD matter. *Physics Letters B*, 519(3-4):199–206, nov 2001.

[19] ALICE Collaboration. ALICE upgrades during the LHC Long Shutdown 2, 2023.

[20] Francesco Noferini and on behalf of the ALICE Collaboration. ALICE results from Run-1 and Run-2 and perspectives for Run-3 and Run-4. *Journal of Physics: Conference Series*, 1014(1):012010, may 2018.

[21] ALICE Collaboration. The ALICE experiment – A journey through QCD, 2022.

[22] ALICE Collaboration. Production of charged pions, kaons and protons at large transverse momenta in pp and Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76\,TeV$. *Physics Letters B*, 736:196–207, sep 2014.

[23] ALICE Collaboration. K^0_S and Λ production in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76\,TeV$. *Physical Review Letters*, 111(22), nov 2013.

[24] Anton Andronic, Peter Braun-Munzinger, Krzysztof Redlich, and Johanna Stachel. Decoding the phase structure of QCD via particle production at high energy. *Nature*, 561(7723):321–330, sep 2018.

[25] Min He and Ralf Rapp. Charm-baryon production in proton-proton collisions. *Physics Letters B*, 795:117–121, 2019.

[26] Plumari S., Minissale V., and Das S.K. et. al. Charmed hadrons from coalescence plus fragmentation in relativistic nucleus-nucleus collisions at RHIC and LHC. *The European Physical Journal C*, 78(348):1434–6052, 2018.

[27] ALICE Collaboration. Λ_c^+ production in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02\,TeV$. *Physics Letters B*, 793:212–223, 2019.

[28] CMS Collaboration. Search for low mass vector resonances decaying into quark-antiquark pairs in proton-proton collisions at $\sqrt{s} = 13\,TeV$. *Phys. Rev. D*, 100:112007, Dec 2019.

[29] S. K. Radhakrishnan (STAR). Measurements of Open Charm and Bottom Production in Au+Au Collisions at $\sqrt{s_{NN}} = 200\ GeV$ with the STAR Experiment at RHIC . *Proceedings of the 27th International Conference on Ultra-Relativistic Nucleus-Nucleus Collisions (Quark Matter 2015): Lido di Venezia, Italy*, May 14-17 2018.

[30] R. Brun, F. Bruyant, and F. Carminati et. al. GEANT Detector Description and Simulation Tool. *CERN Program Library Long Writeup W5013*, 10 1994.

[31] Suljic Miljenko. *Study of Monolithic Active Pixel Sensors for the Upgrade of the ALICE Inner Tracking System*. PhD thesis, Università degli Studi di Trieste, 2017. Presented 02 Feb 2018.

[32] D. Cohen, E. Koltin, M. I. Ayelet, and A. Shacham. Stitching design rules for forming interconnect layers. Available at `https://patents.google.com/patent/US6225013B1/en` (2001/05/01).

[33] D. Kim, G. Aglieri Rinella, and C. Cavicchioli et. al. Front end optimization for the monolithic active pixel sensor of the ALICE Inner Tracking System upgrade. *Journal of Instrumentation*, 11(02):C02042, feb 2016.

[34] D.A. van den Ende, H.J. van de Wiel, R.H.L. Kusters, A. Sridhar, J.F.M. Schram, M. Cauwe, and J. van den Brand. Mechanical and electrical properties of ultra-thin chips and flexible electronics assemblies during bending. *Microelectronics Reliability*, 54(12):2860–2870, 2014.

[35] M. Buckland. Development of the ITS3: A bent-silicon vertex detector for ALICE in the LHC Run 4. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1039:166875, 2022.

[36] A.G. Straatman, N.C. Gallego, B.E. Thompson, and H. Hangan. Thermal characterization of porous carbon foam—convection in parallel flow. *International Journal of Heat and Mass Transfer*, 49(11):1991–1998, 2006.

[37] Derek J. Quade, M. A. Meador, and E.-S. Shin et. al. The Design, Fabrication, and Testing of Composite Heat Exchange Coupons. *Society for the Advancement of Materials and Process Engineering (SAMPE)*, 17-20 Oct 2011.

[38] S. Ramo. Currents Induced by Electron Motion. *Proceedings of the IRE*, 27(9):584–585, 1939.

[39] L. Rossi, P. Fischer, T. Rohe, and N. Wermes. *Pixel Detectors: From Fundamentals to Applications*. Particle Acceleration and Detection. Springer-Verlag, Berlin, 2006.

[40] R.J. Baker. *Circuit Design, Layout, and Simulation.* IEEE Press Series on Microelectronic Systems. Wiley, 2011.

[41] B. Razavi. *Fundamentals of Microelectronics*. Wiley, 2013.

[42] CMOS: Working, Construction and Applications.

[43] CMOS.

[44] R.C. Jaeger, R. Ramani, J.C. Suhling, and Y. Kang. CMOS stress sensor circuits using piezoresistive field-effect transistors (PIFETs). In *Digest of Technical Papers., Symposium on VLSI Circuits.*, pages 43–44, 1995.

[45] R.C. Jaeger, J.C. Suhling, R. Ramani, A.T. Bradley, and Jianping Xu. CMOS stress sensors on [100] silicon. *IEEE Journal of Solid-State Circuits*, 35(1):85–95, 2000.

[46] A.T. Bradley, R.C. Jaeger, J.C. Suhling, and K.J. O'Connor. Piezoresistive characteristics of short-channel MOSFETs on (100) silicon. *IEEE Transactions on Electron Devices*, 48(9):2009–2015, 2001.

[47] Hiroaki Mikoshiba. Stress-sensitive properties of silicon-gate MOS devices. *Solid-State Electronics*, 24(3):221–232, 1981.

[48] D. A. Bittle, J. C. Suhling, R. E. Beaty, R. C. Jaeger, and R. W. Johnson. Piezoresistive Stress Sensors for Structural Analysis of Electronic Packages. *Journal of Electronic Packaging*, 113(3):203–215, 09 1991.

[49] R.E. Beaty, R.C. Jaeger, J.C. Suhling, R.W. Johnson, and R.D. Butler. Evaluation of piezoresistive coefficient variation in silicon stress sensors using a four-point bending test fixture. *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, 15(5):904–914, 1992.

[50] G. Deptuch, J.D. Berst, G. Claus, C. Colledani, W. Dulinski, Y. Gornushkin, D. Husson, J.L. Riester, and M. Winter. Design and testing of monolithic active pixel sensors for charged particle tracking. *IEEE Transactions on Nuclear Science*, 49:601–610, 2002.

[51] A. Dorokhov, G. Bertolone, J. Baudot, and C. Colledani et. al. High resistivity CMOS pixel sensors and their application to the STAR PXL detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 650(1):174–177, 2011. International Workshop on Semiconductor Pixel Detectors for Particles and Imaging 2010.

[52] S. Mattiazzo, M. Battaglia, and D. Bisello et. al. LePIX: First results from a novel monolithic pixel sensor. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 718:288–291, 2013. Proceedings of the 12th Pisa Meeting on Advanced Detectors.

[53] S. W. Glunz, S. Rein, J. Y. Lee, and W. Warta. Minority carrier lifetime degradation in boron-doped Czochralski silicon. *Journal of Applied Physics*, 90(5):2397–2404, 2001.

[54] A. Rivetti. *CMOS: Front-End Electronics for Radiation Sensors*. CRC Press., 1 edition, 2015.

[55] W. Snoeys. Monolithic pixel detectors for high energy physics. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 731:125–130, 2013. PIXEL 2012.

[56] W. Shockley. Currents to conductors induced by a moving point charge. *J. Appl. Phys.*, 9(10):635–636, 1938.

[57] L. Rossi et al. *Pixel Detectors.* Springer-Verlag, 2006.

[58] Jacobus Willem van Hoorn. *Study and Development of a novel Silicon Pixel Detector for the Upgrade of the ALICE Inner Tracking System*. PhD thesis, TU Vienna, 2015. Presented 24 Nov 2015.

[59] Helmuth Spieler. *Semiconductor Detector Systems*. Oxford University Press, 08 2005.

[60] S. M. Sze and Kwok. K. Ng. *Physics of Semiconductor Devices*. Wiley, 3 edition, 2006.

[61] Grzegorz Deptuch. *New Generation of Monolithic Active Pixel Sensors for Charged Particle Detection*. Theses, Université Louis Pasteur - Strasbourg I, September 2002.

[62] N. S. Saks, M. G. Ancona, and J. A. Modolo. Radiation Effects in MOS Capacitors with Very Thin Oxides at 80°K. *IEEE Transactions on Nuclear Science*, 31(6):1249–1255, 1984.

[63] Nelson S. Saks, Mario G. Ancona, and John A. Modolo. Generation of Interface States by Ionizing Radiation in Very Thin MOS Oxides. *IEEE Transactions on Nuclear Science*, 33(6):1185–1190, 1986.

[64] H. Hillemanns, I. Aimo, C. Bedda, C. Cavicchioli, A. Collu, P. Giubilato, A. Junique, T. Kugathasan, P. La Rocca, M. Mager, C. Marin Tobon, P. Martinengo, S. Mattiazzo, L. Musa, G. Pappalardo, C. Puggioni, F. Reidt, P. Riedler, G. Santagati, S. Siddhanta, W. Snoeys, and J. van Hoorne. Radiation hardness and detector performance of new 180nm

CMOS MAPS prototype test structures developed for the upgrade of the ALICE Inner Tracking System. In *2013 IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC)*, pages 1–5, 2013.

[65] S Bonacini. Design and characterization of an SEU-robust register in 130nm CMOS for application in HEP ASICs. *Journal of Instrumentation*, 5(11):C11019, nov 2010.

[66] M. Lupi. *Design, development, and experimental assessment of a highly-reliable, radiation-tolerant readout system for the upgrade of the ALICE inner tracker*. PhD thesis, Goethe University Frankfurt am Main (DE), 2020.

[67] A. Collu. *Development and characterisation of Monolithic Active Pixel Sensor prototypes for the upgrade of the ALICE Inner Tracking System*. PhD thesis, Università degli Studi di Cagliari, Italy, 2015.

[68] W. Snoeys. CMOS monolithic active pixel sensors for high energy physics. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 765:167–171, 2014. HSTD-9 2013 - Proceedings of the 9th International "Hiroshima" Symposium on Development and Application of Semiconductor Tracking Detectors.

[69] F. Morel, C. Hu-Guo, and G. Bertolone et. al. MISTRAL & ASTRAL: two CMOS Pixel Sensor architectures suited to the Inner Tracking System of the ALICE experiment. *Journal of Instrumentation*, 9(01):C01026, jan 2014.

[70] T. Nooney, A. Bevan, and M. Borri et. al. First Results from Cherwell, a Monolithic Active Pixel Sensor for Particle Physics. *PoS*, EPS-HEP2013:507, 2014.

[71] Felix Reidt. *Studies for the ALICE Inner Tracking System Upgrade*. PhD thesis, Heidelberg U., 2 2016.

[72] ALICE ITS ALPIDE development team. ALPIDE Operations Manual. Version 0.3. Available at `https://usermanual.wiki/Document/ALPIDEoperationsmanualversion03.952786419/html` (July 25, 2016).

[73] P. Yang, G. Aglieri, and C. Cavicchioli et. al. Low-power priority Address-Encoder and Reset-Decoder data-driven readout for Monolithic Active Pixel Sensors for tracker system. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 785:61–69, 2015.

[74] ALICE Collaboration. Inventory of DAQ boards used in ITS3. Available at `https://twiki.cern.ch/twiki/bin/view/ALICE/ITS3WP3DAQboards`.

[75] M. Deveaux, S. Amar-Youcef, A. Budenbender, D. Doering, I. Frohlich, C. Muntz, J. Stroth, and F. M. Wagner. Random Telegraph Signal in Monolithic Active Pixel Sensors. In *2008 IEEE Nuclear Science Symposium Conference Record*, pages 3098–3105, 2008.

[76] K.K. Hung, P.K. Ko, Chenming hu, and Yiu Cheng. Random telegraph noise of deep-submicrometer MOSFETs. *Electron Device Letters, IEEE*, 11:90 – 92, 03 1990.

[77] Philippe Martin-Gonthier and Pierre Magnan. RTS noise impact in CMOS image sensors readout circuit. In *2009 16th IEEE International Conference on Electronics, Circuits and Systems, ICECS 2009*, pages 928 – 931, 01 2010.

[78] P. Tricarico. Cylindrical MAPS-based particle detector prototypes for the Inner Tracking System. WP4 Thinning and Bending tests, ALICE ITS3 upgrade, 2021.

[79] I. Tymchuk. ITS3 SpTAB chipcable assembly: results (part 3) of ALPIDE bending tests. ITS3 WP4, weekly meeting, 2021.

[80] Rune Langøy, Magnus Mager, Anh Tuan T. Nguyen, and Jørgen A. Lien. Thinning and readout during bending of a custom silicon IC. In *2020 IEEE 8th Electronics System-Integration Technology Conference (ESTC)*, pages 1–5, 2020.

[81] Susmita Afroz. Noise and Cluster Size Studies of ALPIDE-CMOS Pixel Sensor for pCT. Master Degree Thesis, June 2018.

[82] A. Nobuhiro. Characterization of pixel sensor prototypes for the Muon Forward Tracker at ALICE. Thesis, February 2017.

[83] CERN. The Super Proton Synchrotron. Available at `https://home.cern/science/accelerators/super-proton-synchrotron` (Online; Last accessed July 6, 2020).

[84] ALICE ITS project. First demonstration of in-beam performance of bent Monolithic Active Pixel Sensors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1028:166280, 2022.

[85] D. Dannheim, K. Dort, L. Huth, D. Hynds, I. Kremastiotis, J. Kröger, M. Munker, F. Pitters, P. Schütze, S. Spannagel, T. Vanat, and M. Williams. Corryvreckan: a modular 4D track reconstruction and analysis software for test beam data. *Journal of Instrumentation*, 16(03):P03008, mar 2021.

[86] P. Ahlburg, S. Arfaoui, and J.-H. Arling et. al. EUDAQ—a data acquisition software framework for common beam telescopes. *Journal of Instrumentation*, 15(01):P01038, jan 2020.

[87] Yi Liu. EUDAQ2 User Manual. Technical report, CERN, Geneva, 2018.

[88] P. Becht. *Performance and characterisation of bent monolithic active pixel sensors (MAPS) for the application in future tracking detectors using the example of the ALPIDE chip*. PhD thesis, Heidelberg University, 2021.

[89] Claus Kleinwort. General broken lines as advanced track fitting method. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 673:107–110, may 2012.

[90] ROOT development team. TEfficiency Class Reference. Official website, ROOT class references, 2020. Available at `https://root.cern.ch/doc/master/classTEfficiency.html`.

[91] C. Patrignani. Review of Particle Physics. *Chinese Physics C*, 40(10):100001, oct 2016.

[92] ALICE ITS project. Tracking performance of bent monolithic active pixel sensors mimicking a truly cylindrical barrel configuration. unpublished, 2023.

[93] T. G. Cornelissen, M. Elsing, I. Gavrilenko, and et. al. The global chi**2 track fitter in ATLAS. *J. Phys. Conf. Ser.*, 119:032013, 2008.

[94] Anton Lechner. Particle interactions with matter. *CERN Yellow Rep. School Proc.*, 5:47, 2018.

[95] C. Bierlich, S. Chakraborty, N. Desai, and et. al. A comprehensive guide to the physics and usage of PYTHIA 8.3. *SciPost Phys. Codebases*, page 8, 2022.

[96] C. Bierlich, G. Gustafson, L. Lönnblad, and et. al. The Angantyr model for Heavy-Ion Collisions in PYTHIA8. *JHEP*, 10:134, 2018.

[97] Gianluca Aglieri Rinella, Anton Andronic, and Matias Antonelli et. al. Digital Pixel Test Structures implemented in a 65 *nm* CMOS process, 2023.

[98] R. P. Fabinski, E. J. Meisenzahl, J. E. Doran, and J. R. Summa. Stitching methods using multiple microlithographic expose tools, patent us8728722b2.

[99] Ozhan Koybasi, Kirk Arndt, Gino Bolla, Daniela Bortoletto, Petra Merkel, and Ian Shipsey. Assembly and qualification procedures of CMS forward pixel detector modules. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 638(1):55–62, 2011.

[100] W. Snoeys, G. Aglieri Rinella, and H. Hillemanns et. al. A process modification for CMOS monolithic active pixel sensors for enhanced depletion, timing performance and radiation tolerance. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 871:90–96, 2017.

[101] E. J. Schioppa et al. Measurement results of the MALTA monolithic pixel detector. *Nucl. Instrum. Meth. A*, 958:162404, 2020.

[102] Gianluca Aglieri Rinella. The ALPIDE pixel sensor chip for the upgrade of the ALICE Inner Tracking System. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 845:583–587, 2017. Proceedings of the Vienna Conference on Instrumentation 2016.

[103] M.Mager (ALICE Collaboration). The LS3 upgrade of the ALICE Inner Tracking System based on ultra-thin, wafer-scale, bent Monolithic Active Pixel Sensors. *15 th Trento Workshop on Advanced Silicon Radiation Detectors. Vienna, Austria.*, 17-19, February 2020.

[104] Nor Zaidi Haron and Said Hamdioui. Why is CMOS scaling coming to an END? In *2008 3rd International Design and Test Workshop*, pages 98–103, 2008.

[105] Siva G. Narendra. Challenges and Design Choices in Nanoscale CMOS. *J. Emerg. Technol. Comput. Syst.*, 1(1):7–49, mar 2005.

[106] G. Gielen, Wim. Dehaene, and P. Christie et. al. Analog and Digital Circuit Design in 65 nm CMOS: End of the Road? In *Proceedings of the Conference on Design, Automation and Test in Europe - Volume 1*, DATE '05, page 36–42, USA, 2005. IEEE Computer Society.

[107] R. Dennard et.al. *IEEE Trans. Electron Devices*. Springer Verlag, 31 edition, 1984.

[108] S. Tyagi. Moore's Law: A CMOS Scaling Perspective. In *2007 14th International Symposium on the Physical and Failure Analysis of Integrated Circuits*, pages 10–15, 2007.

[109] S. Cea and et. al. Iedm technical digest. In *IEEE international electron devices meeting*, page 963–966, 2004.

[110] P. Bai et. al. Iedm technical digest. In *IEEE international electron devices meeting*, pages 657–660, 2004.

[111] K. Mistry and et. al. Vlsi tech symp digest 2002, 2004.

[112] Magnus Mager. A truly cylindrical inner tracker for ALICE. *PoS*, ICHEP2022:318, 11 2022.

[113] Szymon Bugiel, Andrei Dorokhov, and Mauro Aresti et. al. Charge sensing properties of monolithic CMOS pixel sensors fabricated in a 65 nm technology. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1040:167213, 2022.

[114] G. Audi, O. Bersillon, J. Blachot, and A.H. Wapstra. The Nubase evaluation of nuclear and decay properties. *Nuclear Physics A*, 729(1):3–128, 2003. The 2003 NUBASE and Atomic Mass Evaluations.

[115] EM Hussein. *Handbook on Radiation Probing, Gauging, Imaging and Analysis: Volume II Applications and Design*, volume 1. Springer Science & Business Media, 2003.

[116] M.J. Berger, J.S. Coursey, M.A. Zucker, and J. Chang. Stopping-Power and Range Tables for Electrons, Protons, and Helium Ions, 2009.

[117] A.C. Thompson. *X-ray Data Booklet*. Lawrence Berkeley National Laboratory, University of California, 2001.

[118] B.G. Lowe. Measurements of Fano factors in silicon and germanium in the low-energy X-ray region. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 399(2):354–364, 1997.

[119] Miljenko Suljic and ALICE ITS3 group. ITS3 WP3 Approved plots.

[120] K Nakamura and (Particle Data Group). Review of Particle Physics. *Journal of Physics G: Nuclear and Particle Physics*, 37(7A):075021, jul 2010.

[121] H. Bethe and W. Heitler. On the Stopping of Fast Particles and on the Creation of Positive Electrons. *Proceedings of the Royal Society of London Series A*, 146(856):83–112, August 1934.

[122] Gabriele Giacomini. Silicon Detectors for Particle and Nuclear Physics. Available at `https://indico.bnl.gov/event/6284/contributions/29056/` (2019/07/09).

[123] W. R. Leo. *Techniques for Nuclear and Particle Physics Experiments.* Springer, 1994.

[124] Lev Davidovich Landau. On the energy loss of fast particles by ionization. *J. Phys.*, 8(4):201–205, 1944.

[125] K Nakamura and (Particle Data Group). Review of Particle Physics. *Journal of Physics G: Nuclear and Particle Physics*, 37(7A):075021, jul 2010.

[126] Hans Bichsel. Straggling in thin silicon detectors. *Rev. Mod. Phys.*, 60:663–699, Jul 1988.

[127] Francesca Carnesecchi. *Experimental study of the time resolution for particle detectors based on MRPC, SiPM and UFSD technologies*. PhD thesis, alma, Aprile 2018.

[128] S. M. Sze and K. K. Ng. *Physics of Semiconductor Devices*. McGraw-Hill Company, 3 edition, 2002.

[129] A. Doolittle. "chapter 2: Density of states and fermi energy concepts", semiconductor devices. Georgia Institute of Technology., 2011.

[130] Carrier Transport in Semiconductors. Available at `https://ebrary.net/192981/engineering/carrier_transport_semiconductors#aftercont`.

[131] B. Van Zeghbroeck. Principles of Semiconductor Devices.

[132] Debi Prasad Dash. P-N junction.

[133] Simon M. Sze. *Semiconductor Sensors*. John Wiley and Sons, 1994.

[134] W. P. Eaton. *Surface micromachined pressure sensors*. PhD thesis, The University of New Mexico, Albuquerque, New Mexico, 1997.

[135] Gary Johns. Modeling piezoresistivity in silicon and polysilicon. *J. Appl. Eng. Math.*, 2:1–5, 01 2006.

[136] D. W. Burns. *Micromechanical integrated sensors and the planar processed pressure transducer*. PhD thesis, University of Wisconsin, Madison, 1988.

[137] Conyers Herring. Transport properties of a many-valley semiconductor. *The Bell System Technical Journal*, 34(2):237–290, 1955.

[138] Yozo Kanda. Effect of Stress on Germanium and Silicon p-n Junctions. *Japanese Journal of Applied Physics*, 6(4):475, apr 1967.

[139] Yozo Kanda. Piezoresistance effect of silicon. *Sensors and Actuators A: Physical*, 28(2):83–91, 1991.

[140] Conyers Herring and Erich Vogt. Transport and Deformation-Potential Theory for Many-Valley Semiconductors with Anisotropic Scattering. *Phys. Rev.*, 101:944–961, Feb 1956.

[141] Yozo Kanda. A graphical representation of the piezoresistance coefficients in silicon. *IEEE Transactions on Electron Devices*, 29:64–70, 1982.

[142] J. C. Hensel, H. Hasegawa, and M. Nakayama. Cyclotron Resonance in Uniaxially Stressed Silicon. II. Nature of the Covalent Bond. *Phys. Rev.*, 138:A225–A238, Apr 1965.

[143] Yozo Kanda and Katuhisa Suzuki. Origin of the shear piezoresistance coefficient π_{44} of n-type silicon. *Phys. Rev. B*, 43:6754–6756, Mar 1991.

[144] Hiroshi Hasegawa. Theory of Cyclotron Resonance in Strained Silicon Crystals. *Phys. Rev.*, 129:1029–1040, Feb 1963.

[145] J. C. Hensel and G. Feher. Cyclotron Resonance Experiments in Uniaxially Stressed Silicon: Valence Band Inverse Mass Parameters and Deformation Potentials. *Phys. Rev.*, 129:1041–1062, Feb 1963.

[146] Katuhisa Suzuki, Hiroshi Hasegawa, and Yozo Kanda. Origin of the Linear and Nonlinear Piezoresistance Effects in p-Type Silicon. *Japanese Journal of Applied Physics*, 23(11A):L871, nov 1984.