The CMS Pixel Detector Upgrades and Novel Timing-Sensor Technology

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Riccardo Del Burgo

aus

Italien

Promotionskommission

Prof. Dr. Ben Kilminster (Vorsitz und Leitung der Dissertation)

Prof. Florencia Canelli

Prof. Olaf Steinkamp

Prof. Dr. Lea Caminada

Dr. Anna Macchiolo

Zürich, 2021

DEDICA

Contents

Part I

The CMS experiment and silicon sensors

Chapter 1

The Large Hadron Collider and the CMS experiment

1.1 Introduction

This dissertation focuses on the development of silicon detectors for precision tracking of charged particles in the Compact Muon Solenoid (CMS) experiment at the CERN Large Hadron Collider (LHC). Part I of the thesis serves as an introduction to my dissertation. The first chapter introduces the LHC, which provides proton-proton collisions to the CMS experiment. The components of the CMS experiment are then explained, including the detector sub-systems relevant to this thesis, and the methods for reconstructing particles from the detector measurements. Following this, chapter two explains the theoretical and operational principles of silicon sensors, which will provide a basis for the dissertation work in the following chapters. Part II of the thesis comprises the research of my dissertation. Chapter three describes my contributions to the construction of the CMS phase-one pixel detector, which has been in operation since 2017. Chapter four explains a simulation I developed to model the propagation of electric charge for a future silicon sensor that is to be used in the CMS phase-two pixel detector. Chapter five covers my research into silicon pixel sensors with precise timing resolution, which are considered as an option for pixel detectors in a stage after the phase-two pixel upgrade. Chapter six reiterates the conclusions of the previous three chapters, and puts them into context within the current and future particle physics landscape.

1.2 The Large Hadron Collider

The Large Hadron Collider (LHC) is the world's largest and most powerful particle accelarator [1]. It first started up on 10 September 2008, and is the latest addition to the CERN accelerator complex. The LHC consists of a 27-kilometre ring of superconducting magnets with two accelerating stages per beam[2] to boost the energy of the particles traveling along the way.

The LHC is the last accelerator in the accelerator complex. The accelerator complex is shown in Figure 1.1. The protons are supplied by a hydrogen bottle shown in Figure 1.2.

The hydrogen atoms are stripped of the electrons and accelerated by the LINAC2 up to 50 MeV. The beam is then injected into the Proton Synchrotron Booster (PSB), which accelerates the protons to 1.4 GeV, followed by the Proton Synchrotron (PS), which pushes the beam to 25 GeV. Protons are then steered via magnetic field to the Super Proton Synchrotron (SPS) where they are accelerated to 450 GeV.
The protons are finally transferred to the two beam pipes of the LHC. The beam in one pipe circulates clockwise while the beam in the other pipe circulates anticlockwise, and both travel at 0.999999991 times the speed of light before they collide. The beam pipes are kept at pressures between 10^{-7} and 10^{-9} Pa (atmospheric pressure is 101325 Pa)[3].

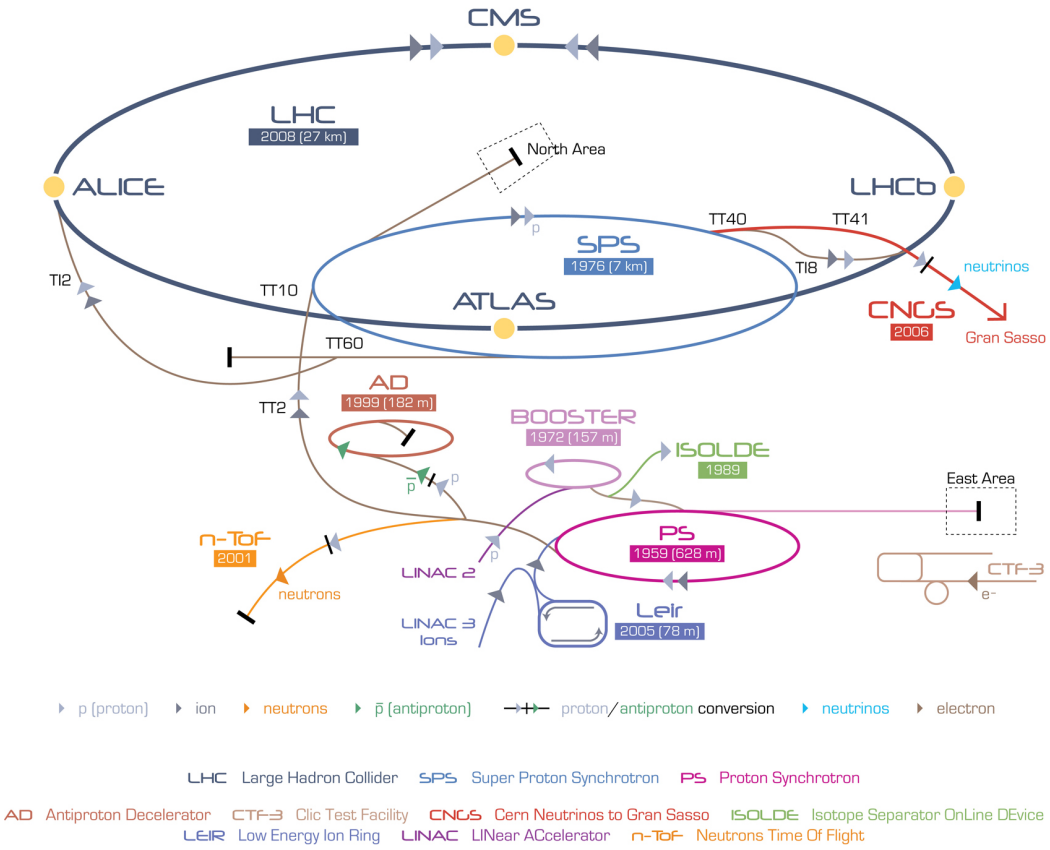


Figure 1.1: The CERN accelerator complex.

The LHC structure is shown in Figure 1.3. Each of the octants hosts a straight section used for an experimental area or beam manipulation, like beam injection, beam dump or beam cleaning. The two beams collide at fixed interaction points. The protons are bent by 1232 superconducting dipole magnets, 15 meters in length, operating at a temperature of 1.9 K, and producing a magnetic field of 8.3 T. The electromagnets are built from coils of filaments made of Niobium-titanium (NbTi) that operates in a superconducting state, efficiently conducting electricity without resistance or loss of energy. Radio-frequency (RF) cavities, whose oscillation frequency is tuned to 400 MHz, are used to accelerate the protons. The energy gain per revolution is about 485 keV, in spite of the 7 keV loss per turn due to synchrotron radiation. The particle beam focus is kept by 392 quadrupole and sextupole magnets which act as electromagnetic lenses. It takes 4 minutes and 20 seconds to fill each LHC ring, and 20 minutes for the protons to reach their maximum energy of 6.5 TeV. At

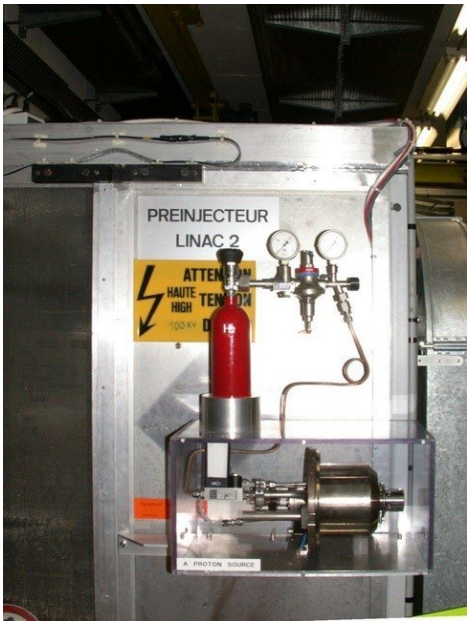


Figure 1.2: The hydrogen bottle that supplies protons to the LHC accelerator complex. [4]

the end of the acceleration chain, protons arrive at the IPs in 2808 bunches, each containing 10^{11} protons. Just prior to collision, special high gradient quadrupoles magnets are used to "squeeze" the particles closer together to increase the chances of collisions. Proton beams with a transverse beam size of 16 microns are steered in opposite directions over the 27 km circumference and collided together within a space of 7 cm.

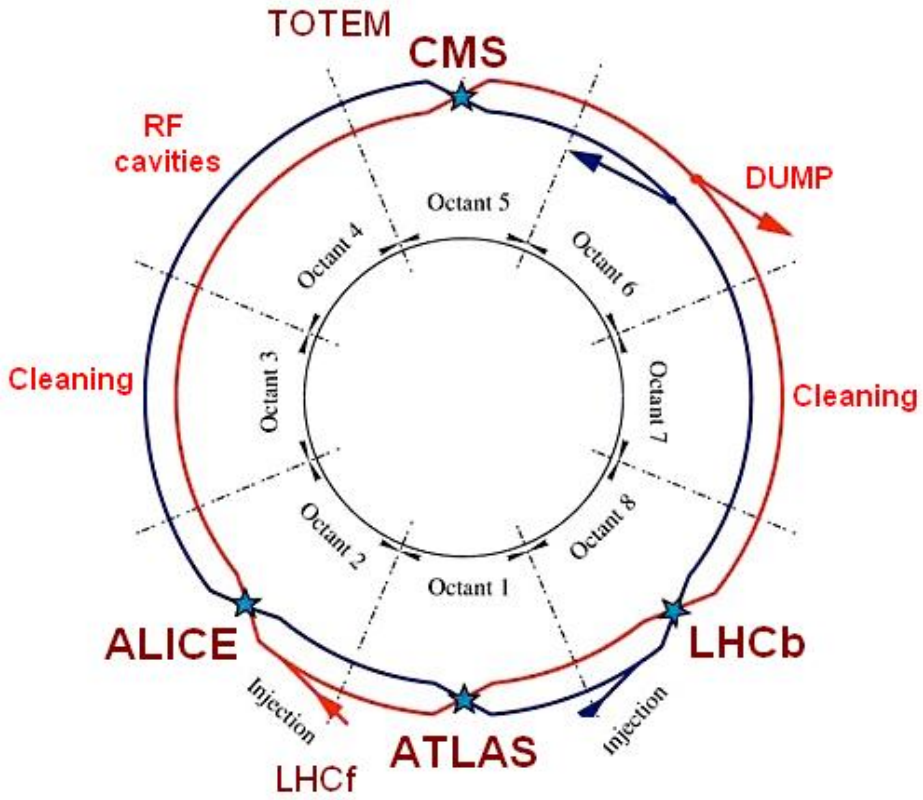


Figure 1.3: Schematic picture of the LHC, layout not in scale.

1.2.1 Proton-proton collisions

The number of events generated per second [5] in the LHC collisions is given by:

$$\frac{dN_{\text{events}}}{dt} = \mathcal{L}\sigma_{\text{events}}, \tag{1.1}$$

where σ_{events} is the cross-section for the event under study and $\mathcal{L}$ is the machine's instantaneous luminosity, measured in $\text{cm}^{-2}\text{s}^{-1}$. The machine luminosity depends on the beam parameters. For the LHC, the beams have a gaussian distribution in the normal plane, so the instantaneous luminosity can be expressed as:

$$\mathcal{L} = \frac{N_1 N_2 f N_{\text{b}}}{4\pi\sigma_{\text{x}}\sigma_{\text{y}}}, \tag{1.2}$$

where N_1 and N_2 are the number of particles per bunch, f is the revolution frequency, and N_{b} is the number of colliding bunches. The beam spread along the x and y directions are introduced with the σ_{x} and the σ_{y} parameters. An important quantity is the integrated luminosity [5], defined as the integral over time:

$$\mathcal{L}_{\text{int}} = \int_0^T \mathcal{L}(t)dt. \tag{1.3}$$

This quantity directly relates to the number of observed events N_{events}. Combining Equation 1.1 and Equation 1.3 one finds that the integrated luminosity can be computed as:

$$\mathcal{L}_{\text{int}} = \frac{N_{\text{events}}}{\sigma_{\text{events}}}. \tag{1.4}$$

The LHC was designed to collide protons at a centre-of-mass energy of 14 TeV and with a peak instantaneous luminosity of 10^{34} $\text{cm}^{-2}\text{s}^{-1}$.
During the period from 2010 to 2011 the centre-of-mass energy was kept at 7 TeV, in order to avoid limitations in the superconducting magnets. Afterwards, it was increased to 8 TeV for the 2012 run, and after a long shutdown to repair the interconnects between the superconducting dipoles, was increased to 13 TeV for the runs between 2015 and 2018. Training of the superconducting magnets to higher fields may make possible higher centre of mass energies in subsequent runs.
The LHC also collides heavy ions at an energy of 2.8 TeV and a peak luminosity of 10^{27} $\text{cm}^{-2}\text{s}^{-1}$.
The collisions created by the LHC are recorded by the experiments and used for the purpose of standard-model and beyond-standard-model physics analyses. Figure 1.4 shows the integrated luminosity delivered by the LHC to the CMS experiment per year, and Figure 1.5 shows the cumulative luminosity delivered to and recorded by the CMS experiment. The recorded data is less than the delivered data due to the requirement that all sub-detectors and data acquisition systems are operational.

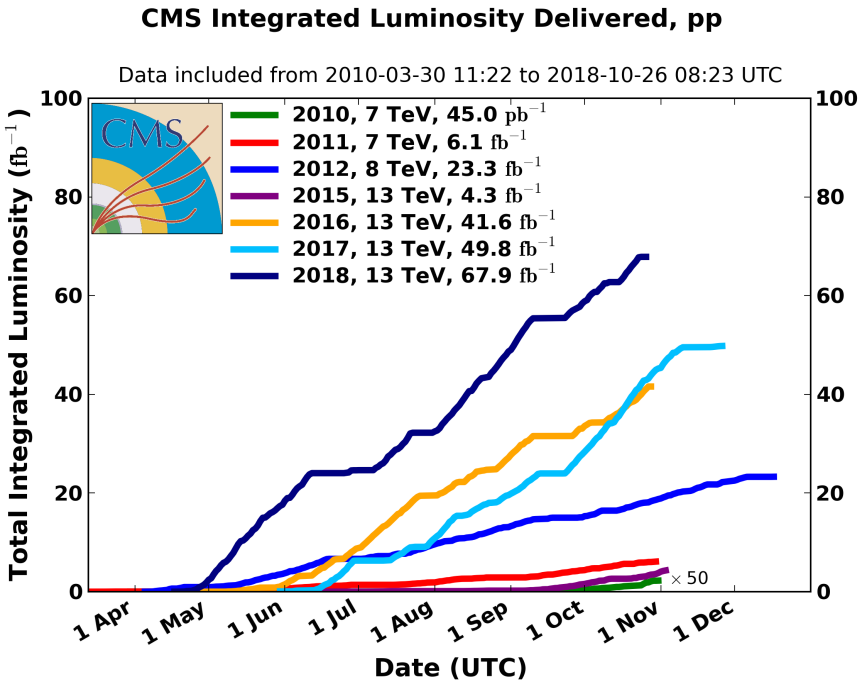


Figure 1.4: CMS integrated luminosity delivered during the different years.

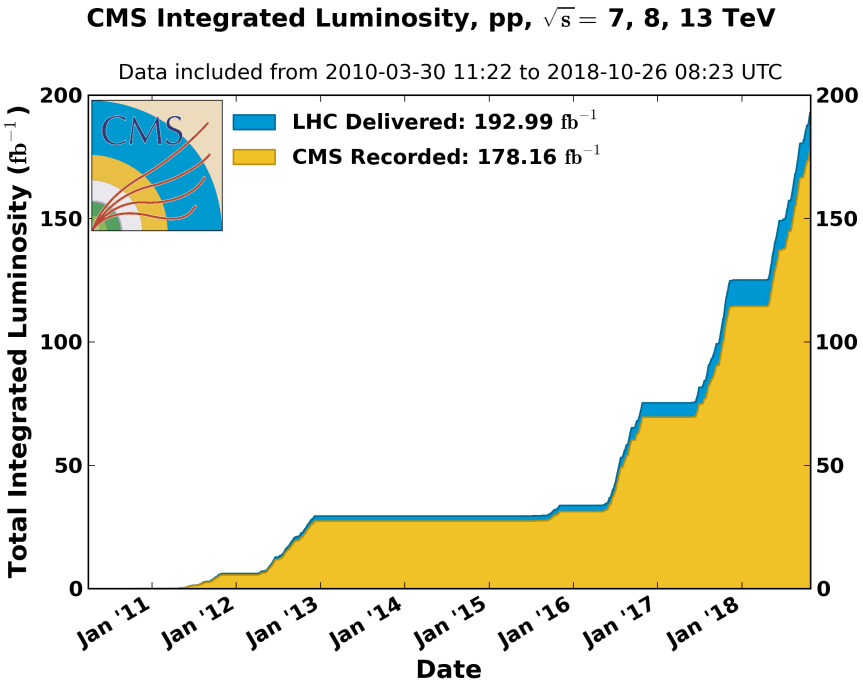


Figure 1.5: CMS integrated luminosity delivered vs recorded for the whole data-taking period.

1.3 The CMS experiment

The Compact Muon Solenoid (CMS) experiment is one of two large general-purpose particle physics detectors built within the Large Hadron Collider (LHC) at CERN. The main goals of the CMS experiment are:

- to explore physics at the TeV scale;
- to further study the properties of the Higgs boson, already discovered by CMS and ATLAS;
- to look for evidence of physics beyond the standard model, such as supersymmetry, extra dimensions, or dark matter;
- to study aspects of heavy ion collisions.

CMS is 21.6 metres long, 14.6 m in diameter, and weighs about 14,000 tonnes. CMS has a symmetric cylindrical shape and its central feature is a superconducting solenoid generating a magnetic field of 3.8 T. Inside the solenoid several sub-detectors are installed: a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL).
Each sub-detector is composed of a barrel section and two end sections to ensure hermetic coverage. Forward sampling calorimeters extend the pseudorapidity coverage to high values ($|\eta| < 5$). Muons are detected in gas-ionization chambers, which are instrumented with the steel flux-return yoke outside the solenoid. Figure 1.6 is a schematic view of the detector. CMS uses a global coordinate system to describe the direction of particles that are identified

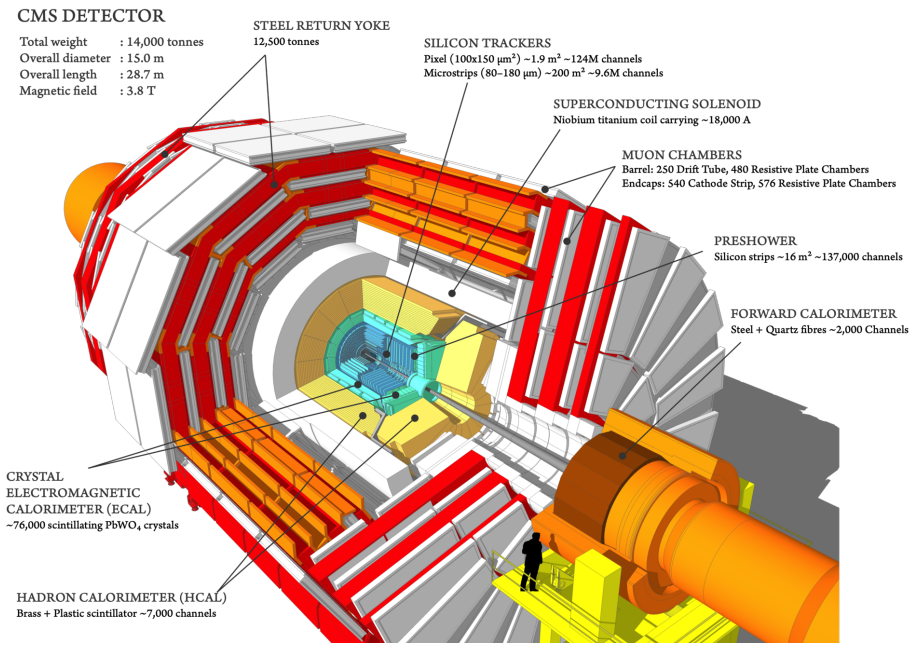


Figure 1.6: Cutaway diagram of the CMS detector after the Phase 1 Pixel upgrade. [6]

in it. The global coordinates with respect to the LHC complex are shown in Figure 1.7. The origin of the coordinates is the geometrical centre of the CMS detector, which roughly

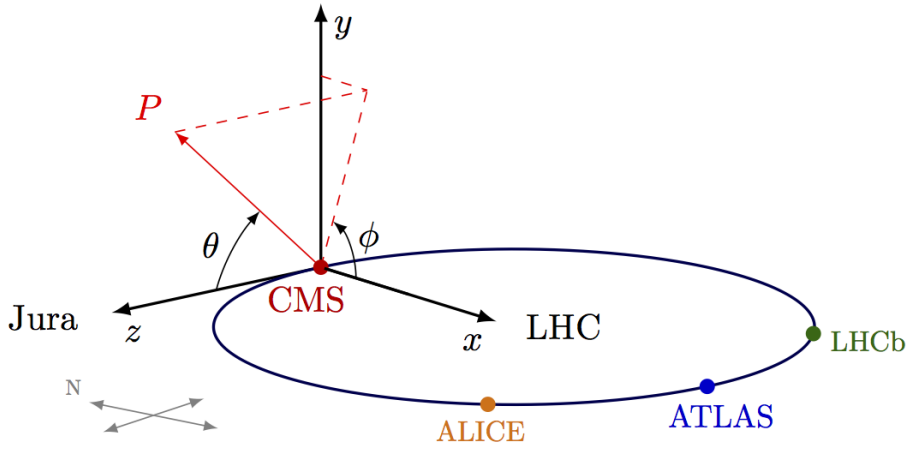


Figure 1.7: CMS global coordinate system [7]

corresponds to the beam interaction region. The z-axis is identified as the direction of the anti-clockwise beam, which is pointing towards the Jura mountains. The x-axis points to the centre of the LHC ring and the y-axis is perpendicular to the ring plane. The azimuthal angle, ϕ, is measured from the x-axis in the x-y plane and its value ranges in the interval $[-\pi, +\pi]$. The polar angle θ is measured from the z-axis in the z-y plane, and it ranges from $[0, \pi]$. The pseudorapidity η is defined as:

$$\eta = -\log(\frac{\tan\theta}{2}). \tag{1.5}$$

1.3.1 Magnet

The CMS magnet is a solenoid with a diameter of 6.3 meters and a length of 12 meters. The magnet generates a 3.8 Tesla magnetic field, which is aligned with the z-axis. A charged particle with charge q traveling in a magnetic field B with a momentum p_T, which is perpendicular to the B field, will bend with a radius r according to:

$$\frac{1}{r[m]} = \frac{0.3B[T]}{p_T[GeV]} \tag{1.6}$$

Since charged particles also have a momentum p_z, parallel to the magnetic field and therefore not affected by it, they will follow helical trajectories inside the CMS solenoid. The measurement of the bending radius is used to extract the transverse momentum of the particles, so it is of paramount importance. Assuming a very large bending radius compared to the magnet length L and considering the finite track resolution of the detector, the momentum resolution $\sigma(p_{\mathrm{T}})$ is found to be proportional to p_{T}^2. It can be approximated as :

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = \frac{p_{\mathrm{T}}[\mathrm{GeV}]}{0.3L[\mathrm{m}]} B[\mathrm{T}] \frac{2\sigma(x)}{d}, \tag{1.7}$$

where $\sigma(x)$ is the error on the track measurements, and d is the distance between the first and last sensors hit by the particle. Typically, for charged particles with p_T >100GeV the transverse momentum resolution is about 1–2% in the central pseudorapidity region [8], but degrades at high pseudorapidities as the radial distance of the last position measurement

decreases such that the lever arm decreases. At lower momenta, the resolution is dominated by multiple scattering and its distribution reflects the amount of material traversed by the track.

1.3.2 Inner tracking system

The inner tracking system of CMS is composed of two sub-detectors, the pixel detector and the strip detector. These two system measure with precision the trajectory of the particle and the interaction point position. The sensor used in both detectors is silicon. The tracking detectors are held within a cylindrical volume 5.8 m long with a diameter of 2.4 m, covering a pseudorapidity region up to $|\eta| < 2.5$. The different sub-detectors that compose a quarter of the CMS tracker are show in Figure 1.8. This pixel detector shown is the phase-one pixel detector installed in 2017, and referred to in detail in chapter 3.

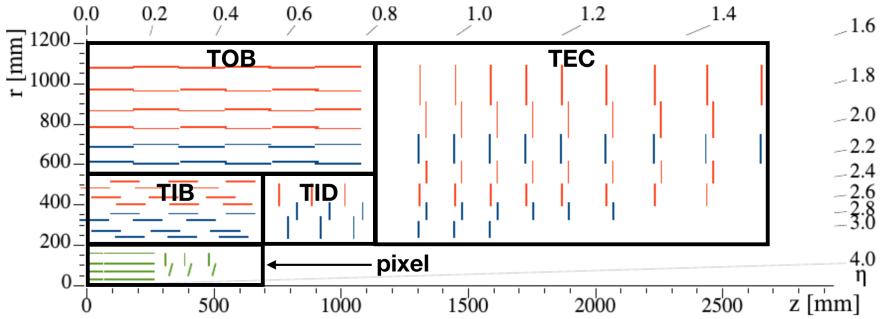


Figure 1.8: CMS tracker. The innermost detector is the pixel detector. The TIB, TID, TOB, and TEC make up the strip detector.

Pixel detector, and its upgrade

The phase-0 pixel detector that operated from 2010-2016 was replaced in 2017 with the phase-1 pixel detector. Among the improvements were an increase in cylindrical layers from three to four in the barrel region, and two to three disks in each forward region as shown in Figure 1.8. The disks are split in inner and outer ring structures. The purpose of the pixel detector is to reconstruct the vertex position of the interaction, as well as any discernible secondary vertices, with high spatial precision [9]. It provides also the starting point for the track reconstruction, providing 4 inner position measurements with higher precision than the strip detector. The diameters of the pixel layers range from 29 mm to 160 mm, and the pixel cell size is 100×150 μm^2.
The pixel detector resolution is of the order of 10 μm and will be discussed in depth in chapter 3.

Strips detector

The strips detector encapsulates the pixel detector and is composed of three main sub-systems [10]:

- TIB/TID, tracker inner barrel and disks;
- TOB, tracker outer barrel, which surrounds the TIB/TID;
- TEC+/TEC-, tracker endcaps.

The innermost strip sub-detectors, TIB/TID, extend in radius up to 55 cm and are grouped in 4 barrel layers and 3 disks at each end. The silicon sensors in this region have a thickeness of about 320 μm and are parallel to the beam axis in the barrel and perpendicular to it in the endcaps. The resolution in $r\phi$ is $\approx$ 13 - 38 μm. The TOB surrounds the TIB and TID. It extends in radius up to 116 cm and is divided into 6 barrel layers of 500 μm thick micro-strips. It has a length along the z-axis of 118 cm. It provides position measurements with a resolution of about $\approx$ 18 - 47 μm. Finally, the TEC+/TEC- cover the region ranging between 124 cm and 282 cm in z and 22.5 cm and 113.5 cm in r. Both TEC sub-detectors are composed of 9 disks and the sensors located here have a a thickness of 320 - 500 μm.

1.3.3 Calorimeter

The CMS calorimetry system measures the energy deposited by particles that interact via the electromagnetic or strong interactions. It is composed of two sub-detectors, the electromagnetic calorimeter and the hadron calorimeter.

Electromagnetic calorimeter

The CMS electromagnetic calorimeter (ECAL) is a hermetic, homogeneous, fine-grained lead tungstate ($PbWO_4$) crystal calorimeter [11]. The detector is shown in Figure 1.9.

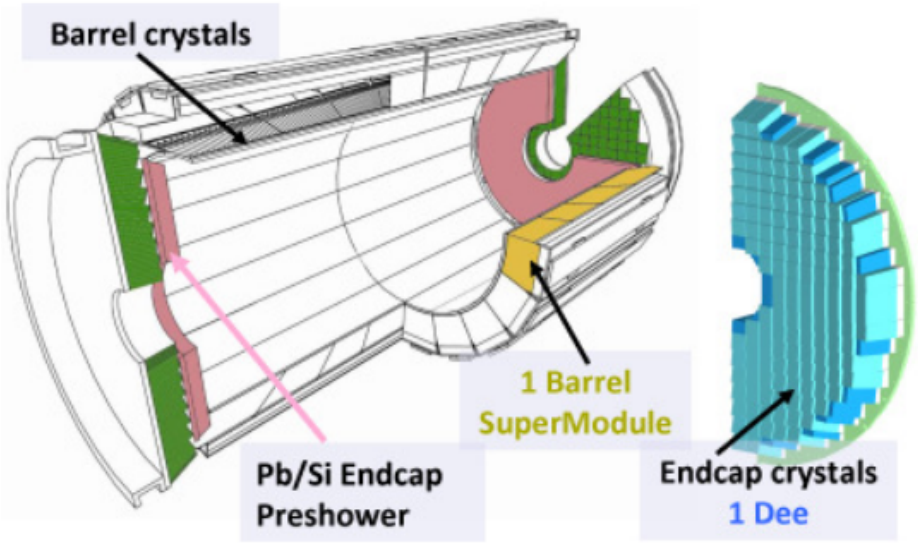


Figure 1.9: Schematic view of the CMS electromagnetic calorimeter.

The choice of a homogeneous medium was made to obtain a better energy resolution by minimising sampling fluctuations. Very dense crystals offer the potential to achieve the required excellent performance and compactness, which was necessary to fit within

the volume of the CMS superconducting solenoidal magnet. The CMS design enabled the electromagnetic calorimeter to fit within the volume of the CMS superconducting solenoid magnet. The 75,848 crystals are arranged in a central barrel section (EB), with pseudorapidity coverage up to $|\eta|= 1.48$, and closed by two endcaps (EE), extending coverage up to $|\eta| =3.0$. A preshower detector (ES), based on lead absorber and silicon strips sensors, placed in front of the endcaps at $1.65<|\eta|<2.6$, improves the photon-pion separation. The total thickness of the ES is $\approx$3 radiation lengths.
The energy resolution of the ECAL detector is factorised in three main contributions: a stochastic term, a noise term, and a constant term [12]. It is expressed as:

$$\frac{\sigma_E}{E} = \frac{2.8\%}{\sqrt{E}} \oplus \frac{12\%}{E} \oplus 0.3\%, \tag{1.8}$$

where the parameter values were estimated using an electron test beam. The stochastic term takes into account the statistical fluctuations in the light collection efficiency and quantum efficiency, the noise term describes the electronic noise, and finally, the constant parameter accounts for the effect of the operational conditions. This estimate corresponds to an energy resolution of 1% for photons and electrons with energy larger than 10 GeV. The energy resolution is illustrated in Figure 1.10 as a function of electron energy.

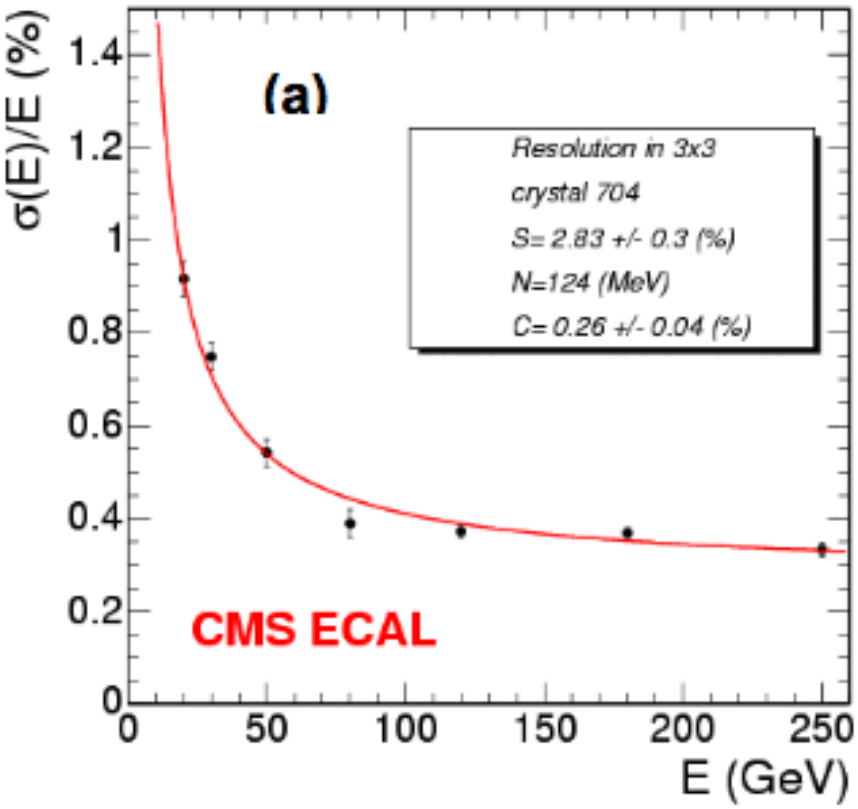
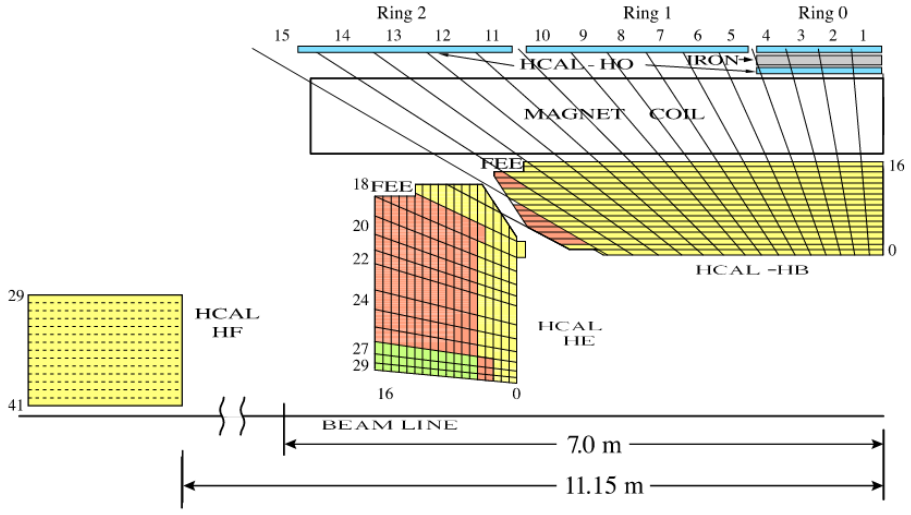


Figure 1.10: The ECAL energy resolution as a function of electron energy from test beam performance [12].

Hadron calorimeter

The Hadron Calorimeter (HCAL) includes four distinct subsystems [13]: the barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters. The HB and HE, which are inside the cryostat of the CMS superconducting solenoid that provides a 3.8 T magnetic field, are sampling calorimeters where the absorber material is brass and the active material is scintillator. The sampling fraction is about 7% . The HB and HE are separated by a gap which is located approximately at a polar angle of 57 degrees but is not projective to the center of CMS, in order to minimize the effect of the un-instrumented gap. HB covers the range in η from zero to approximately 1.4. The $|\eta|$ range between 1.3 and 1.4 is shared by the HB and HE, and HE covers also the $|\eta|$ region between 1.3 and 3.0. The HB is composed of 18 wedges, each of which covers 20 degrees in ϕ, and is divided in 5 degrees sectors. HE is made of brass disks, interleaved with scintillator wedges which cover 20 degrees in ϕ, which are in turn divided in four 5-degrees sectors from the HB. Because of the space constraint within the magnet cryostat, the HB thickness is limited to 5.8 hadronic interaction lengths at η=0 and increases to 10 interaction lengths at $|\eta|$=1.2. To catch the energy leakage from the HB, layers of scintillators are placed outside the solenoid cryostat, which constitute what is called the HO. About 5% of all hadrons above 100 GeV deposit some energy in HO. The



as

Figure 1.11: The CMS HCAL detector (quarter slice). FEE indicates the locations of the Front End Electronics for HB and HE. The signals of the tower segments with the same color are added optically, to provide the HCAL longitudinal segmentation. HB, HE and HF are built out of 36 identical wedges ($\Delta\ \phi = 20$ degrees).

parametrization of the HCAL energy resolution is similar to that of ECAL and is described with stochastic and noise terms [14]. It can be expressed as

$$\frac{\sigma_E}{E} = \frac{115\%}{\sqrt{E}} \oplus \frac{5.5\%}{E}, \tag{1.9}$$

where the parameters have been estimated from test beam measurements. The energy resolution of the CMS ECAL and HCAL detectors for pions, resulting from test beam

measurements, as a function of energy, is parametrised as [14]:

$$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus \frac{5\%}{E}. \tag{1.10}$$

1.3.4 Muon system

Muon reconstruction in CMS is performed combining the information from the silicon strip tracker and pixel detector with the four stations of gas-ionisation muon detectors installed outside the solenoid and sandwiched between steel layers serving both as hadron absorbers and as a return yoke for the magnetic field. The different muons sub-detectors are shown in Figure 1.3.4 [15].

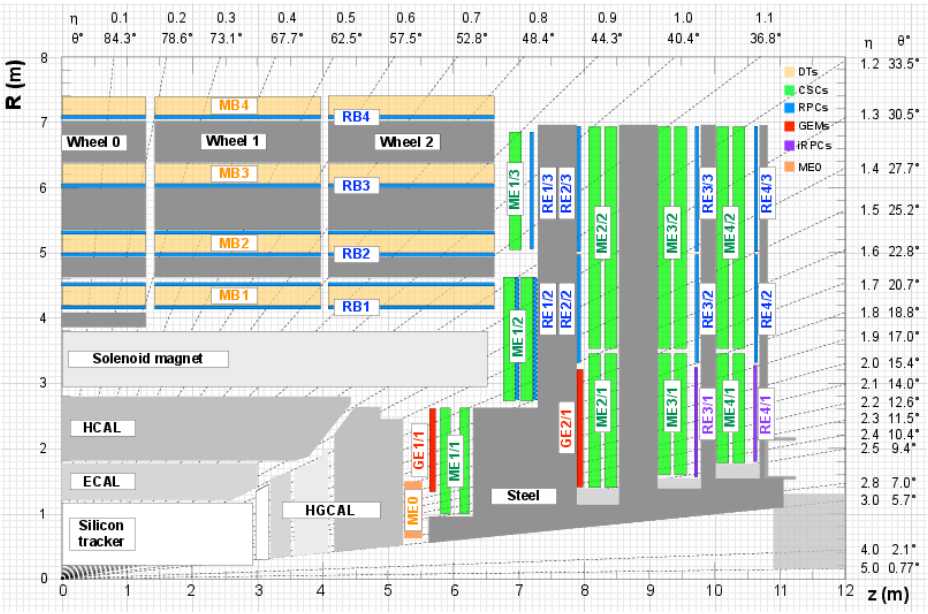


Figure 1.12: Cross section of a quadrant of the CMS detector, including the Phase-2 upgrades (RE3/1, RE4/1, GE1/1, GE2/1, ME0). The acronym iRPCs in the legend refers to the new improved RPC chambers RE3/1 and RE4/1. The interaction point is at the lower left corner. The locations of the various muon stations are shown in colour (MB are Drift Tubes (DTs), ME are Cathode Strip Chambers (CSCs), RB and RE are Resistive Plate Chambers (RPCs), GE and ME0 are Gas Electron Multipliers (GEMs)). M denotes Muon, B stands for Barrel and E for Endcap. The magnet yoke is represented by the dark grey areas.

The muon system covers the pseudo-rapidity region $|\eta|$ <2.4 and performs three main tasks: triggering on muons, identifying muons, and assisting the CMS tracker in measuring the momentum and charge of high p_T muons.
The system uses three different gaseous detector technologies: Drift Tube (DT), Cathode Strip Chambers (CSC), and Resistive Plate Chambers (RPC). DT chambers and CSC are used within $|\eta|$ <1.2 and 0.9 < $|\eta|$ <2.4 respectively, complemented by a system of RPC covering the range of $|\eta|$ <1.6. The basic element of the DT system is the cylindrical drift cell of 42×13 mm^2. Each cell contains a stretched wire within a gas mixture. The gas

mixture (85%/15% of Ar/CO_2) provides good quenching properties and a saturated drift velocity of about 55 μm/ns. The maximum drift time is almost 400 ns.
A CSC consists of 6 layers, and operates as a standard multi-wire proportional counter (MWPC). The cathode strips run radially outward and provide a precise measurement in the r-ϕ bending plane. The wires provide a coarse measurement in the radial direction. All chambers use a gas mixture of 50% CO_2, 40% Ar, and 10% CF_4. The RPC are double-gap chambers, operated in avalanche mode to ensure reliable operation at high rates, that can provide fast and independent trigger signals. The spatial resolution per chamber is between 80 and 120 μm in the DTs, and between 40 and 150 μm in the CSCs, and smaller than 1.2 cm in the RPCs. The time resolution achievable was 3 ns or better per chamber for all 3 systems. Hit and segment reconstruction efficiencies have been shown to be high for all muon detectors, and hits are reconstructed with good resolution. Excellent timing capabilities of the CMS muon system allow for proper bunch crossing identification at the trigger level. The performance of muon identification and isolation algorithms has been studied using the full 2016 data sample [16]. In general efficiencies are high and good agreement is observed between data and simulation.

1.3.5 Trigger and data acquisition systems

Inside CMS, proton proton collisions happen 40 million times every second. Due to data bandwidth limitations, not all events can be recorded, and only a selection of the most interesting ones are recorded. To choose which event to record, CMS relies on the triggering system. The CMS triggering system is composed of two levels, the Level-1 (L1) trigger and the high level (HL) trigger. The L1 trigger uses custom-designed electronics while the HL trigger relies on commercially available PCs.
The CMS L1 trigger has to cope with a very high bunch crossing rate, 40 MHz, and has to reduce this high rate to about 100 kHz [17]. Its architecture is shown in Figure 1.13. The L1 trigger combines information coming from the muon detector and calorimeter to quickly choose if the event is interesting. If the event passes the L1 trigger, the data recorded from all the other sub-detector systems are read out and sent to the HL trigger.
The HL trigger reduces the rate from 100 kHz to about 1 kHz, and has about 10 μs to make a decision on the event. The HL trigger exploits more complex algorithms to reconstruct the events and aims to be as close as possible to the performance of the offline reconstruction. Because the HL trigger uses more specialised algorithms, different trigger paths are implemented in parallel and deployed in what is called the trigger menu. Some software modules of the HL trigger provide input to multiple trigger paths. For instance, the module that identifies muons is implemented once, but the output list of identified muons may be used in various trigger paths that search for B hadron decays, top quark decays, or potential supersymmetric processes. In general the bandwidth available for each path is pre-determined, but as the instantaneous luminosity inside the LHC decreases over the beam lifetime, some trigger path efficiencies can be reduced or enhanced to maximize the use of bandwidth – in jargon this is called "pre-scaling".
Also due to this need to dynamically adapt to LHC conditions, trigger menus are constantly

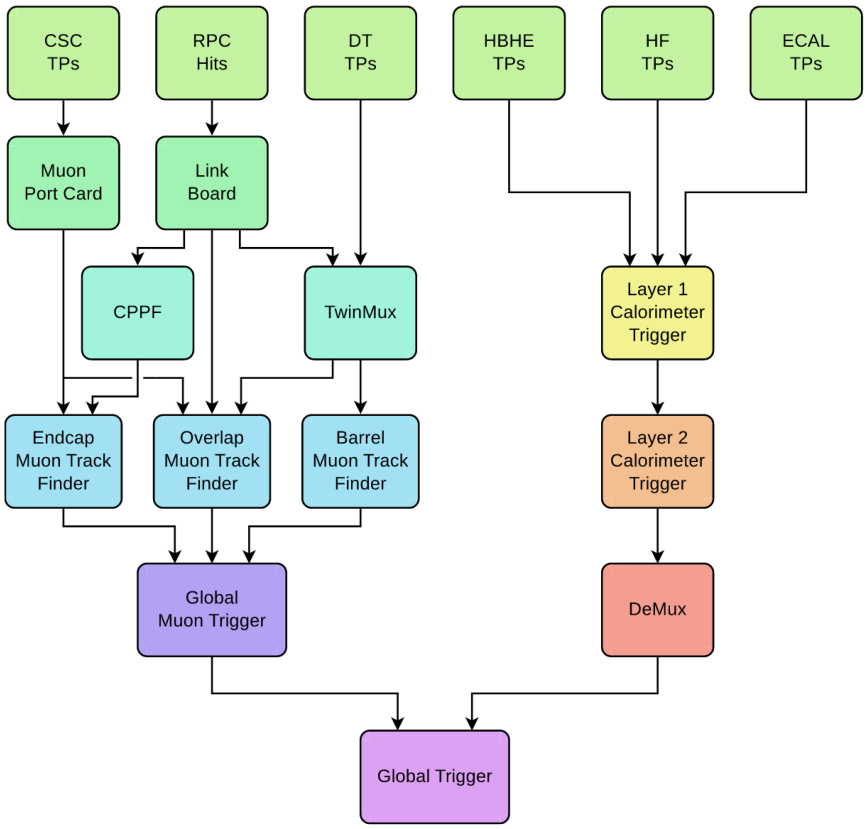


Figure 1.13: Diagram of the upgraded CMS Level-1 trigger system during Run 2. Labels in the diagram correspond to trigger primitives (TPs), cathode strip chambers (CSC), drift tubes(DT), resistive plate chambers (RPC), concentration preprocessing and fan-out (CPPF), hadron calorimeter barrel (HB) and endcap (HE), hadron calorimeter forward (HF), electromagnetic calorimeter (ECAL), and de-multiplexing card (DeMux). Trigger Primitives (TPs) refer to coarse-level determinations of measured objects such as muon tracks or calorimeter deposits.

updated throughout the run period, and new versions often feature versions of previously un-prescaled triggers, in order to satisfy the bandwidth requirements by tightening the restrictions for the events to pass the trigger. For example, an increase in the P_T requirement, or a restriction of the pseudorapidity range, would decrease the rate of triggered muons.

1.3.6 Computing

The LHC produces annually roughly 15 PB of data, which will be accessed and analysed by the scientists within the collaborations. The data are collected, analysed, and stored in the Worldwide LHC Computing Grid (WLCG), which is composed of four different levels or "Tiers", referred to as 0, 1, 2 and 3. Each Tier with the exception of Tier 0 is made up of several computer centres and provides a specific set of services. Between them, the tiers process, store and analyse all the data from the Large Hadron Collider (LHC). Tier 0 is the CERN Data Centre. All of the data from the LHC passes through this central hub, but it provides less than 20% of the Grid's total computing capacity. Approximately ten Tier-1 centres distributed internationally are responsible for storing the huge amount of data in a safe environment. Tier-2 and Tier-3 centres play an important role for analysing the data and running simulations. Tier-2 centres are typically at universities and other scientific institutes that can store sufficient data and provide adequate computing power for specific analysis tasks. Individual scientists can access the Grid through local (or Tier 3) computing resources, interfacing via local clusters in a university department or from an individual PC.

1.4 Object and event reconstruction

As described in the previous sections, CMS relies on independent sub-detector systems to measure the particles in order to reconstruct the events produced inside the detector. All the information coming from different sub-systems have to be combined and used in an optimal way to reconstruct the event [18].

Physics objects are mainly categorised as follows:

- Jets consist of hadrons and photons, the energy of which can be inclusively measured by the calorimeters without any attempt to separate individual particles within the jet. Jet reconstruction can therefore be performed without any contribution from the tracker and the muon detectors.

- The reconstruction of the energy and direction of isolated photons and electrons is determined using the ECAL.

- The tagging of jets originating from hadronic τ decays and from b quark hadronization is based on the properties of the charged particle tracks, and thus mostly involves the measurements from the tracker.

- The identification of muons is principally based on the information from the muon detectors.

An improved event description can be achieved by combining the basic information from all detector layers (tracks and clusters) to identify final-state particles, and by combining the corresponding measurements to reconstruct the particle properties on the basis of this identification. This approach is called particle-flow (PF) reconstruction. In general, the performance of the PF reconstruction will depends on the nature of the particle, which sub-detectors information is available and the quality of those, and finally the performance of the reconstruction algorithm itself.

1.4.1 Charged-particle tracks and vertices

The aim of charged-particle track reconstruction is to measure the momentum of charged particle tracks, which are mainly pions, as well as the charged tracks associated with muons, τ leptons, and b-quark decay. A combinatorial Kalman filter (KF) track finder is used to reconstruct those tracks. The KF uses a three-staged approach:

- initial seed generation starting from hits inside the pixel detector,
- trajectory building: the KF propagates the seed and includes all the compatible hits in the whole tracker volume,
- final fitting.

To be kept for further analysis, the tracks have to be seeded with at least three hits in consecutive layers in the pixel detector, and are required to be reconstructed with at least nine hits in total.
Increasing the track reconstruction efficiency while keeping the mis-reconstructed rate unchanged is critical for PF event reconstruction, and is accomplished by applying the combinatorial track finder in several successive iterations, each with moderate efficiency, but with as high a purity as possible. At each step, the reduction of the mis-reconstruction rate is accomplished with quality criteria on the track seeds, on the track-fit [2], and on the track compatibility with originating from one of the reconstructed primary vertices, as a function of the track p_T, η, and number of hits n_{hits}. The number of iterations is ten, and each one is tuned differently to optimise the reconstruction efficiency while keeping the fake rate as low as possible. In the following table the different target and seeding technique for each iterations are reported:

Iteration	Name	Seeding	Targeted tracks
1	InitialStep	pixel quadruplets	prompt, high p_T
2	LowPtQuad	pixel quadruplets	prompt, low p_T
3	HighPtTriplet	pixel triplets	prompt, high p_T recovery
4	LowPtTriplet	pixel triplets	prompt, high p_T recovery
5	DetachedQuad	pixel quadruplets	displaced
6	DetachedTriplet	pixel triplets	displaced, recovery
7	MixedTriplet	pixel + strip triplets	displaced
8	PixelLess	inner strips triplet	very displaced
9	TobTec	outer strips triplets	very very displaced
10	JetCore	pixel pairs in jets	high- $p_T$jets
11	MuonInsideOut	muon tagged tracks	muon
12	MuonOutsideIn	isolated muons	muon

The first three steps reconstruct triplet and quadruplets seeds used in different PF blocks. To reconstruct these seeds, since 2017 the seeding algorithm is based on the cellular automaton technology, which reconstructs hit pairs and checks the compatibility of the pairs with respect to the interaction point. These first three iterations are the only ones performed by the HLT to reduce the computing time.

1.4.2 Particle identification and reconstruction

A given particle is, in general, expected to give rise to several PF elements in the various CMS sub-detector. The reconstruction of a particle starts with a link algorithm that connects the PF elements from different sub-detectors. A typical event display is shown in Figure 1.14. The probability for the algorithm to link elements from one particle is only limited by the granularity of the various sub-detectors and by the number of particles to resolve per unit of solid angle. The probability to link all elements of a given particle is mostly limited by the amount of material encountered upstream of the calorimeters and the muon detector, which may lead to trajectory kinks and to the creation of secondary particles. The link algorithm can test any pair of elements in the event. To prevent the computing time from exploding, the link algorithm considers only the closest pairs in (η, ϕ).
The specific conditions required to link two elements depend on their nature, and are listed in the next paragraphs. If two elements are found to be linked, the algorithm defines a distance between these two elements, aimed at quantifying the quality of the link. The link algorithm then produces PF blocks of elements associated either by a direct link or by an indirect link through common elements. PF blocks are created when the algorithm links electron bremsstrahlung photons, or for example when it connects energy deposits in the ECAL and HCAL.
In each PF block, the reconstruction proceeds as follow:

- Muon candidates are identified and reconstructed and the corresponding PF elements (tracks and clusters) are removed from the PF block.

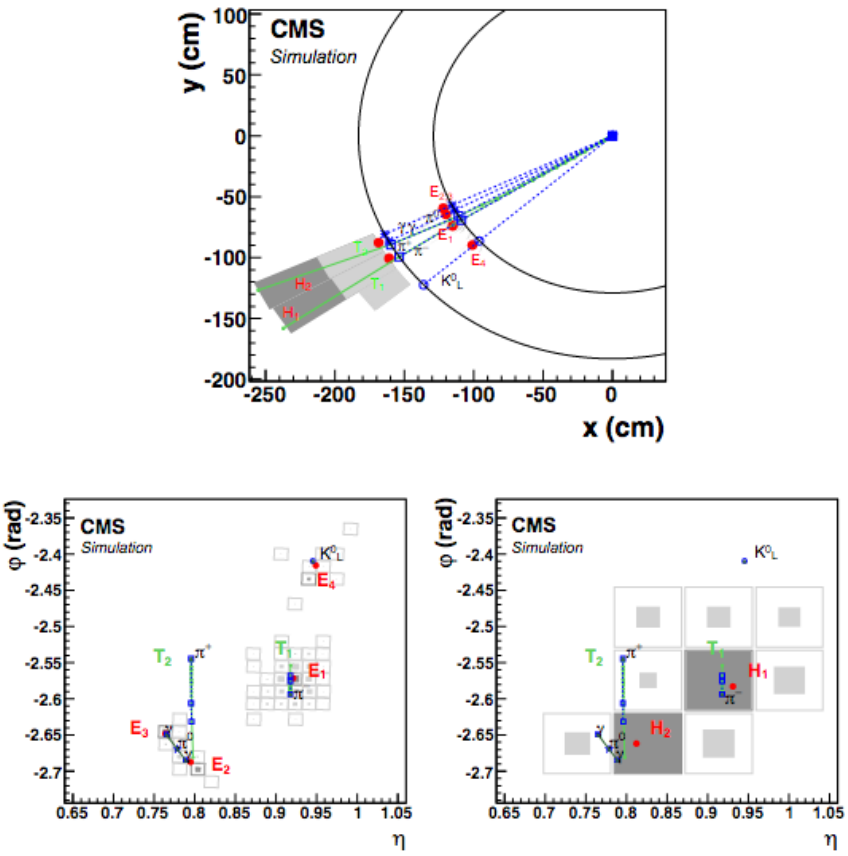


Figure 1.14: Event display of an illustrative jet made of five particles only in the (x,y)view (upper panel), and in the (η, ϕ) view on the ECAL surface (lower left) and the HCAL surface (lower right). In the top view, these two surfaces are represented as circles centred around the interaction point. The K_L^0, the π^-, and the two photons from the decay are detected as four well-separated ECAL clusters denoted $E_{1,2,3,4}$. The π^+ does not create a cluster in the ECAL. The two charged pions are reconstructed as charged-particle tracks $T_{1,2}$, appearing as vertical solid lines in the (η, ϕ) views and circular arcs in the (x,y) view. These tracks point towards two HCAL clusters $H_{1,2}$. In the bottom views, the ECAL and HCAL cells are represented as squares, with an inner area proportional to the logarithm of the cell energy. Cells with an energy larger than those of the neighbouring cells are shown in dark grey. In all three views, the cluster positions are represented by dots, the simulated particles by dashed lines, and the positions of their impacts on the calorimeter surfaces by various open markers [18].

- Electrons are identified and reconstructed with the aim of collecting the energy of all bremsstrahlung photons. Energetic and isolated photons are identified in the same step. The corresponding tracks and ECAL or preshower clusters are excluded from further consideration.
- At this level, tracks with a p_T uncertainty in excess of the calorimetric energy resolution expected for charged hadrons are adequately reduced.
- The remaining elements in the block are then subject to a cross-identification of charged hadrons, neutral hadrons, and photons, arising from parton fragmentation, hadronization, and decays in jets.
- Finally, when the global event description becomes available, all blocks have been processed and all particles have been identified.

Chapter 2

Silicon sensors

2.1 Introduction

Charged particles and radiation are detected by their interactions with matter. At the present day, gaseous, liquid, and solid-state materials are used in a variety of applications for detecting particles in large volumes. The most recent and fundamental detector development began in the 1970s through the work of Josef Kemmer, who was the first to implement silicon solid-state detector technology for the purpose of investigating the rare "charmed" particles that were then discovered in the year 1976. This newly introduced detector is what we know of today as a silicon strip detector, with a resolution of the order of tens of micrometers and the ability to work in a high rate environment. The starting technology was supplied by the growing microelectronics field that was heavily investing in developing the silicon technology. The relatively long lifetime of these particles (10^{-12}–10^{-13} seconds), known today as D mesons, in conjunction with the commensurate position resolution of a silicon detector, made possible the observation of such short-lived particles. Since these first pioneering days we have come a long way, and since their first iteration, the four major experiments at CERN, namely ATLAS, CMS, LHCb, and ALICE all host silicon detectors as their innermost sub-detector.
In this chapter the basic concepts and working principles of silicon sensors will be introduced.

2.2 Intrinsic and extrinsic semiconductors

The most common semiconductor technology uses a monocrystalline silicon lattice. The diamond cubic lattice structure of the silicon is shown in figure 2.1 [19]. The lattice is composed of two interpenetrating face-centred cubic sub-lattices, displaced by a quarter of the lattice constant along the diagonal of the cube. Figure 2.2 shows on the left side the two dimensional bond representation of the lattice structure, while the right side shows the energy band structure of a generic crystal as a function of the lattice pitch. The electrical properties of the material strongly depend on the lattice pitch. If the valence and conduction band are overlapping, the material is a conductor, if they are not connected the material is an insulator. The semiconductor family of materials sits in between the two categories, such

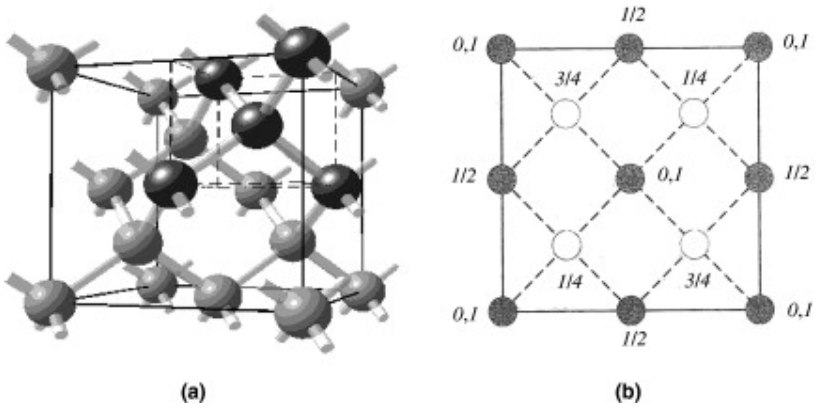


Figure 2.1: The diamond crystal lattice of silicon. (a) Spatial illustration with covalent bonding and (b) projection view[19].

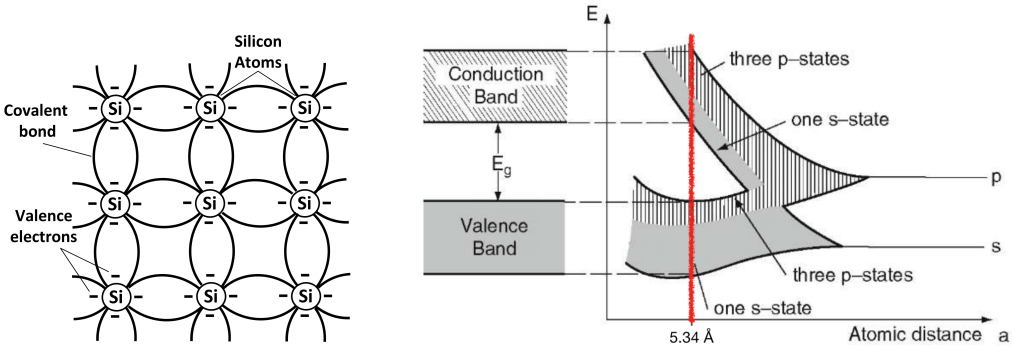


Figure 2.2: Left: Schematic bond representation of a single silicon crystal. Right: Energy bands in a lattice as a function of the lattice pitch. The red line marks the lattice pitch of silicon[19].

that the valence and conduction bands are separated by a small energy gap. The valence band is full. The energy gap between valence and conduction bands in silicon is of the order of eV, and the thermal energy at ambient temperature is sufficient to move electrons from the valence band to the conduction band. As the electrons in the conduction band are free to move along the crystal lattice, also the holes left in the valence band behave as free positive charges. In pure silicon, the number of holes and conduction electrons is the same. Pure silicon is called an intrinsic semiconductor because no impurities are added to the crystal. The silicon crystal can present itself as a single crystal (mono-crystalline) or in a poly-crystalline structure. The electrical characteristics of poly-crystalline silicon depend on the crystal size. In the case of single crystal silicon the electrical characteristics can be controlled by adding specific impurities in the crystal.

An intrinsic semiconductor is rarely used, instead it is common to introduce impurities in the semiconductor and use it as an extrinsic semiconductor. The procedure of adding impurities is called doping. Depending on the type of impurities added, it is possible to obtain two type of silicon: n-type silicon with an excess of electrons in the conduction band and

p-type silicon with an excess of holes in the valence band. The silicon atom itself has four valence electrons. If one atom in the lattice is replaced with an atom of 5 valence electrons (pentavalent) like Arsenic, one electron will be free to move, creating an extrinsic n-type silicon semiconductor. On the other hand, if we replace a silicon atom in the lattice with a trivalent atom like Boron or Gallium, the missing electron in the bond will behave like a hole, with a positive charge, and we will have created a p-type extrinsic semiconductor. The doping locally alters the energy levels of the silicon crystal. In silicon, the use of Arsenic or Phosphorus as an n-dopant creates a new local energy level very close to the conduction band, similarly using Boron will create a new energy level close to the valence band. As a consequence, the additional electron (hole) will be free and the state completely ionised at room temperature.

2.3 Charge carriers drift and diffusion in semiconductors

The charge carriers in a semiconductor, in absence of an external magnetic or electric field, are essentially free particles with a kinetic energy of:

$$E = \frac{3}{2}kT. \tag{2.1}$$

The charge carriers movement is due to thermal energy, and the flow is governed by the diffusion equation:

$$\mathbf{F_{n(p)}} = -D_{n(p)}\nabla n(p), \tag{2.2}$$

where $\mathbf{F}$ is the flux of the carriers, D is their coefficient of diffusion, and $n(p)$ is their concentration. The net average charge transport at equilibrium is zero.
When an electric field is applied, the charge carriers are accelerated in between collisions with an average speed described by the following equations:

$$v_n = -\frac{q \cdot \tau_c}{m_n}\varepsilon = -\mu_n\varepsilon, \tag{2.3}$$

$$v_p = -\frac{q \cdot \tau_c}{m_p}\varepsilon = -\mu_p\varepsilon, \tag{2.4}$$

in which the drift speed v is proportional to the electric field ε, the carrier charge q, the mean free time τ_c between collisions, and inversely proportional to the carrier mass m. The subscripts n and p refer to negative and positive charge carriers. These equations define mobility μ as the proportionality constant between the electric field and the carrier's velocity v.
These equations hold true if the electric field is small enough that the energy acquired by the particles in between collisions is of the order of the thermal velocity. For higher electric fields the simple linear relation between electric field and carriers velocity is not valid anymore. For a high enough electric field (10^5V cm^{-1}) the drift velocity is independent of the field

intensity and saturates [20]. A useful piece of information contained in these equation is that, for a given electric field, the velocity of the electrons and the holes is related by:

$$\frac{v_n}{v_p} = \frac{m_p}{m_n}. \tag{2.5}$$

From experimental data and theoretical calculations, electrons drift at a speed 3-4 times faster than for holes [21].
The drift velocity depends on the local electric field, temperature, and the concentration of dopants in the silicon.
Combining the drift and diffusion equations, the current density for charge carriers can be expressed as:

$$\mathbf{J_{n(p)}} = +q(D_{n(p)}\nabla n(p) + \mu_{n(p)} \cdot n(p) \cdot \varepsilon), \tag{2.6}$$

which describes the charged carriers motion under the effect of an electric field and thermal energy.

2.4 Magnetic field effect on the charge carrier's drift

The effect of the magnetic and electric fields on the charged carriers is expressed by the Lorentz force equation:

$$\mathbf{F} = +q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \tag{2.7}$$

Using equation 2.3 the angle between the velocity v and the electric field vector can be expressed as:

$$\tan\theta_{n(p)} = \mu^H_{n(p)} B, \tag{2.8}$$

where $\mu^H_{n(p)}$ is the Hall mobility. The Hall mobility differs from the mobility, because the presence of a magnetic field removes the randomness of the drift in between collisions. The effect of the magnetic field and its practical consequences will be discussed in section 4.7.

2.5 Carrier generation and recombination in semiconductor detectors

The charge carriers are generated inside a semiconductor detector by three main mechanisms:

- thermal generation,
- absorption of electromagnetic radiation, and
- energy loss by a charged particle.

Thermal generation is a detrimental effect that can be mitigated by operating the semiconductors at lower temperature. Silicon is an indirect semiconductor: the maximum of the valence band and the minimum of the conduction band are located at different momenta. The energy necessary to move one electron from the valence band to the conduction band is

larger than the energy gap between the two bands. The presence of dopants and impurities inside the silicon creates intermediate energy levels which facilitate the jump of the electrons between the valence to the conduction bands.
The photon's absorption effect depends on the energy of the photons. If the photon energy is larger than the band-gap energy, after the absorption of the photon, an electron will be excited to the conduction state leaving a free hole behind. The excited electron will then decay to ground level by phonon or photon emission.
A charged particle traveling through a medium loses part of its energy through inelastic collisions with the electrons of the medium. The energy loss rate is described by the Bethe-Bloch formula [22]:

$$\frac{dE}{dx} = 2\pi N_0 r_e^2 m_e c^2 \rho \frac{Z}{A} \frac{z^2}{\beta^2} \left[\ln \left(\frac{2m_e \gamma^2 v^2 W_{max}}{I^2} \right) - 2\beta^2 - \delta - 2\frac{C}{Z} \right], \tag{2.9}$$

where:

- N_0 is the Avogadro number,
- r_e is the classical electron radius,
- m_e is the classical electron mass,
- $2\pi N_0 r_e^2 m_e c^2 = 0.1535$ MeVc^2/g,
- x is the path length in g/cm^2,
- I is the effective ionisation potential averaged over all electrons,
- Z is the atomic number of the medium,
- A is the atomic weight of the medium,
- ρ is the density of the medium,
- z is the charge of the traversing particle,
- $\beta = v/c$ is the velocity v of the traversing particle normalised to the speed of light c,
- $\gamma = \frac{1}{\sqrt{1-\beta^2}}$,
- δ is the dimensionless density correction,
- C is the dimensionless shell correction which is predominant at low energy,
- W_{max} is the maximum energy transfer in a single collision.

The energy-loss rate of a muon in copper as a function of the muon momentum is shown in figure 2.3 [22].

For low velocity/ momentum, less than 1 MeV, the energy-loss rate is proportional to $1/\beta^2$. Around a momentum of 1 GeV the muon energy-loss rate is at a minimum, and the particles in this momentum range are therefore called Minimum Ionising Particles (MIPs).

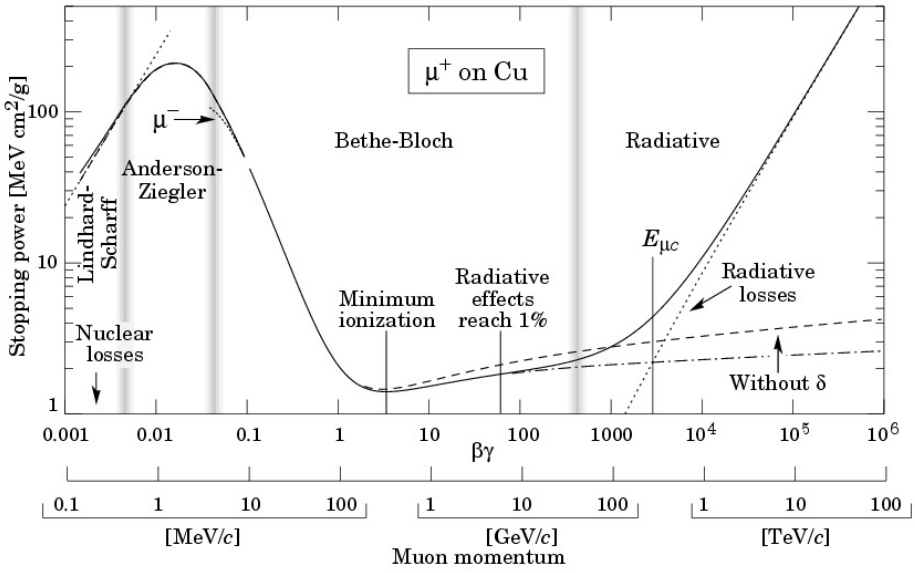


Figure 2.3: Rate of energy loss of muon in copper as a function of the muon momentum [22].

As the momentum of the particle increases, radiative losses begin to become dominant while the rate of energy loss by ionisation remains almost constant.
The effective average energy deposited by a particle traversing a finite material can be obtained by integrating the Bethe-Bloch formula over the particle's momentum range.
Experimentally, it is found that the energy deposit is distributed as a Landau distribution, with the average energy deposited predicted by the Bethe-Bloch formula. Figure 2.4 [23] shows the energy-loss distribution for 12 GeV protons traversing different silicon thicknesses. The continuous lines are the Landau predictions, while the dots are experimental data points. The peak value of each distribution is correctly predicted by the Bethe-Bloch formula.

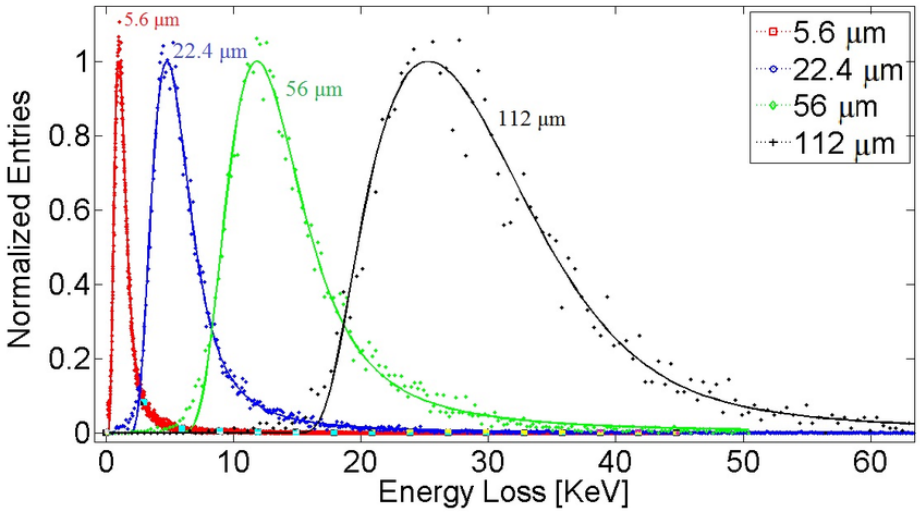


Figure 2.4: Energy loss distributions with fits for 12 GeV protons traversing different silicon thicknesses [23].

2.6 Charge cloud shape

The shape of the charge cloud generated inside a semiconductor depends on the energy and type of particles that generate it. The photon interaction is point-like, and the number of electron-hole pairs is given by its energy divided by the mean ionisation energy of the silicon, 3.6 eV. For visible light (wavelength $\approx \mu$m, energy $\approx$ eV) each photon interacts close to the surface of the silicon, and generates typically one electron-hole pair. An X-ray photon (energy $\approx$ 10 keV) will generate hundreds of e-h pairs in a very small point-like region.
The energy loss rate of α particles, because of their charge (+2) and their mass, is described by the first peak of the Bethe-Block graph in figure 2.3. The energy is deposited along a short path, and the penetration depth is of the order of micrometers. The energy deposit as a function of depth is known as the Bragg curve, and features a prominent peak in energy deposition at the end of the path.
Electrons (β particles), because of their smaller electric charge and mass compared to those of α particles, penetrate the semiconductor, and typically pass through silicon sensors, as they remain relativistic. The energy deposit is spread along its path and is uniform. A relativistic singly charged particle such as the β is often referred as a minimum ionising particle or MIP.
A high-energy charged particle in the relativistic regime behaves as a MIP, in the same way as the previously described β radiation. MIPs deposit a small amount of energy in the silicon and travel through it unimpeded. This behaviour is exploited in HEP experiments, in which large surface, small-mass, high-granularity silicon detectors are used to reconstruct particle tracks without altering the particle's energy in a significant way.
Non-relativistic charged particles behave in a similar way to α particles, depositing all their energy along a short path. The amount of energy deposited is proportional to the inverse of the particle energy and its path length. For this reason, semiconductor detectors are often used as calorimeters and for particle identification of heavy, slow particles such as nuclei.

2.7 Charge carrier lifetime and dark current

In the solid-state physics of semiconductors, carrier generation and carrier recombination are processes by which mobile charge carriers (electrons and holes) are created and eliminated. Carrier generation and recombination processes are fundamental to the operation of semiconductor devices. The electron-hole pair is the fundamental unit of generation and recombination in semiconductors. Carrier generation describes processes by which electrons gain energy and move from the valence band to the conduction band, producing two mobile carriers; while recombination describes processes by which a conduction band electron loses energy and re-occupies the energy state of an electron hole in the valence band. Two important parameters for a solid state detector are the recombination and generation lifetimes.
Silicon is an indirect semiconductor. The recombination and generation lifetime calculation is very complex for indirect semiconductor, so the calculation for the direct semiconductor will be discussed first. For a direct p-type semiconductor, the highest energy state of the valence band and the lowest energy level of the conduction band are at the same momenta.

In this case the passage of electrons from the valence band to the conduction band, as well as holes in the other direction, happen directly, without intermediate jumps.
In the case of a direct p-type semiconductor in thermal equilibrium, although the average charge does not change, thermal electron-hole pairs are continuously generated and recombined. The recombination R rate is proportional to the dopant concentration:

$$R = \beta np \tag{2.10}$$

where β is a proportionality constant, and n and p are the dopant concentrations. At thermal equilibrium, the generation G and recombination R rates are the same:

$$R_{th} = G_{th} = \beta n_0 p_0 = \beta n_i^2, \tag{2.11}$$

where β is the proportionality constant. It is possible to define the recombination rate of the excess as:

$$R - R_{th} = \beta(np - n_i^2) = \beta\left((n + \Delta n)(p + \Delta n) - n_i^2\right), \tag{2.12}$$

and for a p-doped silicon :

$$p_0 >> n_0 \quad , \quad p_0 >> \Delta n, \tag{2.13}$$

$$G_L = R - R_{th} \approx \beta(\Delta n p_0) = \frac{\Delta n}{\tau}, \tag{2.14}$$

where τ is the recombination time constant.

$$\tau = \frac{1}{\beta p_0}, \tag{2.15}$$

considering a depleted p-type semiconductor. The thermal generation rate is:

$$R_{th} = G_{th} = \beta n_0 p_0 = \beta n_i^2, \tag{2.16}$$

and the time constant will be:

$$\tau_g = n_i / G_{th} = \frac{1}{\beta n_i}. \tag{2.17}$$

Inverting the previous formula :

$$G_{th} = n_i / \tau_g, \tag{2.18}$$

and what we obtain is the free charge carrier generation rate for a unit of volume inside a structure.
The free charge carriers generated by this mechanism are responsible for the so-called dark current. Dark current is one of the main sources for noise in silicon sensors. The treatment for an indirect semiconductor is a little bit more complicated. In this case the movement from the valence to the conduction band is mediated by jumps to intermediate energy levels offered by a crystal defect or by the dopant itself. These intermediate levels are called traps, and as the name suggests, they traps charge carriers and re-emit them after a certain time. The traps are of two different kinds, donors and acceptors. Acceptor traps are positive and

become neutral by trapping an electron, while donor traps are negative and become neutral by trapping a hole.
At equilibrium the process of charge carrier trapping and emission have to compensate each other. Starting from this assumption it is possible to obtain the following equation for the generation lifetime:

$$G_{th} = \frac{N_t \nu_{th,n} \sigma_n \nu_{th,p} \sigma_p n_i}{\nu_{th,n} \sigma_n e^{\frac{E_t - E_i}{kT}} + \nu_{th,p} \sigma_p e^{f - \frac{E_t - E_i}{kT}}}, \tag{2.19}$$

where N_t is the defect concentration, E_t the defect energy level, ν_{th} is the thermal velocity, σ the capture cross section, and the n and p subscripts denote negative and positive charge carriers.

2.8 The *p-n* diode junction

The most fundamental and important semiconductor structure is the p-n diode. It can be created by connecting two semiconductors with opposite doping. When a p-type semiconductor and an n-type semiconductor are in contact, holes and conduction electrons are free to move, and the excess electrons from the n side will move to the p side and vice-versa for the holes. This will create a surplus of positive charge in the n side of the junction and of negative charge in the p side of the junction that will oppose the migration of majority carriers until equilibrium is reached. The situation inside the diode is shown in figure 2.5 using the abrupt approximation, where the doping concentration changes abruptly at the junction.
The shift in energy level can be obtained by imposing that the Fermi energy levels should be the same along the junction and in all the structure. In figure 2.6 the current along the junction is split in its different components, namely the thermal generation and the recombination of charge carriers [24].

It is useful to derive the voltage drop along the junction and the width of the natural space charge region in a p-n junction without any voltage applied to understand what happens when an external voltage is applied. The built-in voltage can be derived by the difference in the intrinsic levels in the n and p regions:

$$V_{bi} = \frac{1}{q}(E_i^p - E_i^n) = \frac{kT}{q} \ln \frac{N_A N_D}{n_i^2}, \tag{2.20}$$

where E_i^p and E_i^n are respectively the energy level in the p-doped and n-n doped regions, and N_A and N_D the concentrations of acceptors and donors, andn_i is the product of N_A and N_D. Using this equation, and requiring that the total charge in the free-charge region is zero, the width d of the junction region can be expressed as:

$$d = \sqrt{\frac{2\varepsilon\varepsilon_0}{q} \frac{(N_A + N_D)}{N_A N_D} V_{bi}}. \tag{2.21}$$

The charge-free region is also known as the depletion region. If the diode is placed near an electromagnetic radiation source with photon energies greater than the minimum band-gap of 1.12 eV, the photons will create new electron-hole pairs that drift and generate a current. This is the working principle behind a photo-voltaic panel.

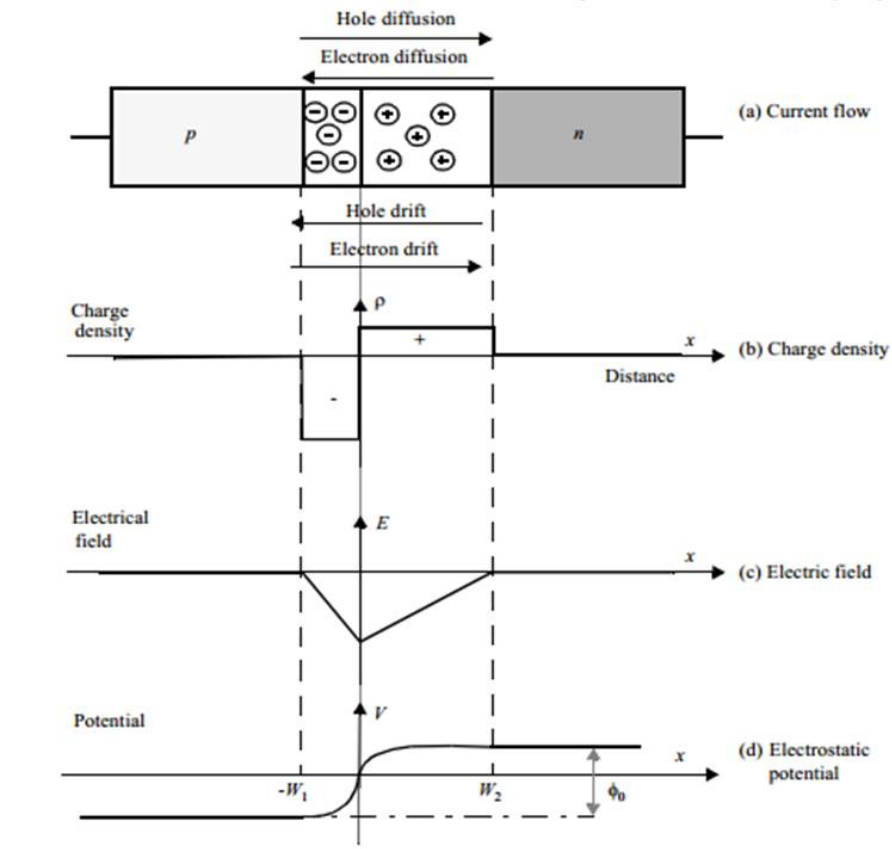


Figure 2.5: Current flow, charge density, electric field and electrostatic potential in a p-n junction under the abrupt approximation [24].

This kind of structure is great for generating energy but does not work well as a detector. Using the typical dopant concentration for a diode and equation 2.21, it is possible to calculate that the width of the sensitive region around the junction is around 25μm thick. Applying a reverse bias V_B to the sensor using a voltage source, it is possible to increase the depletion region width d:

$$d = \sqrt{\frac{2\varepsilon\varepsilon_0}{q}\frac{(N_A + N_D)}{N_A N_D}(V_{bi} + V_B)}. \tag{2.22}$$

The bias voltage necessary to totally deplete the sensor is called the depletion voltage. The depletion voltage scales with the square of the thickness of the sensor, until full depletion. Inverting the last equation we obtain:

$$V_{dep} = \frac{qN_D d^2}{2\varepsilon\varepsilon_0}, \tag{2.23}$$

which is a good approximation not only for diodes but for junction structures in general. Figure 2.7 shows the current-voltage characteristics, also called the IV curves, for a sensor before and after irradiation for two different fluences, and before and after different annealing times. At full depletion the IV curves display a plateau region in which the leakage current

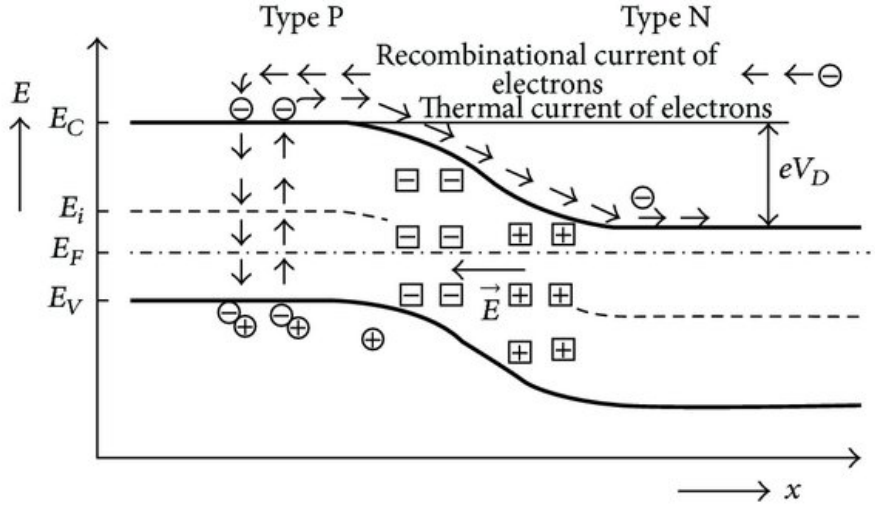


Figure 2.6: Band diagram of energy levels of the p-n junction [24]. The energy level of the valence band E_V, conduction band E_C, and the Fermi level E_F, are shown. E_i is the energy of a generic charge carrier, and x is the spatial coordinate along the junction.

increase is small compared to the voltage increase. The capacitance of the structure can be estimated using the parallel-plates capacitor approximation. The structure is treated as a capacitor of thickness d, where d is the width of the depletion region:

$$C(V) = \frac{\varepsilon\varepsilon_0}{d(V)} = \sqrt{\frac{\varepsilon\varepsilon_0 q N_d}{2V}}, \tag{2.24}$$

$$C(V_{dep}) = \frac{\varepsilon\varepsilon_0}{d}. \tag{2.25}$$

The depletion voltage can be obtained graphically from a C-V scan if we plot $1\backslash C^2$ vs the voltage V, as shown in figure 2.8. The value of $1\backslash C^2$ grows linearly as a function of V while the sensor is not fully depleted and becomes a constant as soon as the detector is depleted, and from the position of the kink it is possible to extract the depletion voltage.
The expressions for the sensor capacitance are valid for a single sensor, and they provide the capacitance per unit area. In the case of a pixel sensor with neighbouring pixels it is also necessary to take into account the capacitance between the single pixel cell and the neighbouring pixel cells, the inter-pixel capacitance. The calculation of the inter-pixel capacitance is very complex and often relies on the use of a specific TCAD simulation. It is the largest contribution to the total capacitance of the structure.
The I-V scan is an important method for measuring the leakage current and the breakdown voltage of the sensor. The I-V curve of a sensor is very similar to the one of a p-n reverse bias junction. The current increases with the square of the applied voltage until depletion and then reaches a plateau region, until the breakdown voltage. If the voltage is high enough, the free carriers that are drifting have enough energy to free other carriers in between collisions, and create a charge carrier avalanche. If the avalanche is not quenched in some way, the sensor breaks with permanent damage. There is not a reliable way to estimate the

breakdown of the structure a priori, because very small defects in the structure back-plane or production or handling differences can change drastically the behaviour of the sensor.

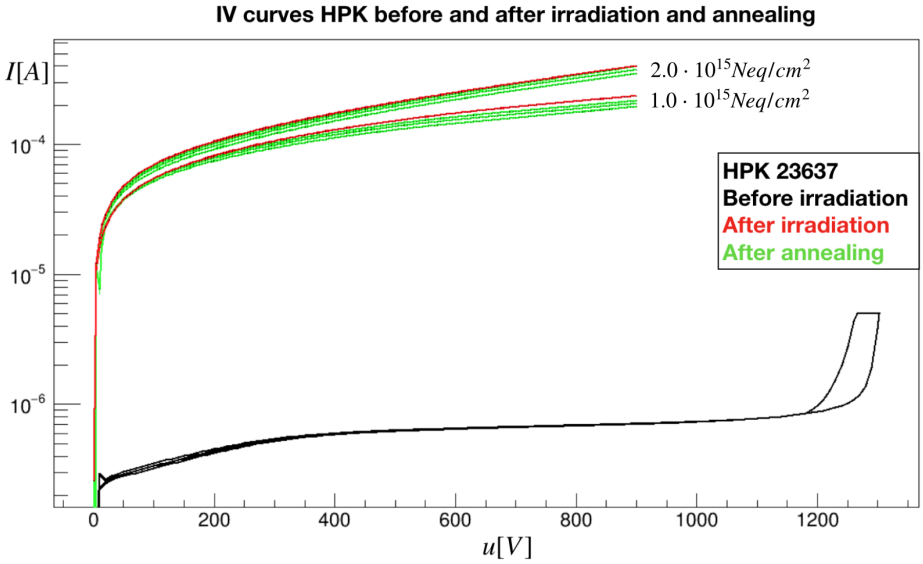


Figure 2.7: IV curves for a HPK pixel sensor before and after irradiation and annealing. The IV curve for the un-irradiated sensor is shown in black, and in red the curves after irradiation at two different radiation doses. In green, the curves after annealing are shown. The data and the plots have been acquired at UZH.

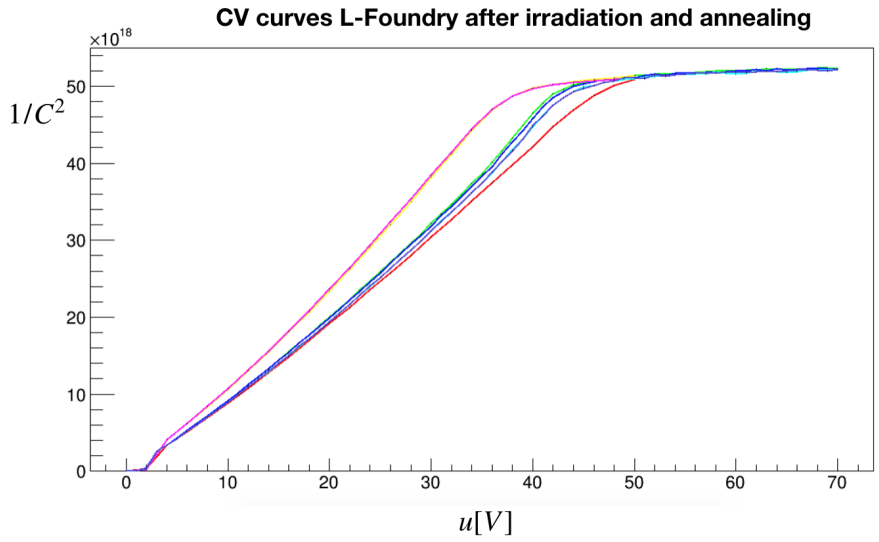


Figure 2.8: CV curves for LFoundry sensors after irradiation and annealing. The data and the plots have been acquired at UZH.

2.9 Pixel detector

A pixel detector is a high granularity sensor typically made of silicon. The sensor type most commonly used in HEP is a silicon pixel hybrid sensor. They are called hybrid because the active part of the sensor and the readout electronics are produced on different structures and then connected using bump-bonds or others technologies. They are the equivalent of the sensor commonly found in a digital camera, optimised for MIP particles instead of for visible photons, and for working in a high-rate, high-radiation environment.
The interest of high-energy physicists in pixelated detectors started from the necessity to study short-lived particles in a high collision-rate environment.
Short-lived particles, with a lifetime of the order of picoseconds, travel for a short length($l = c\tau \approx 3{\cdot}10^8\text{m/s}\cdot\ 10^{-12} = 300\mu\text{m}$) before decaying. To identify this kind of decay a detector with the following characteristics is needed:

- good spatial resolution of the order of 10 μm,
- high granularity, to record a high number of hits in a very track-rich environment and maintain low occupancy,
- high hit-rate sampling (the collision rate in a modern collider is of the order of tens of MHz).

Because of the high number of channels and hits, the amount of data typically delivered by a pixel detector is simply too big to be analysed for every collision. To bypass this problem there are different solutions. The CMS experiment for example uses a multi-level trigger system, and the pixel information is not used at the lowest, fastest level, the so-called level one trigger (L1). If the event passes the L1 trigger selection using information coming from the other sub-detector systems, only then are the data from the pixel sensors read out and further analysed, otherwise they are discarded.

2.10 Signal formation

The most commonly used technology in the production of pixel sensors is the use of fully-depleted radiation sensors. In these sensors the charge induced by one carrier on the readout electrode is described by the Shockley-Ramo theorem:

$$i = q\vec{E_w} \cdot \vec{v}, \tag{2.26}$$

where E_w is the weighting field can be computed by setting the electrode potential to 1 V and all the other potentials to 0. The drift velocity v is a function of the local electric field. In a flat geometry, because the drift direction and field are parallel, the scalar product in between the electric field and the carrier velocity is the product of the magnitudes.
This approximation is not true for a pixel sensor. To obtain the charge deposited in the electrodes, it is possible to integrate the previous equation over the readout time interval:

$$Q = \int_{t_1}^{t_2} i(t)dt = q[\Phi(x_1) - \Phi(x_2)], \tag{2.27}$$

where $\Phi(x)$ is the weighting potential at the initial and final point on the drifting trajectory. The weighting potential is obtained by setting the potential of the electrode to 1 and setting all other electrodes to potential 0. It is important to consider the drift of both electrons and holes in the reconstruction of the signal. The electrons are fast and contribute to the initial part of the signal formation, the fast rising edge. The slower holes, drifting towards the backside, also contribute to the signal formation. The current induced by the two different charge carriers type is shown in figure 2.9. Usually in pixel sensors the generated signal is

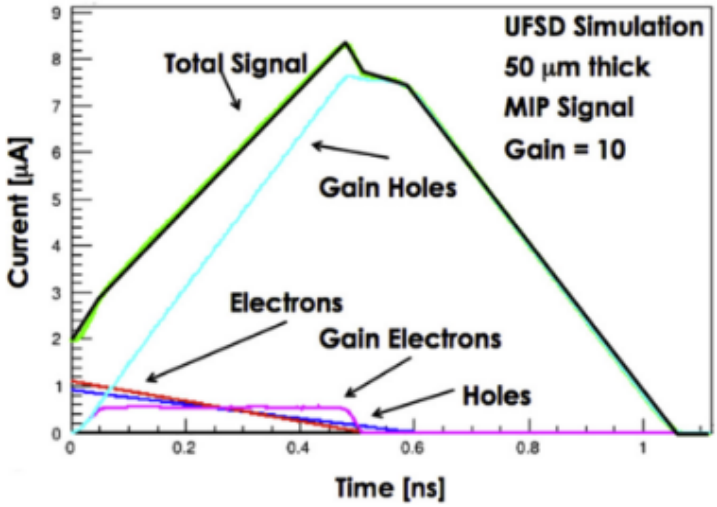


Figure 2.9: Signal formation inside a 50 μm thick low gain avalanche diode sensor [25].

integrated in a long time interval compared to the signal formation time and readout using a shaper, so most of the signal information is lost except for the charge.
In sensors developed especially with time resolution in mind, the readout electronics are faster and the signal shape is important. In this case the sensor capacitance plays an important role, because together with the readout electronics they are the limiting factors to obtaining a good timing resolution. The sensor capacitance in itself works like a shaper, and if the capacitance is too large the signal will be shaped by the intrinsic rise-time of the sensor.

2.11 Spatial resolution

The space resolution of a silicon sensor will be discussed in depth in the chapter dedicated to the space resolution simulation, especially in section 4.9. Here some basic concepts will be introduced, using as an example a detector with one pixel of size d. In this case there are two possible outcomes for the measurement of the position of the passing particle: the particle passes through the sensor, or the particle misses the sensor. In the first case, we can calculate the average distance between the particle and the centre of the pixel. If the particle passes through the sensor, from a statistical point of view, we can repeat the experiment many

times with particles uniformly distributed over the pixel length, and extract the variance of the distribution. The position distribution of a particle passing through the sensor will be flat in the region for x =0 and x =d, because we are assuming that the particle will have the same probability to hit the sensor in any point. It can be shown that in this case the resolution is

$$\sigma_{position} = \frac{d}{\sqrt{12}} \approx 0.28d. \tag{2.28}$$

This configuration delivers a binary resolution, because the two possible outcomes are hit or miss.
In the case of a pixel sensor, for a perpendicular track, almost the entire particle track will pass by a single pixel, and the resolution for this hit will be given by the binary resolution formula. As the particle hits a sensor at different angles, it will pass through multiple pixels. The charge deposit in each pixel is proportional to the length of the track inside the pixel, so if the readout electronics measures the charge inside the pixel, it is possible to derive the position of the particles with a charge weighting method. For a single pixel module, the hit position can be reconstructed using the charge centroid method as:

$$x = \Sigma_i \frac{x_i \cdot ch_i}{\Sigma_i ch_i}. \tag{2.29}$$

It is clear that the lower limit to the resolution is given by the charge resolution and by the statistical fluctuation of the charge deposit. In a typical pixel detector the charge has to cross a certain threshold and it is digitised using a limited number of bits. For a fixed pixel aspect ratio, the threshold, the electronic noise and the buffer size play the most important role in defining the resolution. There are more refined techniques for position resolution other than the charge centroid method. This technique relies on external information that is available in more complex geometries, for example with multiple detector layers.
The measurement of the timing resolution for silicon sensors will be discussed in Section 5.12.

2.12 Radiation-induced damage in the bulk of the sensor

Radiation-induced damage has three important effects on the sensor:

- increase of the leakage current,
- change of space charge in the depletion region and consequently of the depletion voltage,
- charge trapping.

These effects are due to the damage of the lattice structure of the silicon. Electrons and hadrons with an energy of the order of hundreds of keV are able to displace a silicon atom from the lattice. It is important to remember in the first place that the silicon used in HEP applications is mono-crystalline silicon. Any damage to the crystal creates intermediate energy levels in the bulk, which act as new generation and recombination centres, and this kind of damage is usually not reversible.

To describe the effect of radiation damage on the leakage current as a function of the fluence it is possible to use the Hamburg model [26], in which the new levels act as generation/recombination centers and the inverse of generation time τ_g is proportional to the fluence Φ:

$$\frac{1}{\tau_g} = \frac{1}{\tau_g \quad (\Phi = 0)} + k_\tau \Phi, \tag{2.30}$$

where the generation time is the average time between the generation of two charge carriers. Or

$$I = I(\Phi = 0) + \alpha\Phi, \tag{2.31}$$

with the current related damage rate expressed as:

$$\alpha = e n_i k_\tau. \tag{2.32}$$

The defects created inside the silicon lattice are not stable and can move around. The mobility of these defects is proportional to the temperature. To reverse some of these effects it is not uncommon to raise the temperature of the detector: this procedure is called annealing. The α constant depends on the temperature and annealing time, and is usually parametrised as:

$$\alpha = \alpha_i e^{-\frac{t}{\tau_i}} + \alpha_0 - \beta \ln \frac{t}{t_0}, \tag{2.33}$$

in which α_0 is the irreversible damage, and the others terms parametrize the reversible damage annealing on the shorter and longer time scale. The temperature dependence of α is hidden inside the τ_i terms, for which the following equation stands:

$$\frac{1}{\tau_i} = k_{0,i} e^{-\frac{E_i}{kT_a}}, \tag{2.34}$$

where T_a is the annealing temperature.
The change of space charge in the depletion region is complicated and is measured experimentally using the equation:

$$N_{eff} = \frac{2\varepsilon\varepsilon_0 V_{depl}}{ed^2}. \tag{2.35}$$

Experimental results obtained by irradiating an n-type silicon sensor are shown in figure 2.10 [27]. The full depletion voltage was obtained by the C-V curves as explained in the previous section. As seen in the figure, if we start with an n-doped sensor, the defects start building up in the sensor and behave like p-type defects. After a received fluence of $\sim$ 2 times 10^{12} Neq the number of defects introduced, due to radiation, exceeds the initial dopant concentration and the silicon type changes from n-type to p-type.
The change in effective doping concentration is described as:

$$N_{eff} = N_{eff,\Phi=0} - \left[N_C(\Phi) + N_A(\Phi, T_a, t) + N_Y(\Phi, T_a, t)\right], \tag{2.36}$$

where:

- N_C depends only on fluence and describes the constant damage to the sensor,
- N_A term is referred to as the beneficial annealing and describes the annealing behaviour over a short time scale,

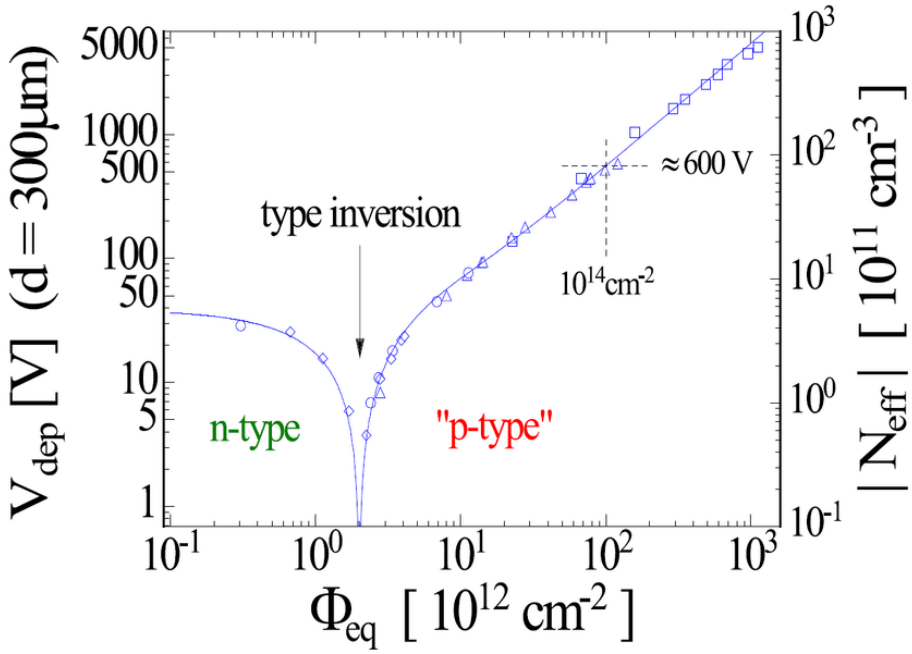


Figure 2.10: Effective doping concentration, and full depletion voltage, as a function of 1-MeV neutron equivalent fluence for a standard FZ n-type silicon detector, illustrating the type inversion phenomena [27].

- N_Y is the so-called reverse-annealing term and describes the sensor behaviour on a time scale longer than a week.

The last radiation effect to be considered is charge trapping. It plays an important role because if the charge is trapped for a time longer than the readout electronics integration time, the charge is not read out and the efficiency of the sensor starts degrading. Trapping can happen anywhere on the charge carrier's drift path, so charge carriers far away from the collection anode will have a larger probability of being trapped because they have to travel for a longer distance inside the sensor. Therefore, the charge collection efficiency of the sensor after irradiation is a function of the depth, and so it depends on the particle impact angle and point. In figure 2.11[28] the charge collection efficiency as a function of depth is shown for an un-irradiated and an irradiated silicon sensor at different voltages. From figure 2.11 we can conclude that the overall charge collection efficiency of the sensor is reduced and a higher operating voltage is necessary to counteract the effect of charge trapping.

2.13 Surface damage

The radiation damage in the silicon sensor can be split into two components, the surface damage and the bulk damage. The surface of the sensor is covered by an oxide layer and the crystal structure on the surface is highly irregular, so the behaviour of these two regions are quite different. In the oxide layer the mobility of the electrons is very high, while the mobility of the holes is very low. After the creation of hole-electron pairs, the holes start

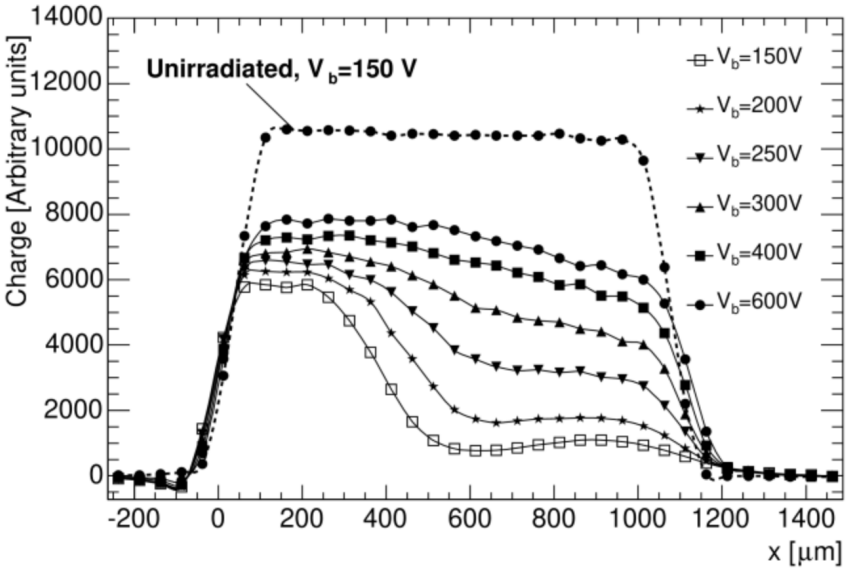


Figure 2.11: Charge collection profiles measured with an unirradiated (dashed line) sensor and a sensor irradiated to 6×10^{14} neq/cm^2 (solid line). The latter is operated at bias voltages between 150 V and 600 V [28].

drifting slowly, while the electrons drift to the closest electrode. When the holes reach the boundary region between oxide and silicon, they become trapped, and the effect is that the oxide boundary layer accumulates positive charge.

2.14 Different pixel technologies

2.14.1 p^+ in n

In the previous sections only p-n junction have been considered. It can be shown that a p^+-p structure or a n^+-n structure behaves in the same way, but there are differences in the production and operation of the structures.
The first choice for building a structure is the kind of substrate, p or n-type.
The first silicon sensors developed for HEP used a p^+ implant in an n resistive substrate, the so called p^+ in n. This approach is technologically easy, and the pixels are electrically isolated because they form a reverse bias p-n junction. Since the electrodes in this configuration are collecting holes that are slower, this kind of structure is less radiation-hard than those described in the following sections.

2.14.2 n^+ in n

The n^+-n offers a number of advantages over the p^+ -n but is more expensive.

In this case, the depletion starts before irradiation from the large back-plane p^+ implantation. Therefore a set of guard rings (GRs) has to be implemented on the back-side, requiring double-sided processing. The GR stop the depleted volume from reaching

the edges of the structure, which are conductive after dicing. These sensors require over-depletion to operate properly before type inversion. After type-inversion, the depletion region starts from the front-side, so that the sensor could be operated also in an under-depleted mode retaining a good collection efficiency. The different behaviour of these sensors before and after irradiation is shown in figure 2.12. It is know by now that type inversion is an approximation and a better description of the phenomena is given by the creation of a double-peaked electric field inside the silicon sensor [29].

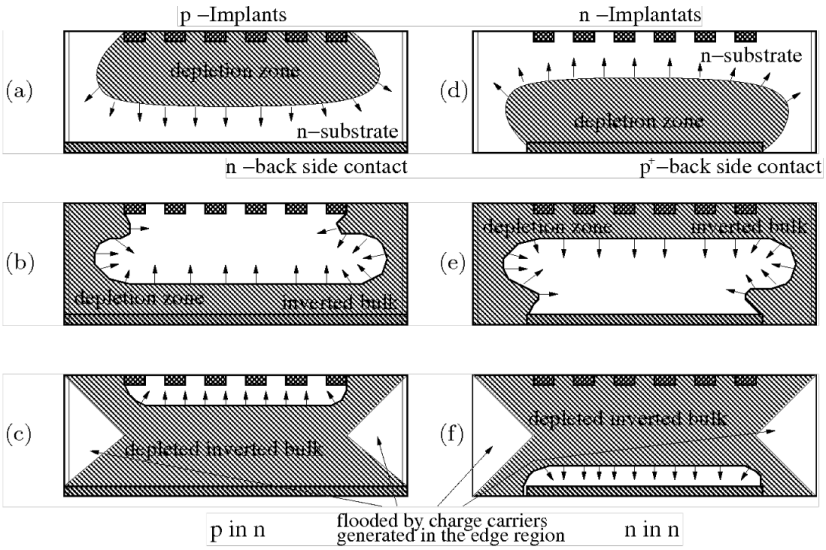


Figure 2.12: Comparison of the p^+ in n (a) and (c) and n^+ in n (d) and (f) sensors. Partial depletion (a) and (d) before and (b) and (e) after type inversion. The generation of electron-hole pairs in heavily damaged edge region is also taken into account [19].

2.14.3 n^+ in p

The n^+-p is an interesting technology. It does not require a double-sided manufacturing process, it does not undergo type inversion, and it collects electrons. This is the sensor technology chosen for the CMS Phase-2 pixel detector upgrade.

2.15 Additional considerations for pixel modules

In the case of pixels, the simple diode model is still valid but additional considerations are needed. The pitch is the distance between neighboring pixels, and the width is the active area of a pixel.Due to design aspects, there is sometimes a non-negligible difference between the two, causing a dead area between pixels. The ratio between the width and the pitch is commonly called the fill factor. A large fill factor is desirable in applications in which a good field uniformity is required, because as explained, the depletion zone grows starting from the n^+ implants, so the different depletion zones of individuals pixels have to merge to create a uniform depletion zone.
On the other hand, the largest contribution to the pixel capacitance is the inter-pixel ca-

pacitance, and it is directly proportional to the length of the implant perimeter.

2.16 Pixel isolation

In n-in-n and n-in-p technologies, the positive charges of the oxide attract electrons which accumulate between the pixel implants. When $n+$ implants are used for the segmented side of the sensor, this electron layer creates a short of all pixels. In order to isolate pixels from each other, a Boron implantation is deposited between the pixel electrodes. This can be achieved by the use of an additional lithographic mask, as in the phase-1 $n+$-in-n sensors, where a higher Boron density is created in a narrow strip between the pixels (p-stop), or with a homogenous Boron implant (p-spray) as in the case of some n-in-p pixel prototypes for phase-2. Depending on the p-stop or p-spray dose, a loss of isolation could be in principle experienced, due to increased positive charge density accumulated in the oxide layer.

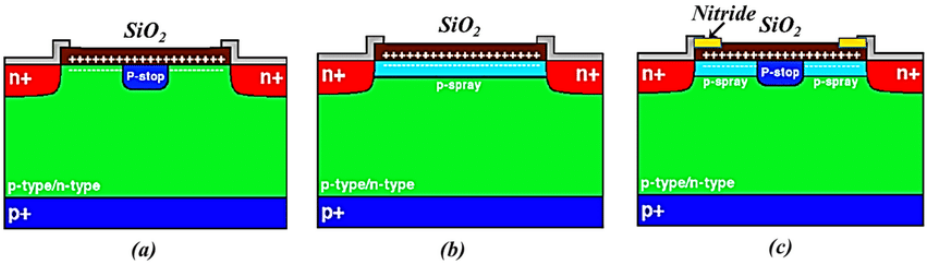


Figure 2.13: Sketch of different implementations of an inter-pixel isolation layer: (a) p-stop, (b) p-spray, (c) p-stop.

Part II

PhD research projects

Chapter 3

Integration of the CMS Phase I Pixel Detector

3.1 Introduction

The CMS experiment at the CERN Large Hadron Collider (LHC) features a silicon pixel detector as its innermost device [30]. It provides precision space-point measurements of charged particle trajectories within a pseudorapidity of up to 2.5. The original pixel detector, which is referred to as the Phase-0 pixel detector, consisted of three concentric cylindrical layers known as the barrel pixel detector (BPIX) and two end-cap disks for both the +z and -z forward regions (FPIX) [31].
The specifications of the pixel detector matched the original LHC design luminosity of $1\times10^{34}\text{cm}^{-2}\text{s}^{-1}$ and a bunch spacing of 25 ns. Due to continuous improvements of the LHC, the design luminosity was already exceeded in 2016. As a result of the increase in instantaneous luminosity and the limited readout bandwidth of the detector, dynamic inefficiencies arose. At the nominal level-1 trigger accept rate of 100 kHz, the data loss would increase to 16 % in the innermost layer for luminosities of $2\times10^{34}\text{cm}^{-2}\text{s}^{-1}$ with a 25 ns bunch crossing [32]. Figure 5.1 shows how the single hit efficiency decreased with increasing luminosity for all layers during the 2016 data taking period. By the end of 2016, a drop of 5 % in the innermost layer was already evident [33].
To maintain and even improve the tracking performance, CMS developed and built a new pixel detector as a part of the wider Phase-1 upgrade campaign [32]. The new pixel detector was successfully installed in March 2017 during an extended year-end technical stop (EYETS) of the LHC. The Phase-1 pixel detector features four pixel layers for the barrel region and three disks for each forward region. The barrel and forward sub-detectors are independent systems, with dedicated electronics, mechanics, cooling and powering systems. The barrel pixel detector is referred as the Phase-1 BPIX and the forward pixel detector as the Phase-1 FPIX.
The upgraded detector improves on many aspects with respect to the previous one, featuring digital instead of analog readout chips, a lighter weight cooling solution, a smaller material budget in the central tracking region, and an increased number of channels, each with a larger readout bandwidth.

The services for the barrel pixel detector are contained in four half-cylindrical shells, two on each side of the detector, called the supply tubes (STs), similarly the FPIX detector and its services are contained in four half-cylindrical shells that are inserted inside the STs. Figure 3.2 shows an exploded-view of the BPIX and FPIX detectors, and of the BPIX and FPIX detector service cylinders.

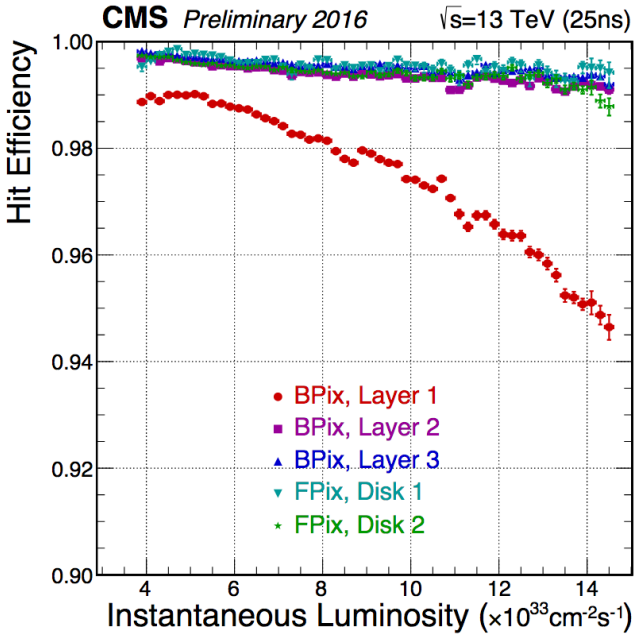


Figure 3.1: Measured single-hit efficiency per layer as a function of the instantaneous luminosity. The data were taken with the Phase-0 pixel detector during 2016. [33].

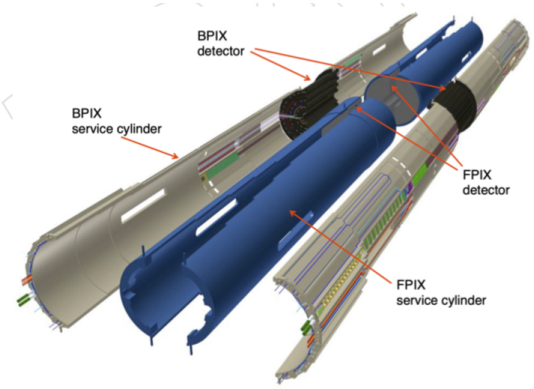


Figure 3.2: Drawing of the Phase-1 BPIX and FPIX detectors together with the service half-cylinders that hold the readout and control circuits as well as the power and cooling lines.[32]

The STs provide cooling and power, while allowing communication with and control of the detector. The STs have been developed, tested and built at the University of Zürich (UZH) in collaboration with the Paul Scherrer Institute (PSI), Eidgenössische Technische

Hochschule Zürich (ETH), RWTH Aachen University, Eötvös Lorand University and CERN. During the development phase, my role was to replicate a section of the ST in a test stand at UZH. The test stand was comprised of the ST electronics, some pixel modules, and a slice of the CMS DAQ system. I tested the ST electronics and developed the necessary verification and calibration procedures. I helped assemble and test the four final ST half-cylinders at UZH before they were shipped to PSI, where they were connected to the BPIX detector modules. I was part of the team that tested the system as a whole after the transport and assembly. The system was then moved to a cleanroom near the CMS experiment at CERN where the final set of calibrations took place. Finally, the system was installed inside the CMS experiment. In the following sections, the different design, assembly, and testing phases will be presented in detail. Finally, the performance of the pixel detector will be discussed.

3.2 Phase 1 pixel detector overview

The Phase-1 pixel detector has been optimised to maintain and improve the performance over the previous Phase-0 pixel detector. The new detector features:

- An increased number of spatial measurements by using a fourth layer in the barrel region and an additional disk in each of the forward regions.
- An innermost layer of the barrel detector that is closer to the beam pipe than the original detector, at a radius of 29 mm instead of the 44 mm.
- An upgrade from the original C_6F_{14} liquid-coolant that was distributed in bulky pipes to a two-phase CO_2 coolant system distributed in thinner, lighter weight pipes.
- A lower material budget, in part because of the new cooling system, and in part to the moving of the electronics to a higher z outside the fiducial detector volume.
- A new readout and backend electronics with a doubled readout bandwidth, moving from a 40 MHz analog readout with 6 discrete levels to a 400 Mbit/s digital readout.
- A new readout chip that includes hit digitization and with higher data rates as compared to the previous design.

In figure 3.3, a comparison of the Phase-0 and Phase-1 pixel detector layouts is shown [32]. The new pixel detector is built from 1856 segmented silicon sensor modules, of which 1184 modules are used in the BPIX and 672 modules are used in the FPIX. Each module consists of a sensor with 160×416 pixels connected to 16 readout chips (ROCs). The total number of channels is almost double that of the Phase-0 pixel system (from 66 million to 124 million). The improved detector layout gives a four-hit coverage in the whole tracking region up to $|\eta| < 2.5$. The smaller radius of the closest pixel layer enhances the vertex resolution, the b-tagging efficiency with respect to the misidentification rate, and the track seeding performance. In order to accommodate the reduction in radius of the innermost pixel layer, a smaller diameter beam-pipe had been previously installed during the CERN

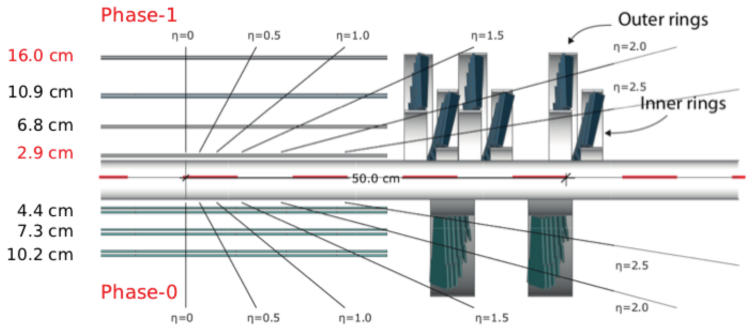


Figure 3.3: Comparison of the layouts of the Phase-1 detector (top) and the Phase-0 detector (bottom) [32].

Long Shutdown 1 from 2013 to 2014. To facilitate installation, the BPIX detector is divided into two mechanically independent half-barrel units, similarly, the FPIX detector consists of two mechanically independent units for each side. Thanks to this design choice it was possible to install the Phase-1 pixel detector in the limited period of time during the extended technical stop during winter 2016/2017 without removing the beam pipe. Furthermore, the mechanical separation allows for an easier extraction of the detector for maintenance work and refurbishment during short periods of LHC EYETS. The new pixel detector fits into the same mechanical envelope as the original system and partly reuses existing services.
The barrel pixel detector is divided in two half-barrels with a length of 540 mm, each one consisting of four layers (called half-shells). The forward pixel detector consists of twelve half-disks with a radial coverage ranging from 45 to 161 mm. Each half-disk, shown in figure 3.3, is further divided into inner and outer half-rings each supporting 22 and 34 modules. The position of the layers and the disks are summarised in table 3.1 [34]. The larger number of channels compared to the original pixel detector leads to a corresponding increase in the current consumption. As a result, the power losses in the cables were expected to increase by a factor of 3.6. To overcome this problem, a new power system based on DC-DC conversion was developed [35]. The existing power supplies inside the service cavern were modified to deliver a voltage of 11 V to the DC-DC converters that are located on the pixel STs, outside the tracking volume but still within the CMS detector. The DC-DC converters provide the required digital and analog voltages of 3.3 V and 2.4 V to the pixel modules. With this new power system the losses were reduced by a factor of around 10. About 1200 DC-DC converters were used in the upgraded pixel detector, with one pair of converters powering up to four modules.
The data acquisition (DAQ) back-end electronics located in a cavern nearby the detector were also upgraded to be compatible with the new detector. The VME-based system was replaced by a MicroTCA-based system that supports higher data rates.
The material budget of the Phase-1 pixel detector was reduced in different ways. The front end electronics located on the STs have been moved outside of the sensitive region of the tracker. The mechanical structure uses lighter carbon fibre composite materials and airex

BPIX			
Layer	Radius	z position	Number of modules
L1	29 mm	-270 to +270 mm	96
L2	68 mm	-270 to +270 mm	224
L3	109 mm	-270 to +270 mm	352
L4	160 mm	-270 to +270 mm	512
FPIX			
Disk	Radius	z position	Number of modules
D1 inner ring	45-110 mm	± 338 mm	88
D1 outer ring	96-161 mm	± 309 mm	136
D2 inner ring	45-110 mm	± 413 mm	88
D2 outer ring	96-161 mm	± 384 mm	136
D3 inner ring	45-110 mm	± 508 mm	88
D3 outer ring	96-161 mm	± 479 mm	136

Table 3.1: Top: For the BPIX detector, the radius of each layer, its Z position range, and its number of modules. Bottom: For the FPIX detector, the inner and outer radii of each FPIX disk, its Z position, and its number of modules [34].

foam [36], which provides less material for particle scattering.
A two-phase CO_2 cooling system replaced the previous C_6F_{14} liquid cooling system. Thanks to the thermodynamic properties of the CO_2, it is possible to use smaller diameter pipes, since the intrinsic lower density of CO_2 with respect to C_6F_{14} further reduced the material budget in the fiducial region of the detector. Figure 3.4 shows a comparison between the material budgets of the CMS Phase-1 pixel detector and the Phase-0 pixel detector, both in terms of radiation lengths and hadronic interaction lengths as a function of pseudorapidity η [34]. The CMS collaboration aims to collect an integrated luminosity of 500 fb^{-1} before upgrading the tracker. During this period, the innermost barrel pixel layer is expected to accumulate a hadron fluence of 3.6 $\times$ 10^{15}neq/cm^2 (fluence is measured in units of 1 MeV neutron equivalents). Because this fluence is about twice the operational limit of the system, the innermost BPIX layer is being replaced during CERN's Long Shutdown 2. Since the collected fluence as a function of r follows the inverse square law, the second and outer layers do not need to be replaced and will operate for the full data-taking period. The same is true for the modules in FPIX, whose inner ring is at a larger radius than that of the inner BPIX layer. The CMS Phase-1 pixel detector has been developed to match and out-perform the original detector at twice the instantaneous luminosity.
In the following sections I will focus on the barrel pixel detector upgrade, in which I was personally involved.

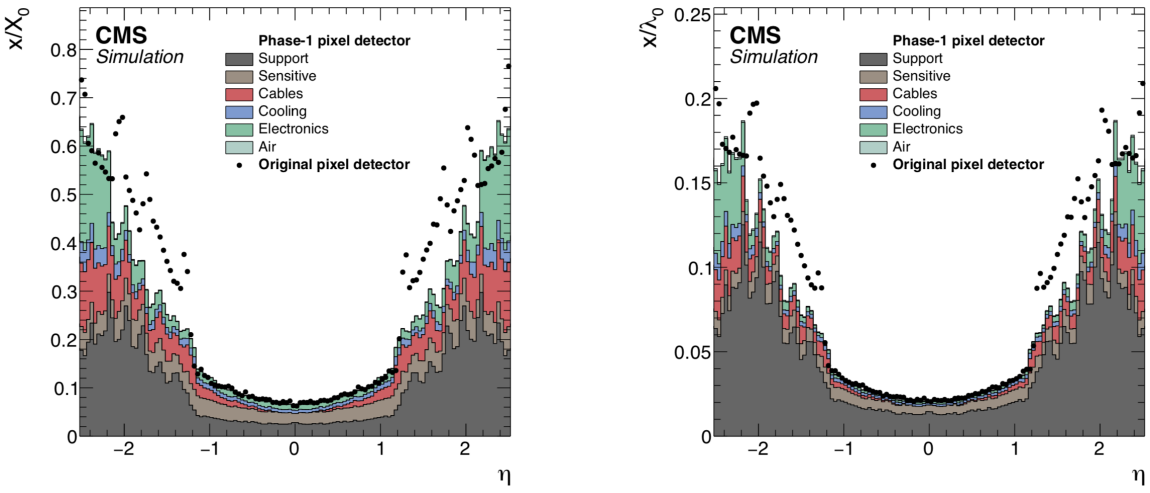


Figure 3.4: The amount of scattering of particles in the Phase-1 pixel detector (stacked histograms) is shown compared to the Phase-0 pixel detector (dots). On the left, this material budget is shown in terms of radiation lengths X_0, whereas on the right, hadronic interactions lengths λ_0 are shown. The different colors in the stack of histograms represent different components of the detector. The largest reduction in scattering is seen to be between $1.2 < |\eta| < 2$. [34]

3.3 Phase 1 pixel module

The Phase-1 pixel module is what is known as a "hybrid" silicon assembly [37]. With this technology, the silicon sensors and the readout electronics are produced on different wafers, and connected with bump bonds. Figure 3.5 shows the components that comprise the BPIX detector modules. The main components are the sensor, the readout chips (ROCs), the high density interconnect (HDI), the token bit manager (TBM), and the twisted pair cable.

The active part of the module is a planar silicon sensor produced with the n^+-in-n technology, and having an active area of 16.2×64.8 mm^2. The sensor is segmented in 66650 individual pixels. Each single pixel cell size is 100×150 μm^2, with a thickness of 275 μm. The silicon sensor is bump-bonded to 16 ROCs. Each ROC hosts 4160 pixels and readout channels. In order to maintain detector sensitivity at the physical interface between readout chips, which has a small gap, the area of the pixels on the sensors is increased by a factor of two along the edges of the ROCs, and by a factor of four on the corners of the ROCs. The ROCs are connected to an HDI that is a flexible printed circuit, which is glued on the backside of the sensor, such that the connections are made by wire bonds. A chip called the TBM is used to receive Level-1 triggers from the global CMS trigger and send tokens to the ROCs in order to receive their data. Layer-1 modules require two TBMs due to the higher rate, while other modules require only one. The ROCs collect the hit information from the sensors and store it until a trigger is received by the TBM. The TBM then receives the data stream from the ROCs, organizes it, and sends it via a twisted-pair cable to subsequent readout electronics. In the following sections, each of these components, as well as the method of data collection and data streaming, will be discussed in more details.

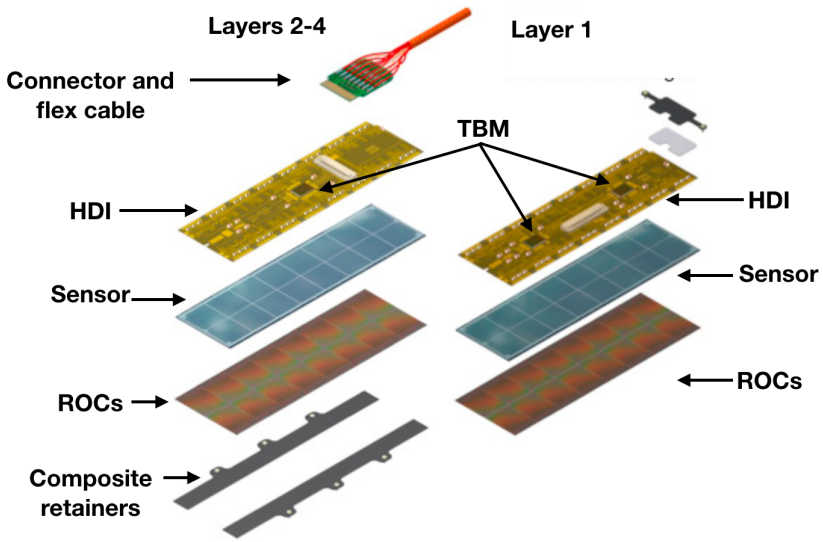


Figure 3.5: The pictures show the stack of different components of the modules [37].

3.3.1 Phase-1 sensor

The Phase-1 pixel detector uses the same sensor used in the original BPIX pixel detector modules. The silicon sensor is produced using the n^+-in-n technique. The strongly n-doped (n^+) pixel implants sit on top of an n-doped silicon bulk and a p-doped back-plane. The pixel cells are isolated by a moderate p-spray termination and are connected to a bias grid. The bias grid is connected to ground using a punch through. This configuration allows the sensor to be biased before any further processing and has been used to test the sensor quality before proceeding to the bump bonding. The BPIX sensors are produced on n-doped 4-inch wafers that were produced using the float-zone (FZ) process. The wafers are approximately 275 μm thick, the same thickness as the final sensor. All BPIX sensors were processed using the same silicon ingot, to minimise variation inside the batch. Wafers were probed and accepted when at least two of the sensors fulfilled the specifications. The most critical requirement was the limit of a maximum current of 2 μA at 150 V reverse bias voltage and a temperature of +17 $°C$.

3.3.2 Phase 1 readout chips

The Phase-1 pixel detector faces twice the instantaneous luminosity of the original pixel detector, and the ROC had to be completely redesigned. The ROCs used in the Phase-1 pixel detector are manufactured using 250 nm CMOS technology. There are two different ROCs, PROC600 for the BPIX layer-1 and PSI46dig for the rest of the detector, and the design specification are shown in table 3.2 [34]. The PROC600, which refers to the specification of 600 MHz/cm, is built to offer the same performance as of the PSI46dig but in a higher hit-rate environment.

	PSI46dig	PROC600
Detector layer	BPIX L2-L4 and FPix	BPIX L1
ROC size	7.9 mm × 10.2 mm	7.9 mm × 10.6 mm
Pixel size	150× 100 μm^2	150 × 100 μm^2
Number of pixels	52×80	52×80
In-time threshold	2000 e$^-$	2000 e$^-$
Pixel hit loss	$<$ 2% at 150 MHz/cm^2	$<$ 2% at 580 MHz/cm^2
Readout speed	160 Mb/s	160 Mb/s
Radiation tolerance	120 Mrad	120 Mrad

Table 3.2: The design requirements for the readout chip PSI46dig, used in BPIX layers 2-4, and PROC600, used in L1 [34].

The ROC pixel matrix consists of an array of pixel unit cells (PUC) arranged in 260 double columns, which are controlled by the double-column periphery that provides a general hit buffer for all double-columns. The three main units of the ROC are the double columns, the double-column periphery, and the chip periphery. When a particle traverses a sensor, the accumulated charge is collected by the PUCs, which are connected to the pixel cells via bump bonds. To calibrate the detector, a calibration signal can be directly injected in the PUC using a dedicated circuit. The signal is amplified and shaped on the PUC, and the analog pulse height is compared in a comparator against a set threshold. Comparator thresholds are set globally for all PUCs by a DAC, but can be modified for each pixel using dedicated 4-bit DAC registers. When the comparator threshold is satisfied, the hit information is stored, and in this moment the corresponding pixel becomes insensitive to new data, while the double column periphery is notified of the hit. Upon notification, the double-column periphery writes a time-stamp that identifies the bunch crossing and issues a readout token to the pixel to initiate the readout. The pixel sends the hit information, including the signal pulse height and the pixel address, using a column drain mechanism. The data are digitised in the double-column periphery and stored in a 64 × 23 bit register until the potential arrival of a Level-1 trigger-accept (L1A) signal issued from the CMS level one trigger. If a L1A is received, the double-column periphery compares the time stamp of the pixel hit with the time counter in the L1A signal message. If the time stamp of the hit and the L1A counter are in agreement, the periphery is set to readout mode and then awaits a read-out token sent from the TBM. When the readout token issued by the TBM arrives at the double-column periphery, the validated data are sent to the chip periphery and the double column is reset. The position and charge of all hit pixels are registered with a time resolution of 25 ns, and the information is stored on-chip up to (4.15 μs), which is the Level-1 trigger latency. The behaviour of the PSI46dig can be tuned using 19 digital-to-analog converter (DAC) registers that are programmed using a 40 MHz serial bus. The ROCs need two different power supplies, +2.5 V and +1.5 V, for the digital and analog internal circuits.

The main improvements for the PSI46dig compared to the PSI46, that was used in the

Phase-0 detector, include:

- The increase in the size of the buffer for hit information and time-stamp storage.
- The implementation of an additional readout buffer stage to reduce dead time during the column readout.
- An 8-bit successive approximation analog-to-digital converter (ADC) running at 80 MHz has been implemented in the PSI46dig. Digitised data are stored in a 64 $\times$ 23 bit first-in-first-out register, which is read out serially at 160 MHz. A phase-locked loop (PLL) circuit has been added to derive the 80 and 160 MHz clock frequencies from the LHC clock.
- The adoption of 160 Mb/s digital readout instead of the previous 40 MHz analog readout rate.

The improvement in the design of the charge discriminator reduces the cross talk between pixels, as well as the time walk of the signal, leading to a lower threshold operation (below 1500 e^- with noise at less than 100 e^-). For comparison the Phase-0 readout chip threshold was set to 2000 e^-.

3.3.3 Phase 1 readout chips radiation tolerance

The ROC's radiation tolerance has been tested after irradiation up to 150 Mrad for the PSI46dig, and up to 480 Mrad for the PROC 600, using the proton beam of the Zyklotron AG in Karlsruhe (ZAG), Germany. The pixel-hit efficiency as a function of the hit rate for the PROC600, PSI46dig, and PSI46 are shown in figure 3.6 [38]. The single-pixel hit efficiency of the chips has been tested by injecting a calibration signal in the pixel, while exposing the PSI46dig sensor to high-rate X-rays up to 300 MHz/cm^2, and the PROC600 up to 600 MHz/cm^2. The threshold and noise characteristics of the chip remains unchanged after irradiation, showing excellent resilience to radiation damage, while the data losses are less than 2 % [38], as required by the ROC specification shown in table 3.2.
To achieve this performance, the PROC600 double column strategy was redesigned. In the PROC600, pixels within a double column are dynamically grouped into clusters of four and read out simultaneously to enable faster readout. The readout is zero-suppressed.

3.3.4 Token bit manager

The main functions of the token bit manager (TBM)[39] are to receive the L1 trigger accept signal, to distribute readout tokens to the ROCs, and to mutiplex the data received from the ROCs. When a L1A signal is received, it will be placed on a 32-deep stack. Normally the stack will be empty but is needed to accommodate high burst rates due to either noise, high track-density events or trigger bursts.
Then the TBM issues readout tokens to the ROCs and the readout starts. The readout tokens issued by the TBM are circulated sequentially to the different ROCs and are then returned to the TBM. The TBM multiplexes the data from the ROCs, and adds a header

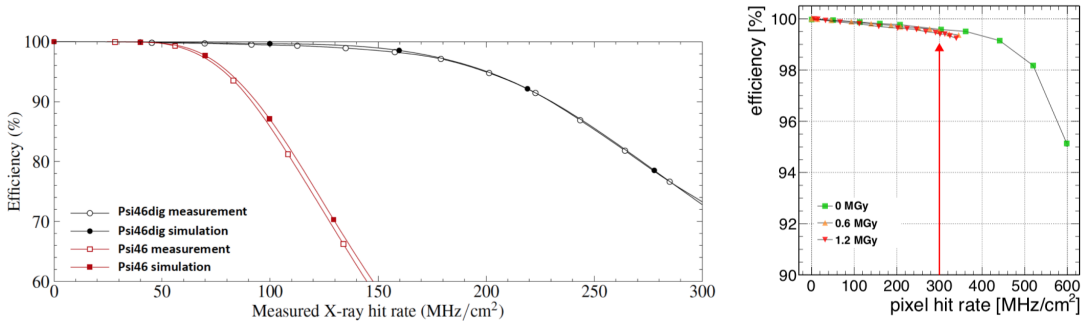


Figure 3.6: Left: PSI46dig efficiencies with x-rays at different hit rates. Right: Measured efficiencies for the PROC in high-rate proton beam. [38].

and a trailer that identifies the beginning and end of the data stream. This is a common feature in variable-length communication protocols. The TBM then drives the data through the readout link.

The multiplexing scheme is shown in figure 3.7 and works as follows: two 160 Mb/s ROC signal paths, with one path inverted, are multiplexed into a 320 Mb/s signal, then encoded into a 4-bit/5-bit Non-Return to Zero Inverted (NRZI) 400 Mb/s data stream. Encoding the data using this scheme randomise the bit patterns ensuring a transition between logical 0 and logical 1. This is necessary because is complicated to accurately count long sequences of 0 or 1 without transitions. There are two possible outcomes for the final encoded data for the symbol A to avoid a sequence of six ones that would be seen as a reset from the FED. The switch from the Phase-0 analog readout, with two analog links per module at 40 MHz, to the Phase-1 digital readout at 400 Mb/s, with four links per module in BPIX L1, yields an increase in readout bandwidth of a factor of four.

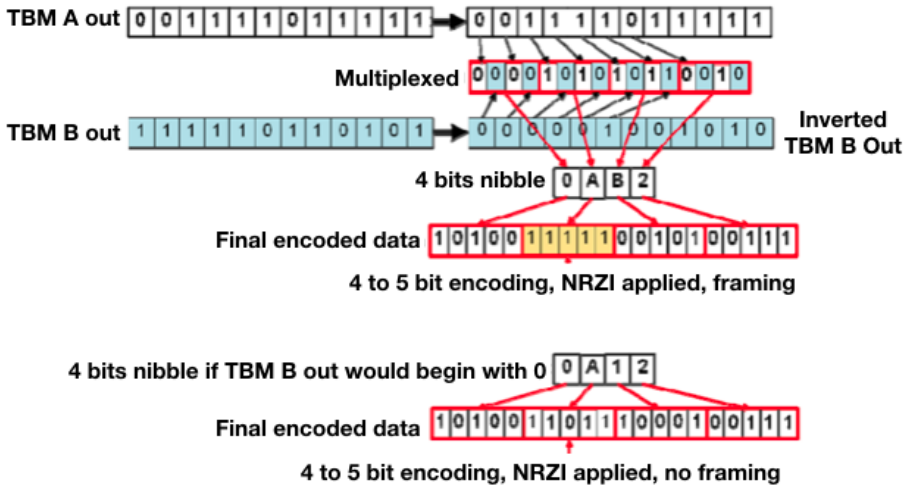


Figure 3.7: Encoding example.

3.4 Phase 1 pixel detector supply tubes

The phase-1 barrel pixel detector is connected to four cylindrical half shells called supply tubes (STs). Figure 3.8 shows the CAD rendering of one supply tube. The STs host the BPIX readout electronics, and interface the sensors to the data acquisition system (DAQ), while providing cooling and power, and allow the programming of the module's readout chip. Each ST is divided into three segments, A, B, and C, and segments are further divided into eight sectors that span the length of the ST, such that four STs have a total of 32 sectors. Each sector is an autonomous unit that controls and serves up to 39 detector modules. Figure 3.9 shows the mechanical structure of the supply tubes during construction. The

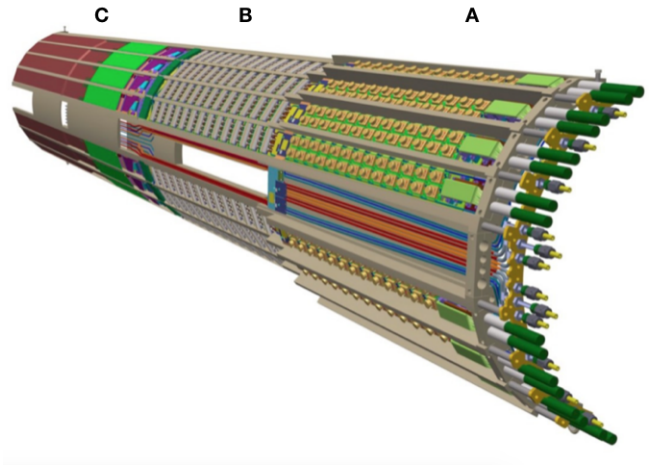


Figure 3.8: One of four pixel barrel supply tubes, showing segments A, B, and C, with each segment having eight sectors. The detector would be located to the left of segment C. Module cables are connected in segment C, both on the inside and outside of the shell. The auxiliary electronics are located in segment B, and the DC-DC converters in segment A, furthest from the interaction region. Cooling pipes are shown entering segment A on the far right, and loop under the auxiliary electronics and DC-DC converters. Cooling pipes for the detector modules can be seen in the central sector, which has an opening in segment B for the beam pipe support. Segment D, not shown, is closest to the detector, and therefore is built as a separate unit using carbon composite material to provide the lowest mass.[32].

mechanical structure is built using composite carbon fiber and AIREX foam, lightweight stiff materials. The supply tubes carry cooling pipes that connect to the detector cooling tubes. The fluid in the cooling pipes absorbs the thermal power of the modules and of the ST electronics. The cooling pipes are arranged in closed loops. The inner diameter of the pipes is 1.7 mm in the detector region with a wall thickness of 50μm. The cooling pipes on the ST have an inner diameter of 2.6 and 1.8 mm for the return and supply lines. The wall thickness is 200 μm. The connection of the cooling pipes between the supply tubes and the detector is achieved using small-mass compression fittings. The cooling loop schematic for one of the supply tubes is shown in figure 3.10 [34]. The total length of one supply tube is 1.7 m and its internal diameter is 175 mm. It is divided in three different segments, A, B, and C, with different outer diameters ranging from 190 to 215 mm. An additional

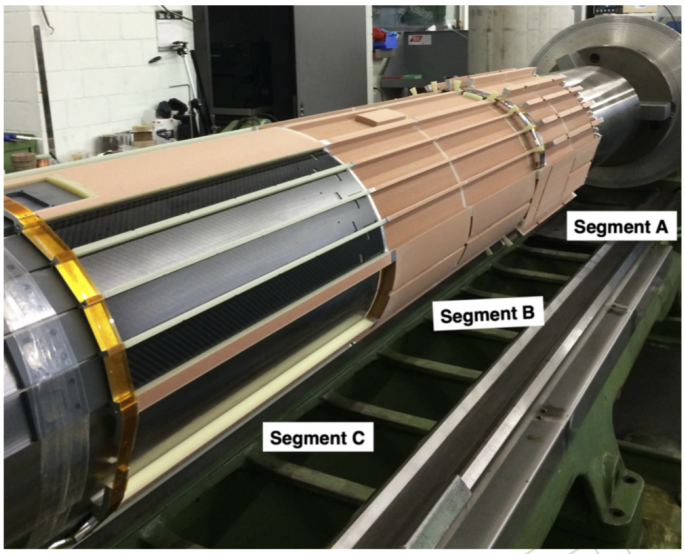


Figure 3.9: Picture of the Phase-1 BPIX supply tubes support structures on the mandrel in the UZH workshop during constructions. [34].

segment, segment D, is inserted between the supply tubes and the detector to arrange the twisted-pair cables that connect the detector modules to the supply-tube electronics.

3.5 Phase 1 pixel detector electronics

The readout architecture of the CMS Phase-1 barrel pixel detector is organised in 32 independent readout groups or sectors. The supply tubes host the front-end electronics that serve as an interface between the modules and the DAQ back-end electronics in the underground counting room. A block diagram of the electronics architecture for each sector is shown in figure 3.11.
The figure shows the different elements of the readout, powering, and programming chains. It is divided into four rows, corresponding to different physical locations on the readout driver chains. Starting from the top we have the modules in the detector region, the connector board and electronic devices on the supply tubes, while on the bottom, the Front End Controllers (FECs) and Front End Driver (FED) in the underground counting room.
The system uses two different controllers, the pixel FEC (pxFEC) and the tracker FEC (trFEC). The trFEC controls the supply tube electronics via standard I^2C communication protocol. The pxFEC is used to program the modules and send the clock and the trigger signals. The control of the modules is achieved using the fast I^2C communication protocol. The FED receive the data stream from the modules.
The data stream is converted from electrical to optical signals by the Pixel Opto Hybrid (POH) on the supply tubes. Optical fibres carry the signal to the counting room. Similarly, the clock, the trigger, and the control signals enter as optical signals from the control room, and are converted to electric signal by the DOHs on the supply tubes. This is necessary because the optical fibres that connect the counting room and the detectors are several tens

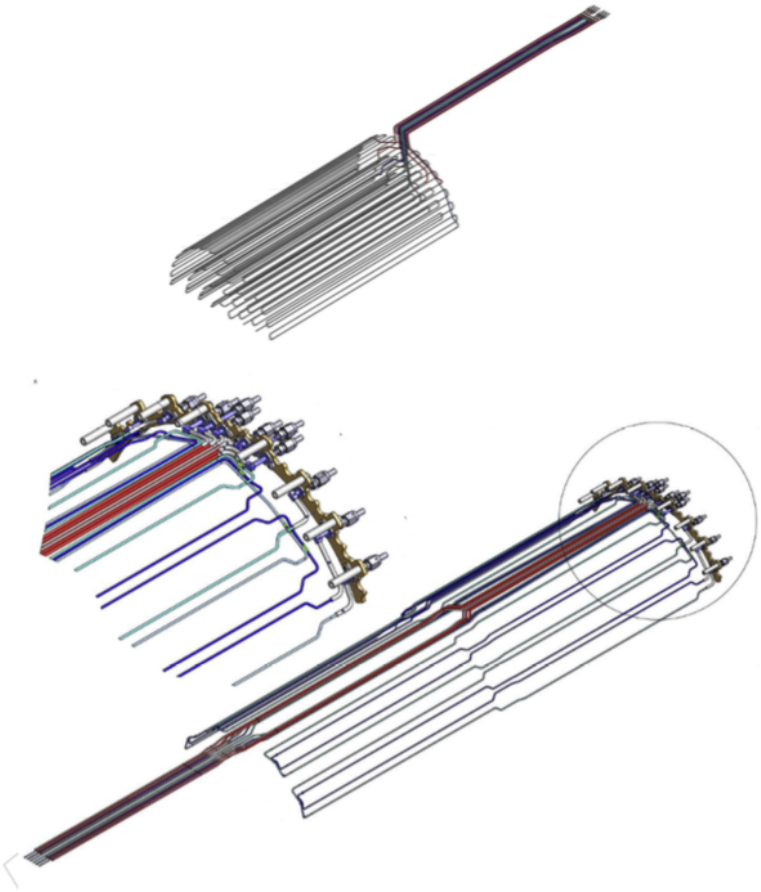


Figure 3.10: Drawing of one quarter of the cooling loop system for the BPIX detector (top) and the service half-cylinder (bottom). The segments of the pipes from the detector and ST are joined through compression fittings [34].

of meters long and optical signals can travel this distance with limited loss, compared to an electrical signal generated using similar power.

I will now describe the flow of the signals and data of the two control chains and readout driver chain shown in figure 3.11. The pxFECs are located in the underground counting room and are connected via optical fibres to the supply tubes. The pxFEC is used to program and read back the status of the modules. This is achieved by I^2C communication, using a configuration of programming and receiver data ports, the send-data A (SDA) and receive-data A (RDA) ports. The optical signals sent from the pxFEC are converted into an electrical data stream by the pixel digital opto-hybrid (pxDOH) and vice versa.
The pxFEC supplies the 40 MHz digital clock to the system and the L1A trigger. The clock and the trigger are encoded in a single data stream. The encoding logic is shown in figure 3.12 [40].

The trigger and the clock are decoded in two separate data streams by a combination of a phase-locked loop (PLL) circuit, and an additional synchronisation circuit, the quartz crystal PLL (QPLL). These two circuits use the LHC clock as reference clock.
The clock, the trigger, and two programming lines, SDA and RDA, pass through a delay chip,

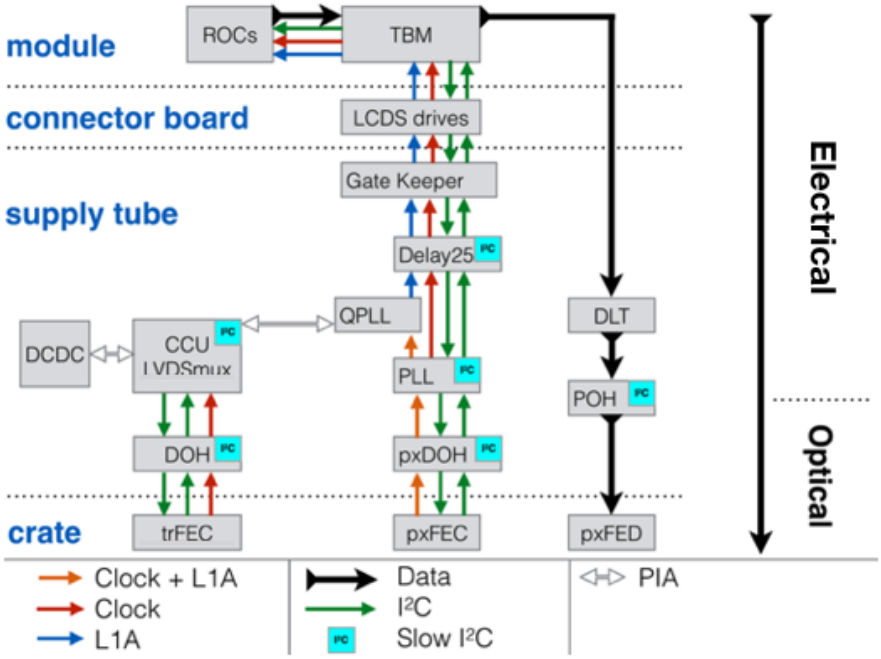


Figure 3.11: Overview of the readout and control system of the CMS Phase-1 pixel detector.

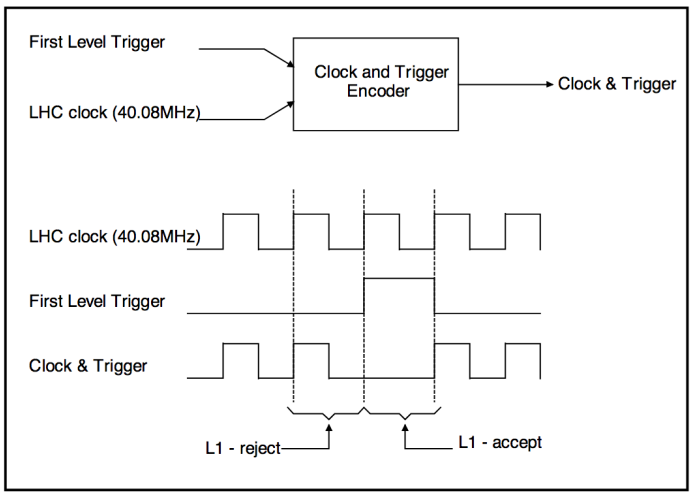


Figure 3.12: Combined coding of the LHC clock and first level trigger signals [40].

the DELAY25. The DELAY25 is a 5-channel programmable delay line that allows the digital signals to be delayed in phase, and a master channel that can be used to phase delay the clock signal. The phase of each channel can be independently changed with a resolution of 0.5 ns through an I^2C interface. The procedure corrects for the delay induced by the electronics and communication lines in the chain, so therefore allowing for the communication of the two devices.

After passing by the DELAY25 chip, the signals are sent to the Gatekeeper, which translates signals from the DOH/PLL LVDS standard to the TBM LCDS standard, and does the

reverse for the return signals. The signals coming from the gatekeeper pass by the LCDS driver chips, one per module, before being sent to the pixel modules. The Low Current Differential Signal (LCDS) driver chips feature an adjustable delay that allows the delay of each module to be tuned to take into account the length of the individual cables and the time-of-flight (TOF) delay. The TOF delay, as the name says, is the delay induced by the time necessary for a particle to cross different layers of the pixel detectors. An example of this effect is shown in figure 3.13. The TOF delay depends on the p_T and η of the particle and is of the order of hundreds of ps. The FED serves a different purpose from the two

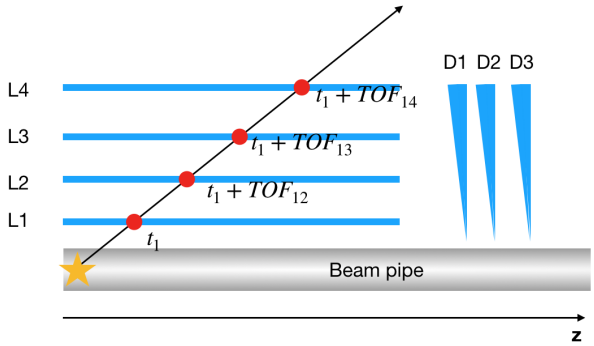


Figure 3.13: Sketch showing the time-of-flight delay of a charged particle traversing the pixel detector. The figure shows a quarter of the pixel detector, and the relative TOF delays between each layer of modules that need to be corrected for, according to the particle's pseudorapidity and p_T.

FECs; it is designed to receives the data from the modules at a 400 Mb/s data rate. The data sent from the modules as an electrical digital signal pass by a digital level translator (DLT) that adjusts the signal to serve as input for the digital laser drivers in the POH. The POHs convert the signal from electrical to optical, using the Transmitter Optical Sub-Assembly (TOSA) component provided by the Versatile Link project [41]. Then, the signals travel through optical fibres and to the FED in the underground counting cavern.
Finally the trFEC controls and program the devices on the supply tubes. The optical signals coming from the trFEC are converted to electrical signals by the digital opto-hybrid (DOH) and sent to the communication and control unit (CCU). The CCU implements the I^2C communication protocol used to control all the electronics device of the supply tubes. It also controls the DC-DC converters, which are radiation-tolerant, synchronous, step-down DC-DC converters mounted on a custom ASIC called the FEAST2 [42] via a peripheral interface adapter (PIA).
In the following sections the structure of a sector of the supply tubes will be described along with its electronic components of the readout and control chain that comprise it.

3.6 Supply tubes sector

Each of the four supply tubes is divided in eight sectors, which are independent units housing the electronics necessary for controlling and reading out the data stream from a group of modules. Figure 3.14 shows an assembled supply tube, in which the eight sectors are visible as slots the length of the tube. Each of the eight sectors is controlled by a CCU. The eight CCUs are connected in series as shown in figure 3.15, via a redundant system that allows a CCU to be bypassed in the event of failure. The power for the modules is supplied

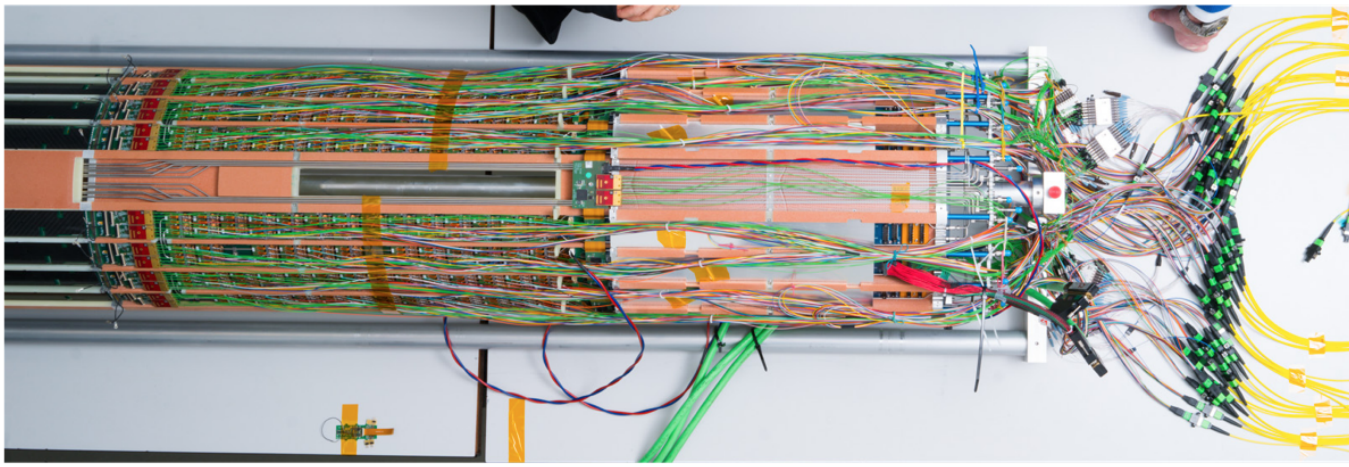


Figure 3.14: Top view of a fully-equipped supply tube. It is possible to see eight different sectors, the two red DOHs in the central sector, and part of the cooling loops.

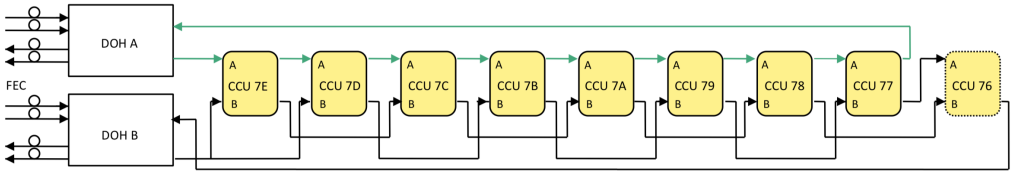


Figure 3.15: CCU ring redundancy system. In normal operation, each CCU uses input port A to transmit. In the case of a failure of one CCU, the previous one in the chain switches to port B for communication, bypassing the damaged CCU and closing the communication ring.

by 32 power supplies hosted inside the CMS cavern. The power supplies deliver 11.0 V to the DC-DC converters. Figure 3.16 shows a single DC-DC converter with and without the shield. Up to thirteen pairs of DC-DC converters are used in a single sector. Each DC-DC pair receives an input voltage of approximately 11.0 V and converts it to 3.5 (2.4)V for the digital (analog) circuits of the modules. The DC-DC converters are connected to the modules by a series of boards with the sequence : DC-DC pairs → extension board → adapter board → connector boards → twisted-pair cables → modules. Each pair of DC-DC converters can power up to four modules in parallel on layer three or four, two or three

modules in parallel on layer two, while for layer-one each module requires a single pair of DC-DC converters.
The modules belonging to layers 1 and 4 are in the same powering group, while the modules belonging to layers 2 and 3 are in a second group. In this way, each of the power groups consist of a similar numbers of modules. Each group is served by a single channel on the power supply, minimising the number of cables required to build the system.
Each sector is populated by up to fourteen POHs. Each POH has four laser drivers connected to optical fibres. The POHs are connected to the POH motherboard, which in turn is connected to the adapter board using flex cables. The hit rate of a module depends on its layer and z-position. Because of the different hit rates, a different number of fibres are used for the readout of modules belonging to different layers. One module on layer 1 uses all four of the fibre connections to the POH, while modules on layer 2 require two fibres, and modules on layer 3 and 4 use only one fibre to stream out the data. The pxDOHs are

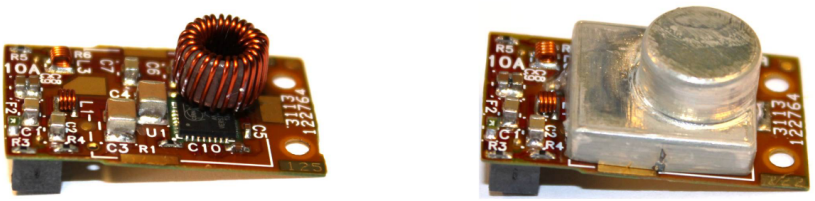


Figure 3.16: Photographs of a PIXV10 DC-DC converter, without (left) and with (right) the shield attached [35].

mounted on the DOH mother board that hosts the DELAY25 chips, the PLLs, the QPLLs and the Gatekeeper. One pxDOH is used for the modules belonging to layer 1 and 2, and another one serves the modules on layers 3 and 4.
Figure 3.17 shows the arrangement of the different boards used to route data and power through the various elements of the supply tubes. The purpose of these boards is to host the electronics devices and route the connections inside the sector. Using figure 3.17 as a reference, the boards used in one sector are:

- In orange is the high-voltage (HV) flex board, which carries the HV to the connector board.
- In blue is the DC-DC board that connects to the DC-DC extension boards.
- In teal the two DC-DCs extension boards.
- In dark green are the POH connector boards, which host and connect the POHs.
- In purple is the adapter board.
- In light green are the connector boards.
- In yellow is the CCU ring.

- In red is the DOH motherboard, which hosts the DELAY25, the PLL, the QPLL and the gatekeeper.

These boards are layered on top of each other in each sector.

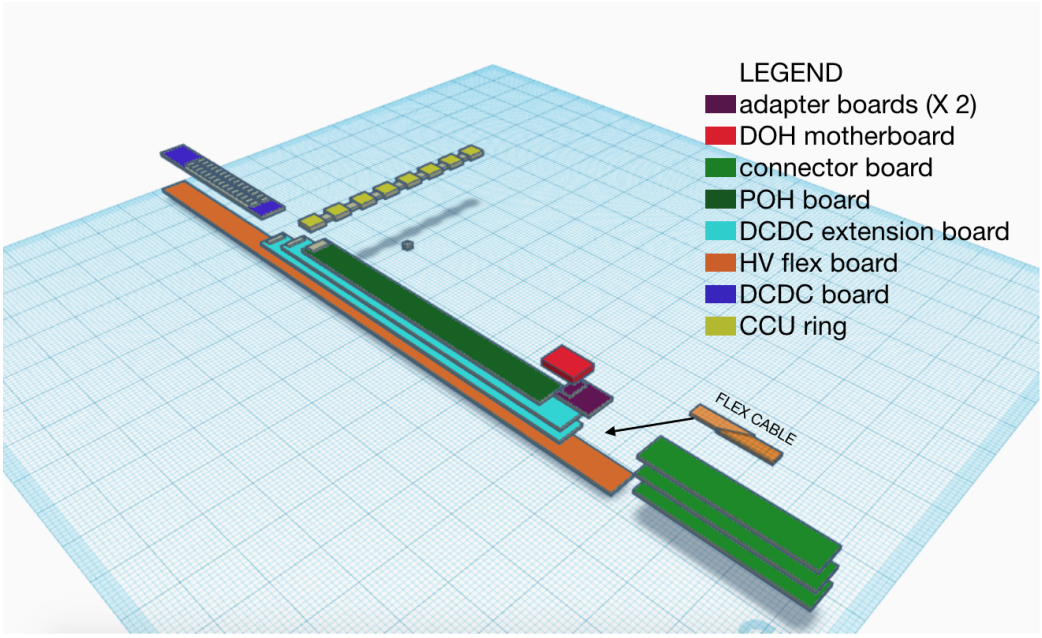


Figure 3.17: Different boards and flex cables used in the assembly of one section of the supply tubes. Not shown are the flex cables used to connect the connector boards to the other boards. The detector is located on the bottom right of the figure.

3.7 Phase 1 pixel detector test stand and testing procedures

A test stand was built to test the detector electronics components and develop the testing procedures. The test stand was situated in the University of Zurich Physics Department, and is shown in figure 3.18. The system replicates a slice of the Phase-1 barrel pixel detector, including the detector modules, the frontend electronics hosted on the supply tubes, and the CMS DAQ system. The test stand serves different purposes:

1. Develop testing procedures. The electronics components used during assembly are custom made, so it was necessary to develop testing procedures that can be used to test the components once installed without the possibility of directly accessing them.

2. Develop and test the initialisation sequence of the system. Single electronics components are linked together in a complex control and readout chain. It is necessary to initialise the components when the detector is turned on or reprogrammed, but also in case an electronics component gets stuck and it is necessary to reset part of the system. The initialisation procedures follow a precise order.

3. Discover the failure modes and explore the possible weak links in the readout and control chains. The supply tubes and the pixel detector are installed inside the core

of CMS, close to the beam-pipe. It is important to understand what are the possible failure modes and how to recover a stuck electronics component remotely.

4. Testing the electronics components before assembly. The electronics component are tested before delivery, but incorrect handling can damage them. Because of the large number of components and the amount of handling required to install them, it was important to test the components before and after the installation.

5. Develop calibration procedures. The different components on the supply tubes have to be calibrated to operate properly. For example, it is possible to tune the power of the laser-drivers of the POHs, or to change the delays on the DELAY25 chips. Furthermore, the readout chip of the pixel modules have a number of settings that have to be optimised in order to optimally collect data.

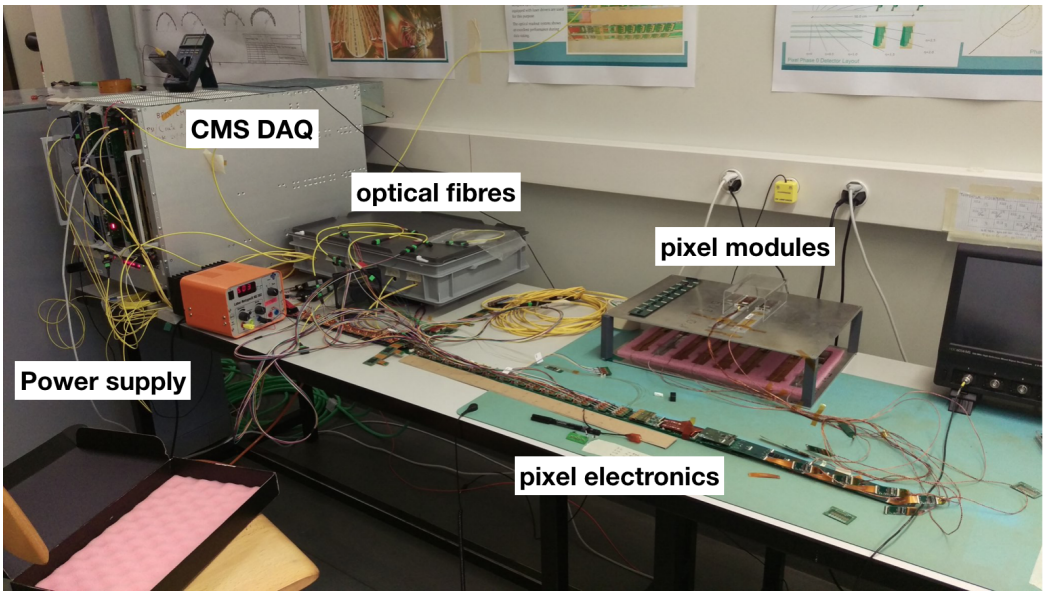


Figure 3.18: Test stand at UZH. The test stand is comprised of the pixel modules, the sector electronics, a slice of the CMS DAQ system, and the power supplies.

I developed or improved the above tests and procedures using a customizable interface software called bpixtools. The software libraries allow the user to write test programs in python, which interfaces to c++ classes, for communication with the FEC and FED application programming interface. The first test developed was the CCU redundancy test. As the name suggests, the test allows communication with the CCU ring, and verifies the communication with the eight CCUs in the chain, using the standard communication path with the input A ports, and then testing the backup redundancy part with the input B ports. The test checks the CCU functionalities for communicating with the electronic devices on the supply tube. The CCU ring schematic is shown in figure 3.15.
The next test sends a reset signal to all the devices on the supply tube. The reset signal is sent to the DOHs, the PLLs, the DELAY25s, and the POHs. For each device, the test writes to the programmable buffers of the devices a known value, reads the buffers, and

then resets the devices. After the reset, the test reads back all the buffers to check that the buffer's values are back to the default initial value. This test ensures that it is possible to program the devices and reset them.
Another test is to check that the DC-DC converters can be enabled and disabled. The DC-DC converters generate low-voltage from the power supply for channels that supply either layers 1 and 4, or 2 and 3. The DC-DC converters can be enabled or disabled by acting on a set of peripheral interface adapter registers on the FEAST2 control chip. The status of the DC-DC is then read back from the "power good" flag.
After the reset test, all the devices are initialised and the next testing phase starts. The next test in the chain checks the functionality of the POHs, the quality of the fibre connection, and the mapping of the fibre connections. The first test measures the intensity in the FED of the laser signals as a function of the laser bias. An example of the test results for different POHs is shown in figure 3.19. The plot shows the intensity of the laser signals as a function of the laser bias. The turn-on of the curves is different depending on if the POHs are sending logical signals of level zero or one. Dirty or bad connections generate power loss across the transmission lines and the laser intensity as function of the bias is lower. The slope of the curves is the figure of merit to evaluate if the fibre connections are clean or require inspection.

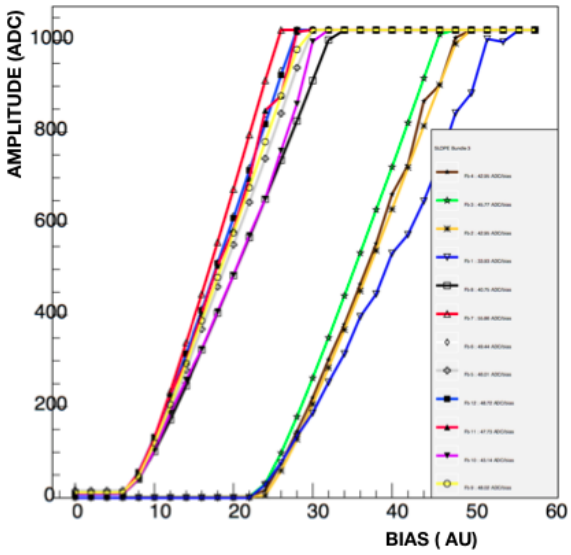


Figure 3.19: POH bias scan for 12 fibers coming from 3 POHs. The turn-on of the curves is different depending on if the POHs are sending logical signals of level zero or one. The slope of the curve defines if the connection of the fibres is clean. In this case the blue curve, has a low slope value, and cleaning and inspection are necessary.

If the slope of the curve is smaller than a certain threshold, the test prints a warning and the user can clean the fibre connections and repeat the test. This test is performed simultaneously for all the fibres to optimise it time-wise.
The test is then repeated powering-up one laser at a time, and reading and comparing just the corresponding FED address, which is mapped in a txt file for each sector. If the POH

lasers work, the fibers are clean, and the mapping of the fibres from the POHs to the FED channels is correct, then the test will be successful.
The testing procedures described up to now do not require the use of detector modules. The next test is the "RDA-SDA scan", to check the communication with the module and the programming of the RDA and SDA lines delays on the DELAY25 chip. As explained before, the LHC 40 MHz clock is used both by the FECs, the FEDs, and the TBM on the detector modules. The DELAY25 chip uses four delay lines, one for the clock, one for the trigger, one for the SDA line, and one for the RDA line. The test checks all the possible delay settings for the SDA and RDA lines by sending to the TBM a query and reading back the answer. When the RDA and SDA delays are properly set, the FEC command is received by the TBM, and the FEC is able to decode the package sent back by the TBM. There are two DELAY25 chips for each sector, one for all the modules on layers 1 and 2, and one for all modules in layers 3 and 4. With this test, it is possible to find out the right delay settings that allow programming of and communication with the modules. The delays can be adjusted in steps of 0.5 ns from 0 up to 32 ns. Figure 3.20 shows the result of an RDA-SDA scan for five modules and the combined result. The x and y axes are respectively the RDA and SDA delay in steps of 0.5 ns, the yellow regions are the regions in which the system is able to communicate with the sensors. The last plot shows the combined regions, and the narrow yellow lines are the regions in which the system can communicate with all five modules. Each plot shows four regions which have a periodicity of 25 ns (or 50 delay units) both in x and y, consistent with the period of the LHC clock.

The two final tests check the readout of data from the modules. The first tests the PLL delays inside the TBM chip. The schematic of the delays is shown in figure 3.21 [43]. Using the TBM delays it is possible to delay the TBM header and trailer by 6.25 μs. The ROC tokens can also be delayed by 6.25 μs. Additionally the data streams from the ROCs can be delayed up to 7 ns, in increments of 1 ns. The phase between the 40 MHz clock and the 160 MHz clock can be adjusted using two adjustable PLL delays with a range of 7 ns in increments of 1 ns.
The scope of this test is to scan the delay configurations and find the optimal one. If the delays are not set properly, non-proper synchronisation between the TBM header, ROC header and TBM trailer can lead to invalid bit patterns and thus decoding issues. The optimal setting of the delays allows for a seamless readout of all the data. The test is performed by scanning the adjustable delays for the two data streams of the two ROC ports, port 0 and port 1. Subsequently, if there are problems reading the TBM header, trailer, or first or last ROC headers, the delay between the token and the TBM header and trailer are adjusted.
After this test, all the electronic elements of the readout chain have been tested, and the modules delays have been optimised. The modules can be now used as drivers to test the full readout chain. The final test checks if the full chain works from DAQ to detector module. A trigger is sent to the modules, and the data stream is read back. The test checks that all the modules send back data, and that the data contains the correct sequence of information, including the TBM header, the ROC headers, and the TBM trailer. This sequence of tests checks the functionality of all the elements of the readout chain, including the powering,

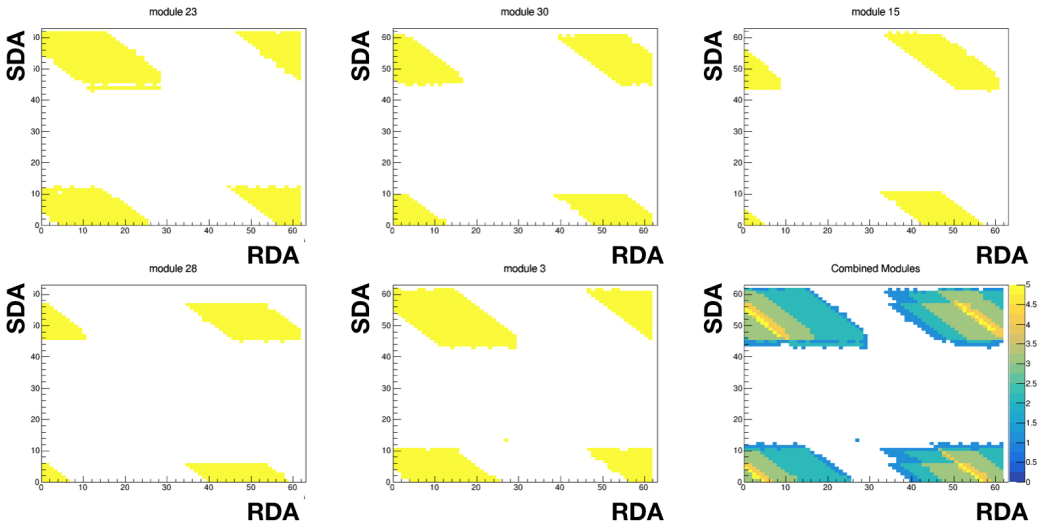


Figure 3.20: The RDA-SDA delay scan for 5 modules and the combined result for all 5 modules. The x and y axes are respectively the RDA and SDA delay in steps of 0.5 ns, and the yellow regions are the ones in which the system is able to communicate with the sensors. The last plot shows the combined regions and the narrow yellow lines are the regions in which the system can communicate with all five modules. Each plot shows four regions which have a periodicity of 25 ns (or 50 delay units) both in x and y, consistent with the period of the LHC clock.

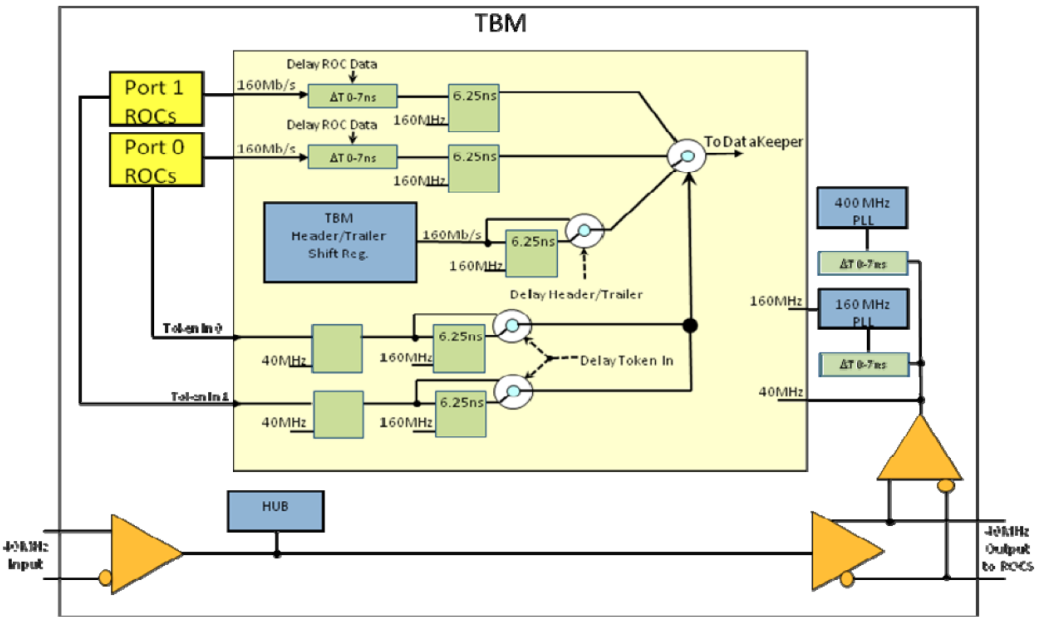


Figure 3.21: Block diagram of readout delay adjustments [43].

the quality of the optical, electrical and electronics connection, and the ability to control the elements of the readout chains, and finally, the programming of the modules and the

readout of the data stream. The high voltage distribution is tested using a multimeter investigating for possible shorts or broken connections. To test the low voltage distribution, a custom load board was used, shown in picture 3.22 [44]. The load board features high power resistors that replicate the loads of the modules under normal operations. Using the board it was possible to test the DC-DC converters, the electrical connections of the low voltage distribution, and the voltage drop along the lines. The test procedures have been

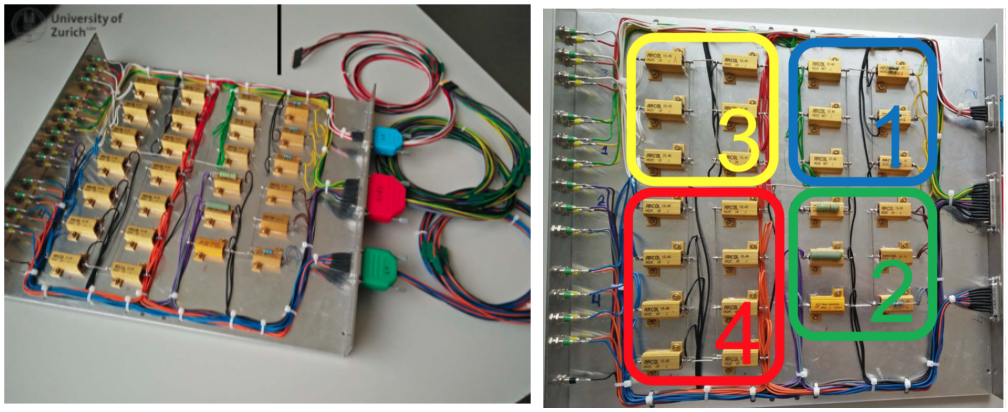


Figure 3.22: Load board used to test the low power distribution for the supply tubes sectors. The high power resistors (in bronze) have been chosen to reproduce the load of the modules under data taking [44].

developed using the UZH test bench, and then used for commissioning the four supply tubes after the electronics components and cables were connected.

3.8 Integration at UZH

The integration of the four BPIX supply tubes consisted of the assembly and testing of the components for power, cooling, and readout. The mechanical structures shown in figure 3.9 were placed on a rotatable axle on a bench to facilitate the assembly. In the first step, the HV flex cables were positioned on the mechanical structure. Then the DC-DC bus board, the extension boards, the POH motherboards, and the adapter boards were screwed in place. Then the cooling tubes were placed and fixed with cooling bridges, which hold the supply tubes in place and serve as thermal connections between the cooling pipes and the DC-DC converters. The DC-DC converters were then installed on top of the cooling bridges. In the next step, the CCU ring was connected before the optical components (POHs, the DOH motherboard and the DOHs) and flex cables were attached. The POHs also used cooling bridges to improve the thermal connection to the cooling pipes. Three POHs are connected to a single bundle of 12 optical fibres before mounting them onto the supply tube. The 1984 optical fibres that run along the service half-cylinders were carefully arranged within to fit into the tight space, together with the DOH fibres. Pictures of the BPIX service half-cylinders during the assembly are shown in 3.24. All electronics parts of the BPIX readout and power system were tested for functionality before, during, and after mounting. With the components of the pixel detector DAQ and powering system available for the testing,

one sector could be tested at a time. In order to guarantee that the detector modules would not be damaged when connecting to power, the power distribution was tested by measuring the low voltage at the module connectors by applying a constant load corresponding to the expected maximum power consumption of the system during operation. The load board used is shown in figure 3.22. Different areas of the board host sets of resistors that replicate the DC-DC load for different layers. Modules on layer 1 are powered by one DC-DC converter pair each, while on layer 2 each DC-DC converter pair can power up to a group of 2 or 3 modules. Modules on layers 3 and 4 are grouped into four, and each group is powered by a single DC-DC converter pair. This translates in different loads for the different DC-DC converters pairs depending on which layers they power. The load board uses high power resistors and takes into account the resistance of the copper cables used in its construction. The final value of the resistance for the different layers is shown in table 3.23, together with the relevant information like the number of modules in each group and number of module groups for each layer. The board is connected to the DC-DC extension board using using modified connector boards

Analog Signal										
Pixel Modules								Load Board		
Layer	# of modules per group	# of groups	module current (A)	Cable resistance (Ω)		module voltage	Module + cable resistance (Ω)	cable resistance (Ω)		module resistance (Ω)
				active	ground			active	return	
1	1	3	0.38	0.30	0.08	1.60	4.59	0.035	0.035	4.52
2	3	2	0.38	0.34	0.12	1.59	2.32	0.028	0.028	2.27
2	2	1	0.38	0.34	0.12	1.59	1.55	0.032	0.032	1.49
3	4	3	0.38	0.34	0.12	1.60	1.17	0.042	0.042	1.08
4	4	4	0.38	0.34	0.12	1.61	1.17	0.042	0.042	1.09

Figure 3.23: Load board resistance value for different layers and module groups [44].

The high-voltage distribution was checked up to voltages of 1000 V, which are necessary to maintaining signal integrity after irradiation. In addition, the supply tubes have been tested cold at -20 degrees ˚C. The STs were enclosed in an air-tight container and flushed with nitrogen gas to avoid condensation. A portable cooling station was used to achieve the required temperature. The portable cooling station denominated "CO_2lean" was developed by different institutes, including RWTH Aachen and ETH [45]. During the cold test, one POH was found faulty and replaced, all the other components performed as expected, and the supply tubes were packaged and shipped to PSI for integration with the barrel pixel detector modules.

3.9 Integration and testing at PSI

While the supply tubes were being built at UZH, the pixel detector was being assembled and tested at PSI. The pixel detector mechanical structure is composed of four cylindrical carbon structures, one per layer, each divided in two half shells. The eight carbon support structures

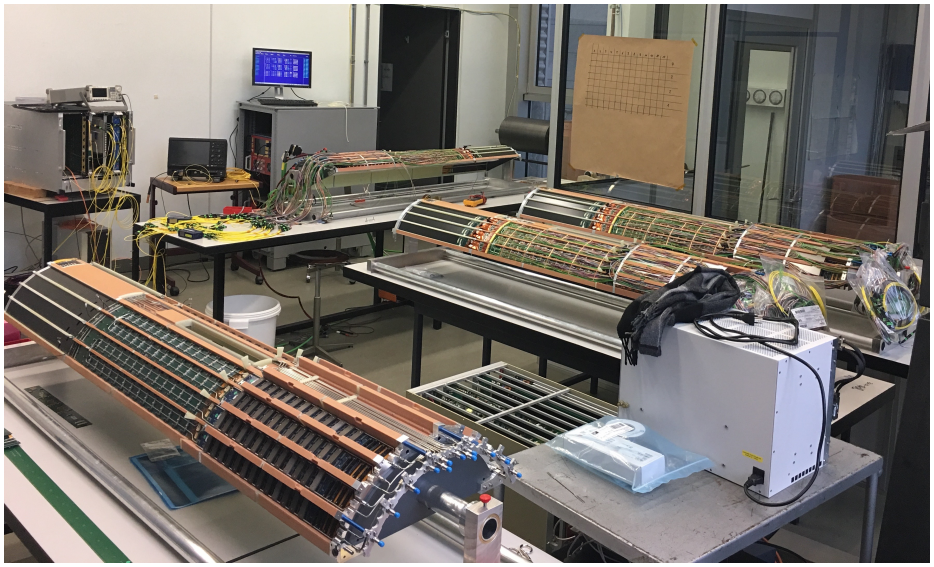


Figure 3.24: The four supply tubes in the UZH assembly room. The two on the right are already assembled and tested. The one on the bottom left corner is partially assembled and the cooling pipes and manyfold are visible. The supply tube in the back is fully assembled and under test, the yellow optical fibres used to connect the supply tube to the DAQ system are visible on the table.

host the cooling pipes, on top of which the modules were connected using carbon flanges and screws [34]. In successive assembly steps, the twisted-pair cables were connected to the modules and the populated half shells were connected together to form the two detector halves, which are mechanically independent. Figure 3.25 shows one detector half [34]. In the picture, it is possible to see the four concentric carbon layers populated by modules, the twisted-pair cables, and the cooling loops.

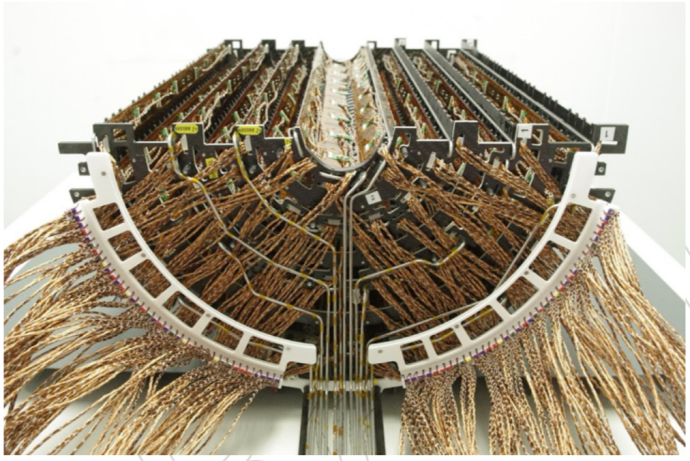


Figure 3.25: One half of the BPIX detector. It is possible to see the four concentric carbon layers populated by modules, the twisted-pair cables, and the cooling loops [34].

The twisted-pair cables were connected, on the other end, to the connector boards.

The mounting of the modules was completed with only four non-working modules. These modules were, at this integration stage, not accessible. The number of dead channels was less than 0.7%[34]. The STs and the detector were connected at PSI inside transport boxes with a built-in rail-mechanism; the system was also used to facilitate the final installation of the detector inside CMS. During the merging of the STs and the detector, the connector boards were connected on the supply tubes, and the cooling loops on the detector side were connected on the STs cooling lines, closing the cooling circuits. The same portable cooling station used for the cold tests at UZH was used at PSI for testing the full detector system at a room temperature of 15°C. The setup at PSI allowed for the testing of one quarter of the detector at a time. The final check at PSI was performed by using multiple test procedures, including the "pixelAlive" test routine of the full detector. In this test procedure, each pixel cell is injected with a fixed number of calibration pulses, and the output is an efficiency map of the modules.

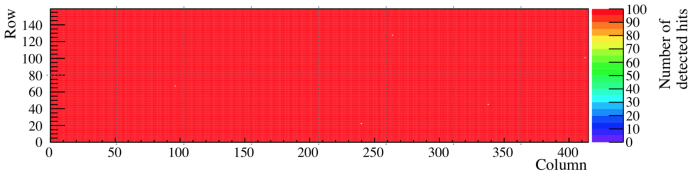


Figure 3.26: PixelAlive efficiency map for a full silicon module. The picture shows a schematic view of the pixel module, the 16 readout chips are outlined in dots. The number of registered pulses is color-coded. here, only five defective pixels show no hits. [46]

The results from a pixelAlive test are shown in figure 3.26[46]. The test confirms the functionality of all the elements on the detector and the results are used to construct an efficiency map of the detector. The results are also used after the final installation to confirm that all the modules are still functional. After all the tests were performed, the two halves of the detector were moved to the CMS experiment at CERN for the final installation.

3.10 Final installation and calibration

On February 7, 2017, the two halves of the BPIX detector system were transported to CERN. After the transport, the system was tested in a cleanroom at the surface of the CMS experimental area [34].
Detailed tests and calibrations were carried out at room temperature, and the detector was operated inside cold boxes at -20°C . A portion of the twisted-pair cables was found faulty, due to friction damage. The careful insertion of isolation and buffer materials solved the problem.
The two halves of the detector were then transported in the cavern and inserted simultaneously inside CMS using the rail system. Prior to this event, the endcaps of the CMS detector was moved and an access platform was installed. The same platform was used to extract the original pixel detector. The necessary preparation for the services was already in place. After installation and the closing of the detector volume to avoid condensation

and ice formation, the detector was cooled to the nominal operational temperature of -22°C and a number of calibration and functionality tests were performed [34]. The sequence of tests described in the previous sections were performed, including measuring the digital and analog current drawn by the modules, checking the functionality of the POHs, and the delay settings of the single modules and of the Delay25 chips with the TBMPLL and TBMDelay25 tests. Calibration routines were executed to optimise the module settings, to trim the noise and threshold of the pixels, and to mask the noisy pixels. The CMS Phase-1 pixel detector was exposed to the first collisions in May 2017.

3.11 DC-DC converters failure

On October 2017, a DC-DC converter broke inside the pixel detector. This was an unexpected event, because from previous testing the DC-DC converters did not show any obvious failure mechanism. In the following two months of operations, about 5% of the 1200 DC-DC converters failed. While dedicated teams of people started investigating the problem, the only practical solution at our disposal was to replace the broken DC-DC converters. It was suspected, and later confirmed, that operating the converters at a higher voltage, 11 V instead of 10 V, would mitigate the leakage current inside the FEAST chip and contain the DC-DC failure problem. Therefore, the operating voltage of the DC-DC converters was quickly changed from 10 V to 11 V. It was decided that all DC-DC converters would be removed and replaced in December 2018. The detector was extracted and the DC-DC converters of the BPIX and FPIX were all replaced with new (but identical) ones. The extracted DC-DC converters underwent a number of electrical tests to find out the origin of the problem. The mechanism that led to those failures was finally identified in May 2018 by a dedicated team of people. The problem was a leakage-drain current mechanism inside the low-voltage protection circuit of the FEAST2 chip [47]. The low-voltage protection circuit is used during the powering-up phase of the DC-DC converter. When the voltage is low the DC-DC converter drains more current to reach the expected power level and this can be harmful. The protection circuit kicks in in the startup phase, moderating the amount of current drained by the DC-DC converter during the start up phase. To reproduce the problem, the DC-DC converters were irradiated, under load, and power-cycled for a period of 6 months, roughly equivalent to the time necessary to observe the first DC-DC converter failure in the pixel detector.

The problem was aggravated by the repeated power cycles, necessary to restart stuck module. Because of a design flaw, when a TBM was stuck due to a single-event upset (SEU) from the deposition of ionizing energy in a transistor, the only way to recover the stuck modules was to power-cycle the DC-DC converters of the corresponding power group.

As a consequence of the DC-DC problem, some modules were damaged by a radiation-induced charge accumulation inside the pixel unit cell (PUC) at the ROC level. As discussed in the previous sections, the LV and HV group granularity is quite different, and during the DC-DC problem some modules were still connected to the HV while the LV was off. As a consequence, the charge created by minimum ionizing particles inside the silicon sensors accumulated at the PUC pre-amplifiers and damaged it. This effect was proportional to

the pixel area, and pixels at the corners of the ROC suffered the most, especially the ones situated on layer 1 and 2, closer to the beam pipe. The damaged modules that were directly accessible (6 out of 8) were replaced [34]. Overall, these problems only affected a small part of the collected datasets due to the rapid response of CMS collaborators in troubleshooting and mitigating the affects. Two improved version of the DC-DC converter have been developed, and all the DC-DC converters have been replaced during LS2.

3.12 Conclusion

The construction of the supply tubes, and their testing, commissioning, and merging was a success. The schedule was tight, but the detector performed as expected, with the exclusion of the DC-DC crisis period. Multiple analyses benefited from the improved vertex resolution, and in a more general sense, all the analyses were able to benefit from the additional tracking layers and improved readout speed to improve efficiency with respect to fake rate, as well as improved track resolution parameters. The spatial resolution performance, extracted using the triplets method [48], is shown in figure 3.27 [34]. The cluster hit efficiency is shown in figure 3.28 [34]. For comparison, the efficiency of the original detector, at the same instantaneous luminosity, would have been below 95%, right at the bottom of the plot. Since the first insertion, the detector has been extracted and inserted multiple times for ordinary and extraordinary maintenance, and the installation and commissioning procedure has proven to be reliable, while the detector system has proven its solid performance during these operational years.

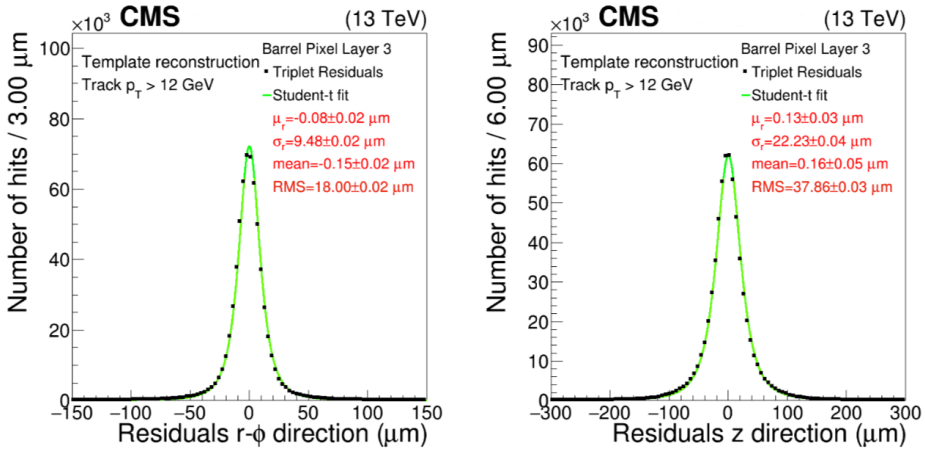


Figure 3.27: Pixel hit residuals of the Phase-1 barrel pixel detector for BPIX L3 in the r-Φ (left) and the z direction (right)[34].

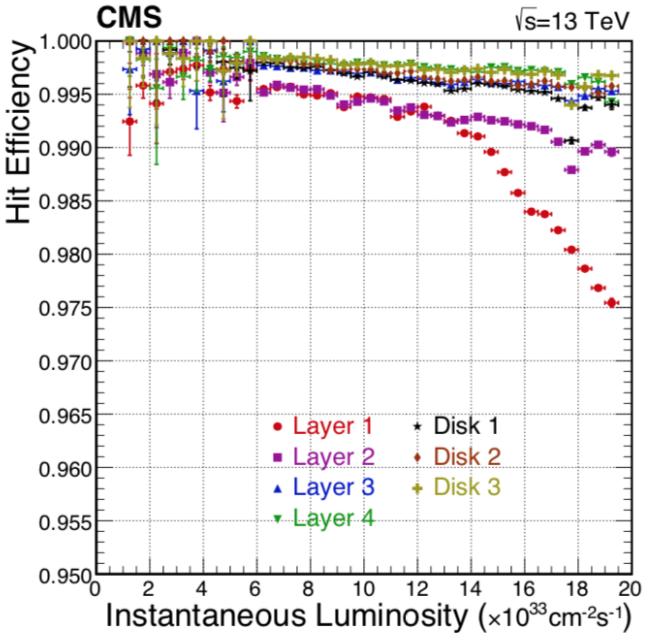


Figure 3.28: Cluster hit efficiency of the Phase-1 barrel pixel detector as a function of instantaneous luminosity [34].

Chapter 4

Simulation of single-layer position resolution for the CMS Phase-2 pixel detector

4.1 Introduction

In view of the high luminosity LHC upgrade, the current CMS pixel detector will be replaced by the CMS pixel phase-2 detector. The goal of this upgrade is to maintain the physics performance of the present detector during the high luminosity era, when there will be 200 collisions per bunch crossing, which is 5 times the present so-called "pile-up". The new detector is designed to efficiently operate at higher pile-up, and extending the η coverage from 2.4 to 4.0. The new detector will feature improved technologies for the pixel modules, a more optimized layout, and a faster readout. A simulation software called tkLayout has been used to estimate the performance of different detector layouts. The tkLayout parametric simulation relies on basic physics principles, a detector layout configuration, and the resolution of the detector's sensors. In the following sections, I will described how I simulated and predicted the sensor's resolution performance years before test beam results were available, the implementation inside the tkLayout software, and the relevance of my work.

4.2 CMS phase 2 pixel detector

The CMS phase 2 pixel detector is the next upgrade of the inner tracker, and will be installed around 2025 before the beginning of the HL-LHC era. The new detector will be upgraded in almost every aspect with respect to the current one. The new pixel size, 100×25 μm^2 (or 50×50 μm^2) is 6 times smaller in area compared to the current pixel size of 150×100 μm^2. The sensor will also be thinner, 150 μm vs 275 μm. The detector layout is shown in figure 4.1. For comparison the present pixel phase-1 detector features 4 layers in the barrel region and 3 disks in the forward region, with a coverage of $|\eta| < 2.4$, while the phase-2 detector will number 4 layers in the barrel region, 8 disks per side in the forward regions, and 4 disks per side in the end-cap region, and extend to $|\eta| < 4.0$. All the different barrels and disks

will use the same readout chip technology. Planar sensors are the default sensor choice, but ongoing studies are testing the possibility of using 3D sensors in the innermost layers and rings to improve the radiation hardness of the system. In the following sections I will describe how I simulated the single-layer resolution of this detector and how the simulation has been used to study the performance of different layouts.

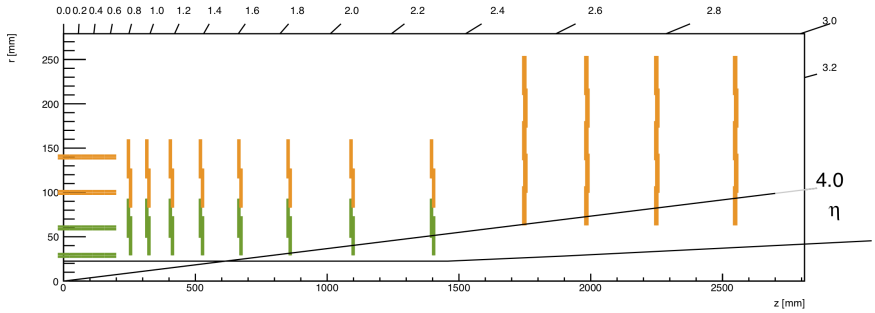


Figure 4.1: A quarter of the CMS pixel phase-2 detector layout. The system feature 4 layers in the barrel region, 8 layers per side in the forward region, and 4 additional disks per side in the end-cap region.

4.3 Resolution study

The goal of this simulation is to study the performance of a single layer of the CMS phase-2 pixel detector. The performance of the sensors depends on many different parameters. The most important ones are as follows.

- Pixel size: two different dimensions with the same area are being evaluated, 50 $\times$ 50 μm^2 and 100 $\times$ 25 μm^2. The thickness of the sensor in both cases is 150 μm. The ratio between pitch and thickness plays an important role in the resolution as a function of the incident angle, as will be explained in this section.

- Charge collection efficiency: the charge collection efficiency of a sensor depends on numerous parameters, including the geometry, the irradiation, and the bias voltage. In general, the charge collection efficiency is quite flat as a function of the depth in un-irradiated sensors, while radiation damage changes the shape of the collection efficiency function. This effect can be accounted for and corrected offline using different methods that will not be discussed here. For this simulation only un-irradiated detectors have been considered and a flat charge collection efficiency of 80 % has been used, which is a typical value.

- Digitizer characteristics: usually complex simulations are employed to simulate how the signal is collected and shaped by the readout electronics. These simulations emulate the circuit time and charge response to take into account effects like time walk (described in section 2.12) or inefficiencies of the system. The purpose of the simulation described in this chapter is to estimate the spatial resolution of the sensor and

not its efficiency, therefore the only parameters taken into account are the digitizer threshold, which is tunable and the digitizer resolution, which is 4 bits. Since the hit position is reconstructed using the centre-of-charge method, the spatial resolution is strictly correlated to the charge resolution.

- Operating temperature and bias voltage: the drift velocity inside a silicon sensor depends on the sensor doping, irradiation, geometry and finally on the operating temperature and bias voltage. While the doping is kept fixed and the irradiation effect is not taken into account, different operating voltages are investigated. The operating temperature is fixed at -30°C.

- Magnetic field. The value and orientation of the magnetic field play an important role in the charge drift inside the sensor and its final spatial resolution. In the pixel detector region, the magnetic field is constant, parallel to the beam axis and equal to 3.8 T. The Lorentz force affects the charge carrier's movement, and this effect is taken into account. Two different cases were evaluated, corresponding to the two sections of the detector. In the barrel region, the electric field inside the sensor is perpendicular to the magnetic field, while in the forward region of the detector the electric field inside the sensor is parallel to the magnetic field.

The work-flow of the simulation is structured as follows. The sensitive part of the silicon sensor is simulated in Geant4 as a slab of silicon with a thickness of 150 μm. The size of the silicon slab is several times larger than the size of a single pixel. Inside Geant4, it is possible to define a particle gun that generates muons with a fixed momentum of 10 GeV. The points from which they originate, and the flight direction, are also defined. In this particular case, the muon direction was always the same, while the origin point was smeared using a uniform distribution, to emulate a beam spot. This reflects what happens in LHC conditions, and is necessary to avoid always hitting the same pixel at the same point. After each hit, the entry and exit point of the muon inside the silicon slab are registered, together with the number of electron-hole pairs created inside the sensitive region. This procedure is repeated for a large number of events and for different silicon slab orientations. The trajectory of the particle inside the silicon slab and the number of electrons created are inputs in the following steps. Geant4 does not include a treatment of the electron-hole pairs drift physics so I developed a separate python module to include this. This simulation, starting from basic principles, describes the electron-hole pair drift mechanism. These steps are fundamental to take into account the effect of the combined magnetic and electric field on the charge carriers. The final step is the digitization, in which the deposited charge is read out by the electronics and digitized. This step gives as output the reconstructed hit, taking into account the threshold and the resolution of the chip. The position of the reconstructed hit is then compared to the true hit position, and from the distribution of the residuals the module resolution is extracted. The residual for the ith event is defined as:

$$r_i = x_{reco} - x_{truth} \tag{4.1}$$

where x_{truth} is the true hit point obtained from the Geant4 simulation and x_{reco} is the reconstructed hit point using the charge centroid method.

The pixel resolution depends on the impact angle of the particle. Each step will be described in the following paragraphs.

4.4 Geant4

Geant4 is a toolkit for simulating the passage of particles through matter. It includes a complete range of functionality including tracking, geometry, physics processes and hits. The physics processes offered cover a comprehensive set, including electromagnetic, hadronic and optical processes, a large set of long-lived particles, detector materials and elements, all over a wide energy range, in some cases extending to the TeV energy range. It has been designed and constructed to allow the physics processes utilised to be interchanged, to handle complex geometries, and to enable its easy adaptation for optimal use in different sets of applications. The toolkit is the result of a worldwide collaboration of physicists and software engineers. Its software exploits object-oriented technology and is implemented in the C++ programming language. It has been used in applications in particle physics, nuclear physics, accelerator design, space engineering, and medical physics. The key domains of the simulation of the passage of particles through matter are:

- Geometry and materials. These classes allow for the definition of the different components of the detectors and surrounding materials. The different volumes of material fall under two categories, sensitive and inactive. The active volumes are part of the detector sensor, while all the services, ambient fluids and gases, and the support structure fall under the inactive material category.

- Particle interaction in matter. Depending on the requirements, different physics processes can be enhanced or suppressed. Because of that, different libraries have been developed, from general cases to ones specifically targeting low or high energy interactions.

- Tracking management. Geant4 simulations operate in steps. At each step, the simulations update the particle's position and interactions. The tracking manager keeps track of the particle position.

- Digitisation and hit management. The hit management use the position and trajectory of the tracks and the geometry to evaluate if a particle is inside an active volume. In this case, the detector response can be simulated and, if necessary, digitised.

- Event and track management. An event is a collection of tracks and the interactions that occurred.

- Visualisation and visualisation framework. The visualisation framework is used to create a GUI and event display, which are useful both for debugging and geometry optimisation.

- User interface.

The toolkit offers the user the ability to create a geometrical model with a large number of components of different shapes and materials, and to define "sensitive or active" elements that record information (hits) needed to simulate detector responses (digitisation). The primary particle of the events can be derived from internal and external sources. Geant4 provides a comprehensive set of physics processes to model the behaviour of particles. The user is able to choose from different approaches and implementations, and to modify or add to the set provided. In addition, the user can interact with the toolkit through a choice of (graphical) user interfaces and visualise the geometry and tracks with a variety of graphics systems through a well-defined as well as implement it over other favoured systems. In general, the classes in the toolkit are designed in a highly reusable and compact way so that the user can extend or modify their services for his/her specific applications. The user can realise this by following the discipline of object-oriented technology.

The design has evolved during development. It currently includes 17 major categories. These are grouped into categories that describe the tracking of particles and the physical processes they undergo. The track category contains classes for tracks and steps, used by processes which contains implementations of models of physical interactions. Additionally, one such process, transportation, handles the transport of particles in the geometry model and, optionally, allows the triggering of parameterisations of processes. All these processes may be invoked by the tracking category, which manages their contribution to the evolution of a track's state, and is used to provide information in sensitive volumes for hits and digitisation. Over these, the event category manages events in terms of their tracks and the run category manages collections of events that share a common beam and detector implementation. A readout category allows the handling of pile-up. Finally capabilities that use all of the above and connect to facilities outside the toolkit (through abstract interfaces) are provided by the visualisation, persistency, and user interface categories.

4.5 Pixel sensor simulation

I developed a Geant4 implementation to simulate the interactions of muons with the active portion of a phase-2 silicon pixel detector. The simulation includes the following.

- The silicon sensor: the silicon sensor is simulated as a uniform slab of silicon with a thickness of 150 μm. The slab is not segmented to optimize the computational time.

- The particle gun: in this simulation the particle gun creates a uniform beam of muons with a momentum of 10 GeV/c. The beam is collimated and the beam spot size is several times larger than a single pixel.

- The physics library: the physics library used is optimized to simulate the interaction between the muons and the silicon. The main interactions are the creation of electron-hole pairs and delta-ray creation. For this purpose, the standard emstandard_opt4 library has been used. This physics constructor was designed for any applications requiring higher accuracy for electrons, hadrons and ion tracking. It is based on the most accurate standard and low-energy models.

- The stepping action manager: Geant4 allows the user to define the maximum length of a step, which is the maximum distance a particle can travel inside the simulation before the program stops and check its position and interaction. Poorly tuning the step length can lead to a poor simulation or a computationally heavy one. To choose the right step length, the guideline is to refer to the interaction length of the particles inside the detector material.

- The tracking event and run manager: the tracking manager uses the step length and hit manager information to update the tracks, while the event manager collect all the tracks inside an event. Similarly, the run manager collects all the events inside a run.

An event display of the simulation is shown in Figure 4.2. The Geant4 part of the simulation determines for each track:

- the entry point of the muon inside the silicon sensor,

- the exit point of the muon inside the silicon sensor, and

- the number of -hole pairs created inside the sensor.

These data are saved in a ROOT Tree format and are used in the following steps. As a sanity check, the distributions of the number of electron-hole pairs created inside the sensor for different lengths of traversed path was tested against theoretical predictions and published values [49]. The values are distributed as a Landau, and the peak value matches the expected value of 80 electron-hole pairs per micron.

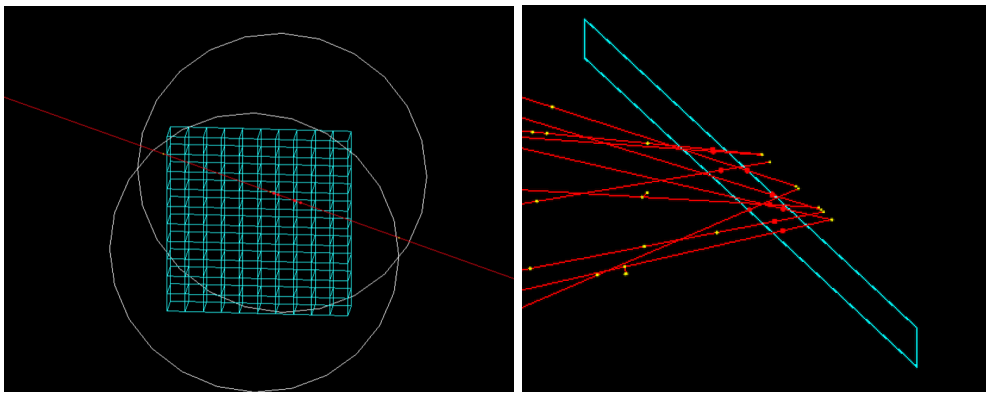


Figure 4.2: Event display examples. On the left: the track of the muon is shown in red, while in blue is the pixel sensor layout. On the right: different tracks in the same event. The red and yellow dots symbolise the ionization and bremsstrahlung interactions.

4.6 Charge-carrier propagation

Geant4 does not handle charge-carrier propagation. To simulate the effect of the electric and magnetic field in the silicon sensor, a separate simulation has been developed. This simulation relies on basic physics principles, on the silicon sensor characteristics and on

the information obtained from the Geant4 simulation. The multiple scattering of the muon passing through the silicon sensor is not taken into account, and so the trajectory of the muon inside the sensor is a straight line between the entry and exit point, as described by the Geant4 simulation. It is assumed that the electron-hole pairs are uniformly created along this path, and that the sensor is fully depleted. The pair's components then separate and start to drift in opposite directions under the effect of the electric field, and the magnetic field, which is to be discussed later. The drift of electrons (holes) inside the silicon sensor can be separated in two components, one parallel to the electric field and one in the plane perpendicular to it. The first one is described by the following equation:

$$\vec{v} = \frac{d\vec{r}}{dt} = \mu\vec{E}, \tag{4.2}$$

where μ_e (μ_h), is the electron (hole) mobility. The second component is described in term of the diffusion coefficient, and is related to the multiple scattering between the silicon reticule and the charge carriers. The drift of electrons (holes) inside the silicon sensor, in the plane perpendicular to the electric field, is described by the equation:

$$< x(t) >= \sqrt{D}t, \tag{4.3}$$

where $< x(t) >$ is the mean free path of the electron in the plane. D is the diffusion coefficient defined as:

$$D = \frac{\mu_e k_B T}{q}, \tag{4.4}$$

where μ_e is the electron mobility, k_B the Boltzmann coefficient, T the temperature in Kelvin, and q the charge in Couloumbs. The combined effect of the drift and the diffusion is shown in Figure 4.3.

These effects are simulated in the following sequence:

- The track of the muon inside the sensor is defined using the information from the Geant4 simulation.
- The total number of electrons generated by the particle is weighted by the charge collection efficiency of the sensor, assumed to be in this work 80% with a flat distribution across the thickness of the sensor.
- The weighted number of electrons generated by the particle is divided in a finite number of steps, equally spaced, and distributed along the track.
- The electrons belonging to different steps will travel a different distance, and the trajectory is defined by the combination of equations 4.3 and 4.2.
- The effect of the diffusion is emulated by smearing the charge distribution of each step using a Gaussian generator. Each packet of electrons in each step is smeared around its coordinate with a 2D normal distribution with a variance of $\sigma^2 = D \times t$, where t is the travel time estimated starting from equation 4.2, and D is the distance of the packet from the readout plane.

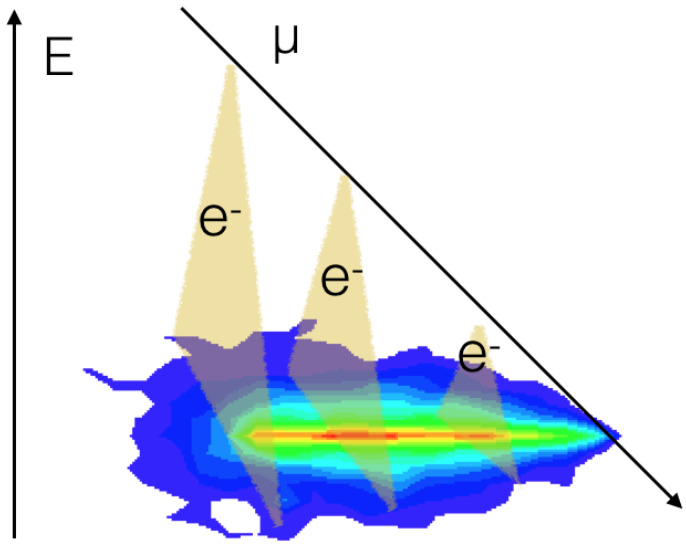


Figure 4.3: Electron drift and diffusion inside the sensor. The electron-hole pairs are uniformly randomly generated along the track of the charged particle. Depending on the point at which they are generated, they travel a certain distance inside the sensor. In yellow is the electron path, which is a combination of the effects of the lateral diffusion and the drift in the electric field. On the bottom is the charge distribution on the top of the sensor, where the charge is read out. The holes drift in the opposite direction and they are not considered in this sketch.

- The charge distribution on the readout plane is saved and passed on to the next step of the simulation, the digitizer.

The number of steps in which the total number of electrons is split is critical for the simulation. Too many steps heavily slow down the simulation, while too few steps do not reproduce accurately the physics processes taking place inside the sensor. The minimum number of steps used in this simulation is 100.

4.7 Magnetic field effect

Inside CMS a powerful magnet generates a 3.8 T solenoidal field. The effect of the magnetic field on the 10-GeV muon tracks is negligible inside the sensor, where the length of the track is of the order of 100- 200 μm and the deflection of the order of a nanometer. This is not true for the charge carriers, which have a much lower energy, of the order of the KeV. The drift of electrons (holes) inside silicon, under magnetic and electric fields, is described by:

$$\vec{v} = \frac{d\vec{r}}{dt} = \frac{\mu(\vec{E} - \mu r_H \vec{E} \times \vec{B} + \mu^2 r_H^2 (\vec{E} \cdot \vec{B})\vec{B})}{1 + \mu^2 r_H^2 |\vec{B}|^2}, \tag{4.5}$$

where r_H is the Hall coefficient, which has different values for electron and holes, μ is the mobility, which depends on the detector material, and $\vec{E}$ and $\vec{B}$ are the electric and magnetic

fields. In the simulation, different values of electric and magnetic fields have been used to emulate the condition inside the real CMS detector. Two particular cases are of interest:

- Case 1: the electric field and magnetic field are parallel. This is the case inside the forward end-cap pixel detector. In this case, equation 4.5 is simplified and becomes equation 4.2.
- Case 2: the electric field and magnetic field are perpendicular. This is the case inside the barrel detector, and in this case equation 4.5 is simplified and becomes:

$$\vec{v} = \frac{d\vec{r}}{dt} = \frac{\mu(\vec{E} + \mu r_H \vec{E} \times \vec{B})}{1 + \mu^2 r_H^2 |\vec{B}|^2}. \tag{4.6}$$

To understand the effect of the magnetic field on the charge carriers, is possible to split equation 4.6 in two components, one parallel to the magnetic field and one parallel to the electric field. Assuming that the two fields are perpendicular:

$$v_{\parallel E} = \frac{\mu E}{1 + \mu^2 r_H^2 |\vec{B}|^2}, \tag{4.7}$$

$$v_{\perp (E \times B)} = \frac{\mu^2 r_H E B}{1 + \mu^2 r_H^2 |\vec{B}|^2}, \tag{4.8}$$

and the ratio between these two components of the velocity gives the angle at which the charge carriers drift, and is known as the Lorentz angle:

$$\frac{v_{\perp (E \times B)}}{v_{\parallel E}} = \tan \theta_L = \mu r_H B. \tag{4.9}$$

Since the magnetic field and the Hall coefficient are fixed, the only way to change the Lorentz angle inside CMS is to change the electric field applied to the sensor or the temperature of the sensor. The relations between drift velocity and mobility as a function of the temperature and electric field are shown in figure 4.4 [50]. Here, the electron (hole) mobility is defined as the ratio between the electron (hole) drift velocity and the electric field. The relation between the electric field and the electron drift velocity is linear for low values of the electric field, and the mobility is constant. For large values of the electric field, the drift velocity saturates.

4.8 Digitization

The final step of the simulation is the digitization. The readout electronics of a pixel module are quite complex, and in the simulation a number of simplifications have been used. In the real module, the charge deposited inside the pixel is detected through its induced current, which can be calculated by using the Ramo-Shockley theorem [24]. The charge carrier drift induces an electric current on the readout pad, which after being amplified in a circuit, is converted by an ADC counter. If its value passes a fixed threshold, the pixel hit is streamed out at the periphery of the chip and registered. A realistic simulation of the readout circuit would take into account the time domain, to simulate time-dependent effects like the time walk.

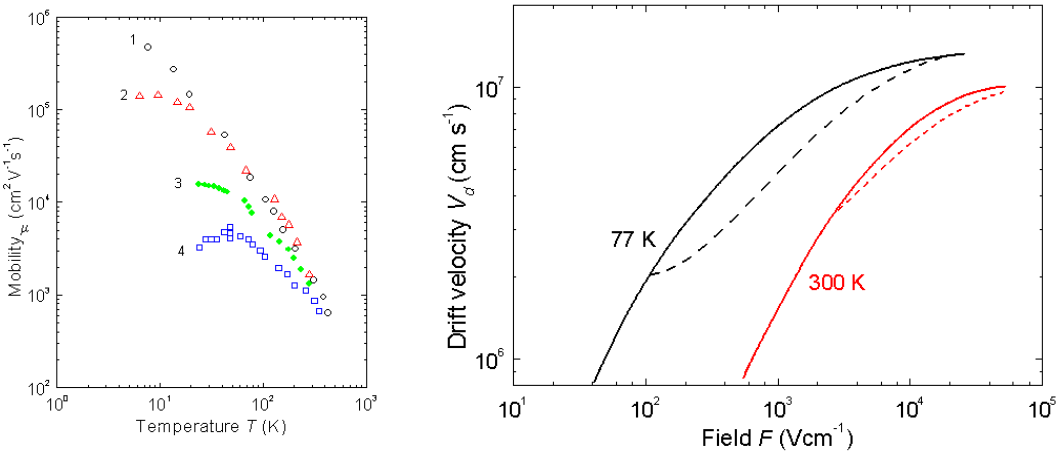


Figure 4.4: Left: Electron mobility versus temperature for different doping levels 1.High purity Si (Nd< 10^{-12} cm^{-3}) 2. High purity Si (Nd< $4{\cdot}10^{-13}$ cm^{-13}) 3. Nd= $1.75{\cdot}10^{16}$ cm^{-3}; Na = $1.48{\cdot}10^{15}$ cm^{-3} 4. Nd= $1.3{\cdot}10^{17}$ cm^{-3}; Na = $2.2{\cdot}10^{15}$ cm^{-3}). Right: Si. Electron drift velocity vs. electric field[51].

In this simulation, the time walk and other time-dependent effects are neglected, assuming the sensor has no radiation damage, and the readout circuit is properly calibrated and all the internal delays properly trimmed. Under these assumptions, the number of electrons that are finally read out corresponds to the real number of electrons inside the sensor. Next, the software emulates the digitization process, according to the number of available bits, which determine the digitizer resolution, and its threshold. The number of bits is fixed and equal to 4, which allows for 15 different charge values to be stored, plus one overflow bit. An example of the output of the digitizer is shown in figure 4.5. Each bin corresponds to a single pixel in the cluster. In this particular example, the threshold was set to a reasonable expected value of 1000 e^-. The pixels that pass the charge threshold are shown in green, while the ones under threshold are shown in red. The central pixel in red is an example of the under-fluctuation that can happen inside the detector. The algorithm treats these cases as if the cluster was reconstructed correctly (with a hole) and not as two separated clusters.

The energy deposit inside a pixel on average is proportional to the length of the particle track inside the pixel. Because of that, the central pixel's charge differs only by the small statistical fluctuation of the charge deposit, and gives a rough estimation of the hit position. The most important pixels for the position reconstruction are the external ones.

The digitizer produces as output the reconstructed muon hit position. This is compared to the real hit position, and the sensor resolution is extrapolated from the distribution of the residuals. To reconstruct the hit position the centroid-of-charge method was used. This method takes into account the available information from the cluster and extrapolates

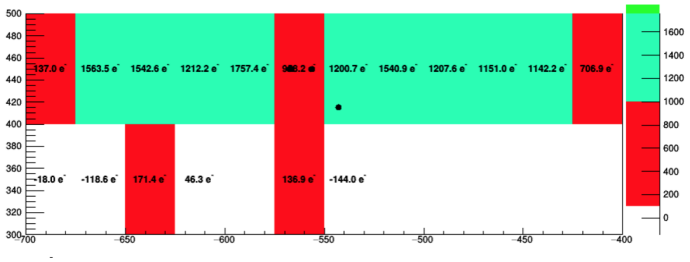


Figure 4.5: Single-event cluster as seen at the digitizer stage. In this case, the angle of the muon was 75°, and the cluster is quite elongated. Each bin corresponds to a single pixel in the cluster. In this particular example, the threshold was set to 1000 e^-.The pixels that pass the charge threshold are shown in green, while the ones under threshold are shown in red.

the hit position as:

$$x_c = \frac{\sum_i x_i \cdot c_i}{\sum_i c_i}, \tag{4.10}$$

where x_c is the reconstructed hit position, and x_i and c_i are the position and the charge of the i^{th} pixel inside the cluster. Figure 4.6 shows a selection of the residual and cluster size distributions for different incident angles. These results are for the 25 $\times$ 100 μm^2 configuration, along the smaller side, with a threshold of 1500 e$^-$ and a flat charge collection efficiency of 80%. The angles are 0°, 9°, and 45°, the corresponding to three cases of interest:

- at 0°, the tracks are perpendicular to the sensor and the cluster size is very close to 1. In this case, the centre of the cluster coincides with the centre of the pixel, and the residual distribution is almost flat. This residual distribution is commonly referred to as "digital resolution" , since if the particle is detected, the reconstructed position is by construction in the centre of the pixel. The two horns on the side are created by particles passing in neighbouring pixels, the charge is still collected when the track is close to the border of the pixel thanks to the transverse diffusion of the charge carriers during drift.
- at 9° the tracks cross the sensor at an angle so the cluster size is larger than 1 and the center-of-charge method can be used to reconstruct the hit position.
- at 45° the tracks cross the sensor at an angle so large that the cluster size is close to 6, the center-of-charge method can be used to reconstruct the hit position but the large cluster size will significantly degrade the final sensor resolution.

In the next section, the correlation between resolution, impact angle, and cluster size will be explained and analysed. The impact of using different pixel sizes, digitizer thresholds, and magnetic field settings will also be discussed.

4.9 Validation

The final step of the simulation is to parametrize the resolution of the single pixel module as a function of the incident angles for the barrel and forward pixel sub-detectors. The dip

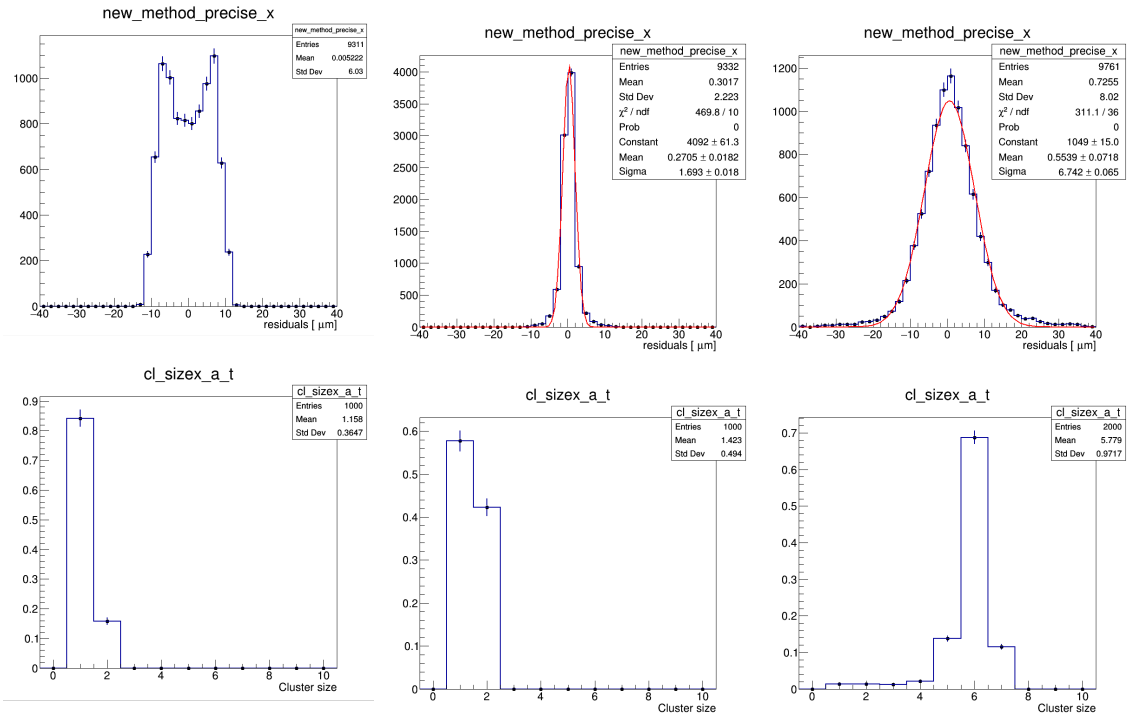


Figure 4.6: Residual (top row) and cluster size (bottom row) distributions for different incident angles. The cluster size increases with the incident angle, while the resolution as a function of the dip angle will be discussed in the following section.

angle α is defined as the angle between the particle and a line perpendicular to the sensor surface, and is shown in figure 4.7. The angle β is the angle between the particle trajectory projection on the surface of the sensor and a longitudinal line passing through the centre of the sensor, as shown in figure 4.7. In the first stages of this study, the simulation was

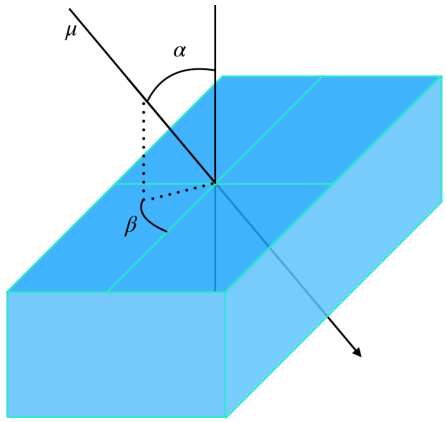


Figure 4.7: Definition of the dip angles α and β. The incident angles are the angles between the particle trajectory and a line perpendicular to the sensor surface, passing by the entry point of the particle in the sensor.

validated by comparing the result against the test beam study for the phase-1 pixel detector modules. I made the assumption that the resolution of the sensor can be factorised in two different components along the x and y axis. The simulation successfully predicts the trend in the resolution as a function of the dip angle. The curves qualitatively agree, as shown in figure 4.8, and the minimum occurs at the same place in both, at around 9 degrees. In a later stage, results from a test beam study for the phase-2 pixel detector modules were available, and it was possible to compare the simulation prediction with the test beam measurements. The comparison is shown in figure 4.8. The simulation successfully predicts the behaviour of the sensor. The curve that describes the resolution as a function of the dip angle is well understood. Starting from a dip angle equal to zero, at the left side of the plot, the resolution is equal to 6 μm. In this case the muon beam is perpendicular to the sensor, and the muons hit only one pixel inside the sensor. In this case, the resolution is dictated by the size of the pixel and is equal to:

$$\sigma = \frac{\text{pitch}}{\sqrt{12}} = \frac{25\mu\text{m}}{\sqrt{12}} \approx 6\mu\text{m}. \tag{4.11}$$

Increasing the incident angle, the particle starts hitting two pixels. In this case it is possible to use the charge centroid technique to reconstruct the centre of the cluster, and the spatial resolution improves. The resolution reaches a minimum for a value of the incident angle equal to 9°. This is the angle at which most of the particles passing through the sensor pass through exactly two pixels. This angle is related to the pixel size by the formula:

$$\theta = atan\frac{\text{pitch}}{\text{thickness}} = \text{atan}(\frac{25\mu\text{m}}{150\mu\text{m}}) \approx 9°. \tag{4.12}$$

For larger incident angles the mean cluster size becomes greater than two and the resolution degrades. This is expected, since the information about the hit is carried by the charge in the two external pixels, while the charge is almost uniform in the pixels in the middle of the cluster. Furthermore, for larger angles the particle traverse a small portion of the pixel, releasing a smaller charge. Nevertheless, the resolution has a small improvement every time the incident angle is such that the cluster size is very close to an integer number. We can generalize the previous formula 4.12 as:

$$\theta = atan\frac{(N\text{-}1)pitch}{thickness} = \text{(N=2) } 9°,\text{(N=3) } 18\ °,\ \text{(N=4) } 26\ °, \tag{4.13}$$

where N is the cluster size. This effect is strictly related to the threshold. Because of the digitizer threshold, particles that traverse a small portion of the pixel do not leave enough charge and the hit is lost. That is why the resolution improves every time the incident angle reaches a value described by formula 4.12. Another important aspect of the simulation is to faithfully emulate the effect of the magnetic field on the sensor resolution. The effect of the magnetic field has been studied for different magnetic field intensities, up to 10 Tesla. The magnetic field affects the charge carrier's motion as described in formula 4.6. The charge carrier's drift path is no longer parallel to the electric field direction, but instead they drifts at the Lorentz angle, described in equation 4.9. The final effect of the magnetic field is to translate the resolution curves. The effect is described by the following equation:

$$\sigma(tan(\theta), B_0) = \sigma(tan(\theta) - \mu r_H B_0, B = 0) = \sigma(tan(\theta) - \theta_L, B = 0), \tag{4.14}$$

sim vs phase 1.png

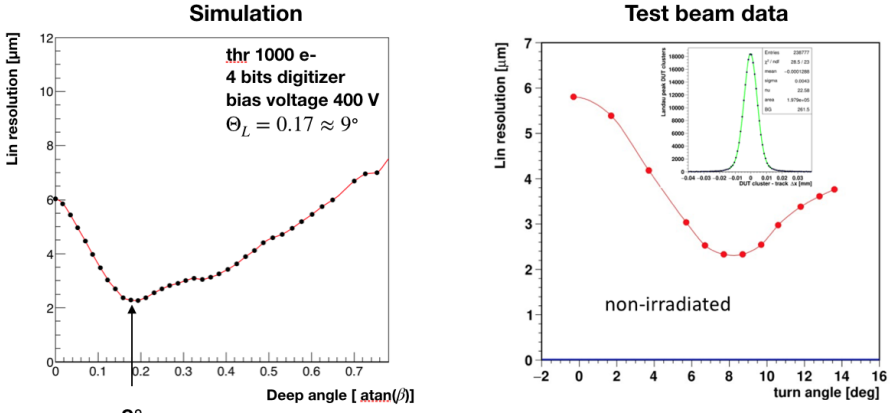


Figure 4.8: Left: simulation of the phase-2 pixel silicon sensor resolution as a function of the inverse tangent of the dip angle. Right: test beam resolution study result of the phase-2 pixel silicon sensor. The simulation successfully predicts the behaviour of the sensor.

which relates the resolution as a function of the dip angle for an arbitrary magnetic field B_0 to the resolution with a magnetic field B =0. The results for different values of magnetic field strength are shown in figure 4.9.

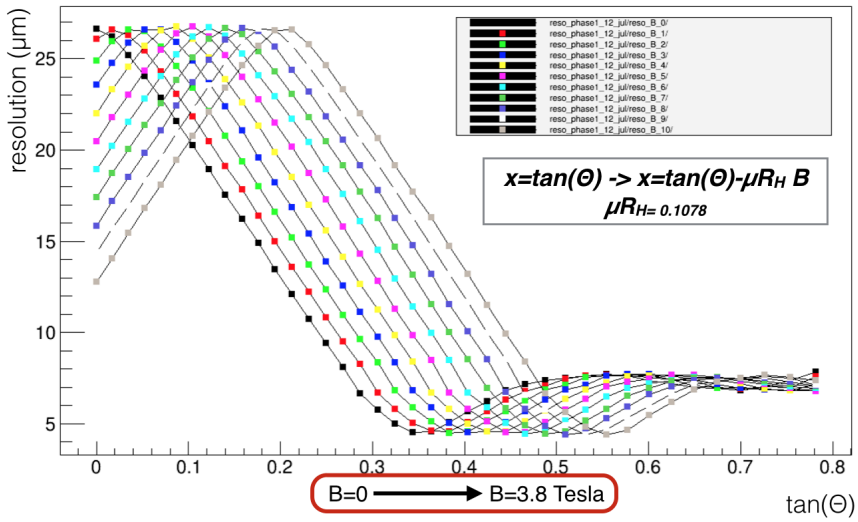


Figure 4.9: Simulation of the phase-1 pixel silicon sensor resolution as a function of the inverse tangent of the dip angle. The different curves show the effect of magnetic fields of different intensities. The final effect of the magnetic field is to translate the resolution curve by a factor $\mu r_H B$.

4.10 Resolution

The resolution as a function of the dip angle has been extracted for different configurations:

- Two different sensor sizes: 50×50 μm^2 and 100×25 μm^2, which are the two candidates

for the upgrade. The second option is the preferred one, because it improves the resolution along the r-ϕ plane while decreasing the cluster size along the z coordinate.

- Two different digitizer threshold values: 1000 and 1500 e^-, as possible options for the phase-2 detector.
- Two different Lorentz angle values: 0.055 and 0.110, corresponding to operational bias voltages of 150 V and 400 V [52].
- Two different magnetic field orientations, parallel and perpendicular to the electric field, corresponding to the two cases for the barrel and forward end-cap pixel detector.

The position resolution performance for the barrel region is shown in figures 4.10 and 4.11. The first figure shows the single-layer spatial resolution along the z axis for two different pixel lengths, 50 μm and 100 μm. In both cases the threshold is fixed to 1500 e^-. The second figure shows the resolution along the r-ϕ plane for two different pixel lengths, 25 μm and 50 μm. In both cases the threshold is fixed to 1500 e^-. The r-ϕ plane is perpendicular to the magnetic field, while the z axis is parallel to B. The effect of the magnetic field is to translate the resolution curve, as explained in the previous section. The effect of changing the threshold is shown in figure 4.11. The change of threshold affects the resolution curve in different ways. First, when increasing its value, some hits are not recorded. This effect is more visible at higher angle, when the charge deposit inside the pixels in a cluster is lower.

Another effect is the shift in the minimum of the resolution, which moves to higher incident angle. This is because the minimum in the distribution is strictly correlated to the cluster size, and with a higher threshold, a larger angle is required to register the charge deposit inside the two pixels. The behaviour for an incident angle $cotg(\beta) > 1.5$ is not very important for the proper detector operations, because almost all particles have a smaller incident angle.

4.11 Integration inside tkLayout

The resolution parametrisation extracted using the simulation was used inside tkLayout. Among other functions, the tkLayout workflow keeps track of the orientation of each sensor inside the detector, and transfers the resolution of each sensor from the local reference system to the global one, by using the relative orientation between sensors and the particle tracks. For each track, tkLayout takes into account the single-layer resolution, how many layers the particle track crosses, and the local dip angle for the hit on each layer. The final detector resolution is computed by performing an error propagation from the single hit resolution to the error on the track parameters, taking multiple scattering into account.. The number of tracks used to reconstruct the detector performance is $N = 3000$, if not stated otherwise. The latest pixel layout available at the moment is shown in figure 4.12. Figures 4.13, 4.14, 4.15, and 4.16 show the local resolution, the distribution of the resolution on the local $X/\ Y$ coordinate, and incident angle distributions for the barrel and end-cap pixel sub-detectors.

It is useful to analyse in depth one of these figures. Looking at figure 4.13 we are presented with four plots that describe different aspects of the tracks that hit the barrel

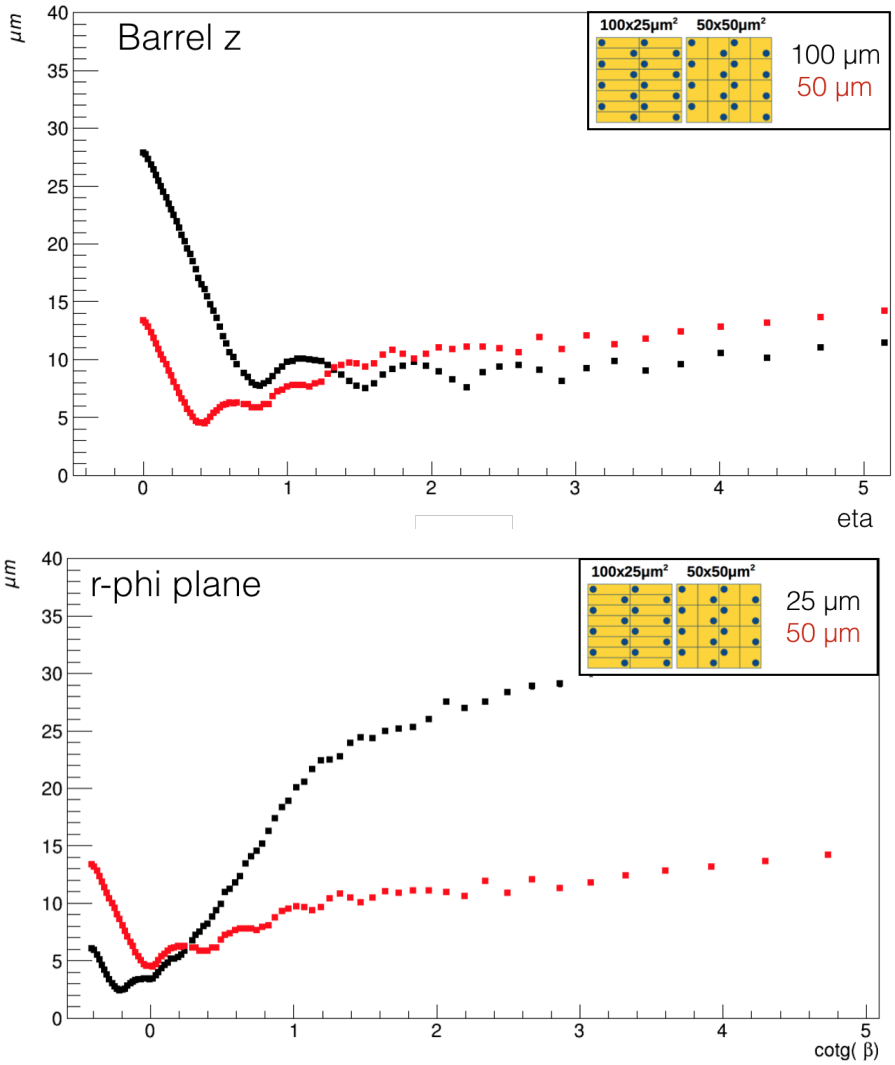


Figure 4.10: Single-layer resolution in the barrel of the CMS phase-2 pixel detector.

detector and their resolution:

- Top left: Resolution on the local X coordinate vs cotg(α)for a barrel module. This plot shows the parametrisation of the resolution obtained in the simulation. It is exported to tkLayout after applying the effect of the magnetic field.
- Bottom right: Distribution of the incident track ϕ (barrel). This plot shows the angles of the track in the r-ϕ plane, for all barrel modules.
- Bottom left: Distribution of the incident angles α for the barrel modules. This plot shows the distribution of the incident angles of the tracks with respect to the barrel modules in the *r-ϕ plane*. As can be seen from the plot, most of the tracks hit the sensor at a small angle, $\alpha < 0.4$.
- Top right: Distribution of the resolutions on local X coordinate (barrel modules). The information about the local resolution and incident track angles α are used to create

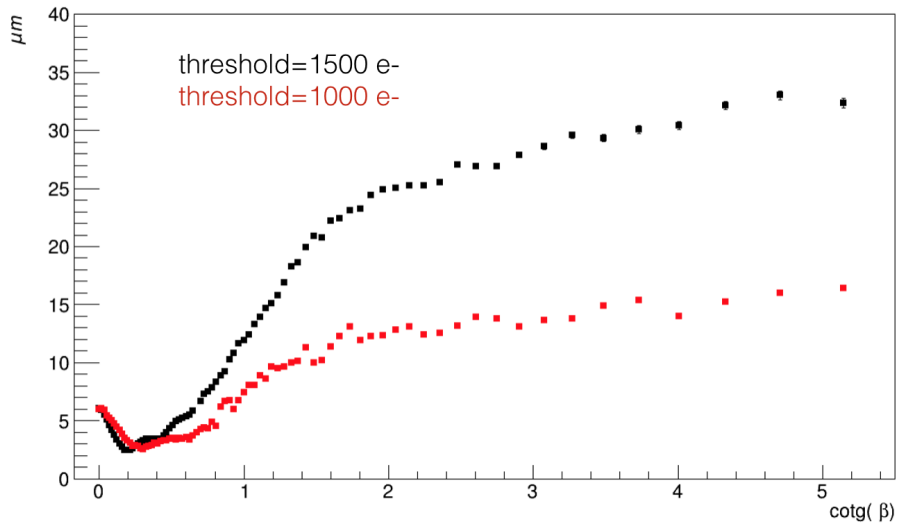


Figure 4.11: Resolution vs dip angle for the 25 μm pixel for two different digitizer threshold values, in red, 1000 e^- and in black, 1500 e^- .

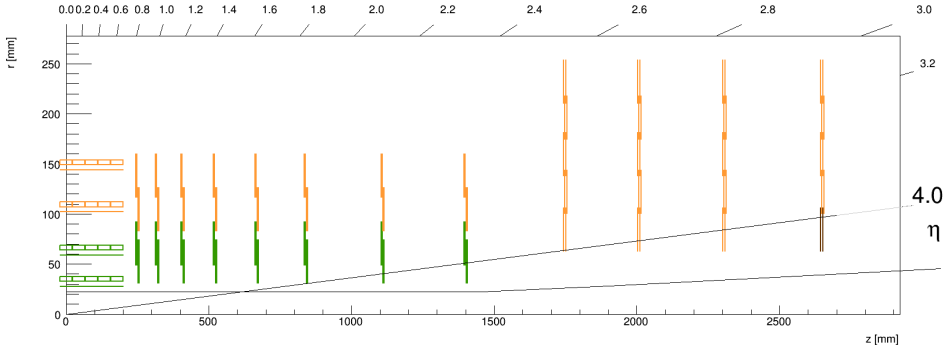


Figure 4.12: CMS phase-2 pixel detector layout iteration used in the TDR (IT614). A quarter of the pixel detector is shown, featuring the four pixel barrel layers, and the eight forward small disks and four end-cap large disks, which are duplicated at -z. The two different colours, green and yellow, identify two different pixel module sizes, while the ring in brown in the rightmost disk is the portion of the detector out of the nominal acceptance, which will be used for online instantaneous luminosity measurements.

this plot. This plot shows the distribution of the local resolution value extracted for for each track dependent on the incident angle.

Plots like these can be obtained for different geometries, pixel sizes, and module settings. These will then be used to compare different configurations. In the following section, as an example, the performance for two different pixel geometries will be compared.

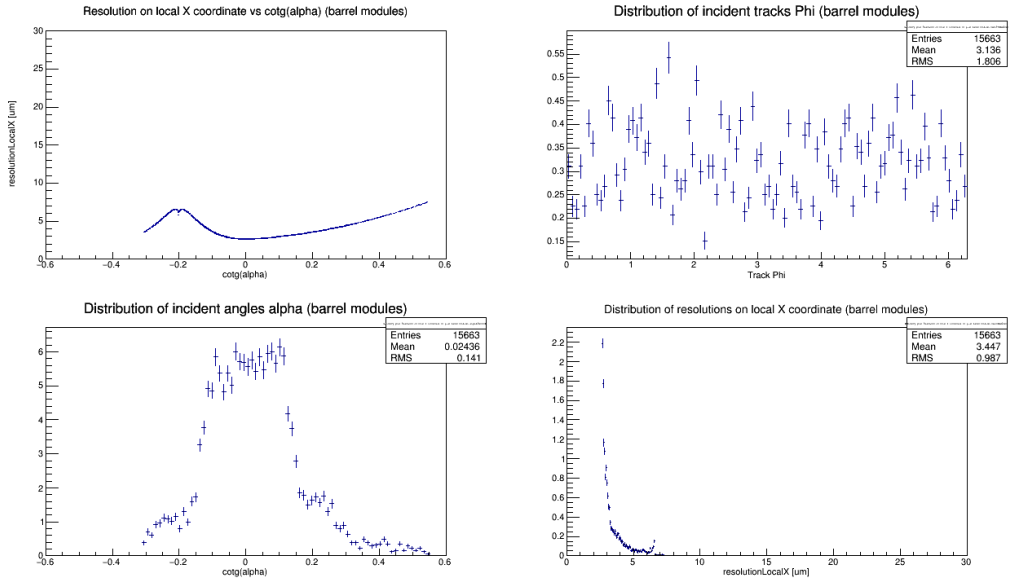


Figure 4.13: Top left: resolution on the local X coordinate vs cotg(α) for the barrel modules. Top right: Distribution of resolution on the local X coordinate. Bottom left: Distribution of incident angle α for the barrel modules. Bottom right: Distribution of incident track ϕ for the barrel modules.

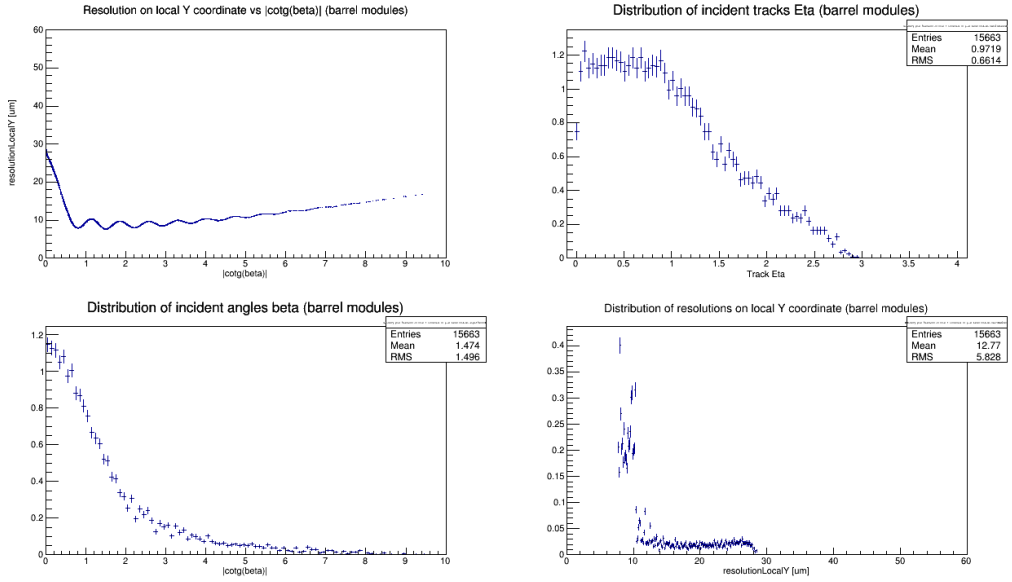


Figure 4.14: Top left: resolution on the local Y coordinate vs cotg(β) for the barrel modules. Top right: Distribution of resolution on the local Y coordinate. Bottom left: Distribution of incident angle β for the barrel modules. Bottom right: Distribution of incident track η for the barrel modules.

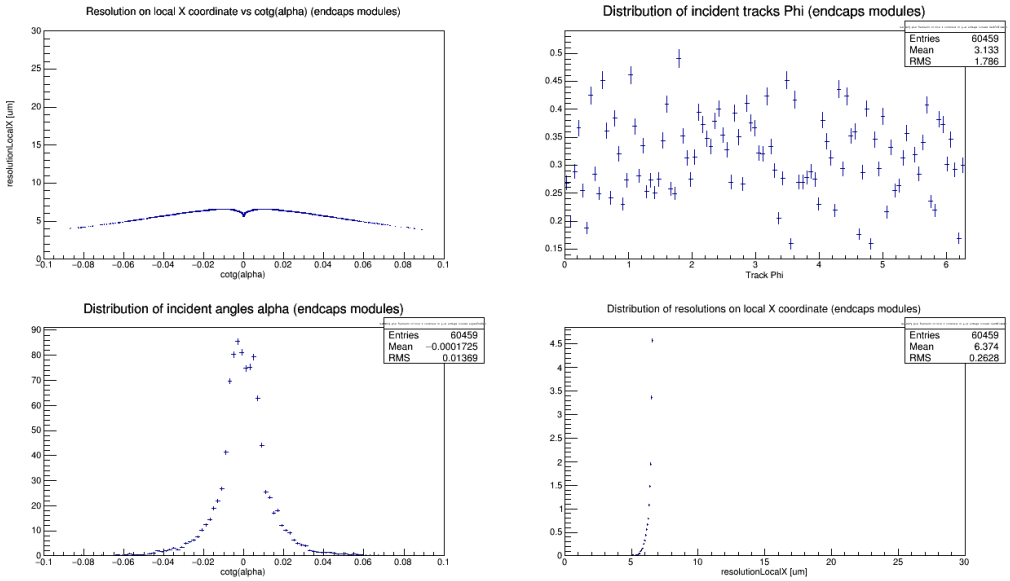


Figure 4.15: Top left: resolution on the local X coordinate vs cotg(α) for the end-caps modules). Top right: Distribution of resolution on the local X coordinate. Bottom left: Distribution of incident angle α for the end-caps modules. Bottom right: Distribution of incident track ϕ for the end-caps modules.

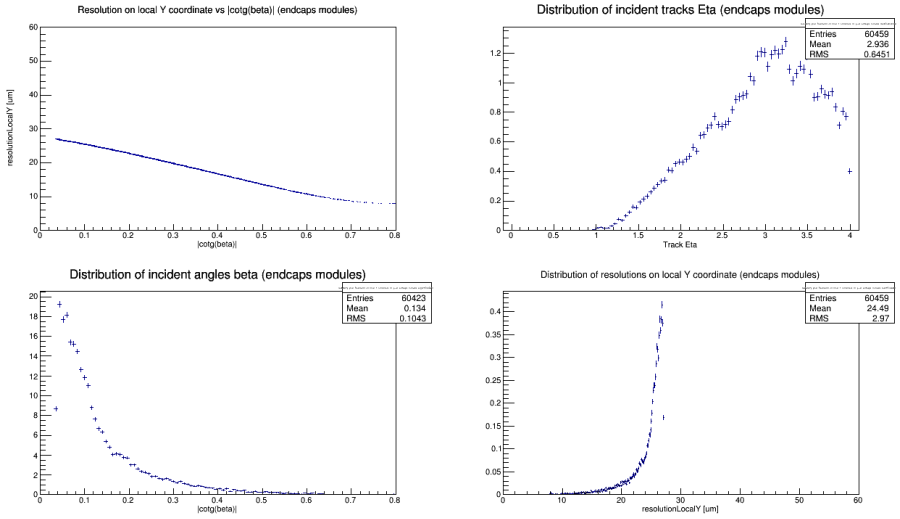


Figure 4.16: Top left: resolution on local the Y coordinate vs cotg(β) for the end-caps modules. Top right: Distribution of resolution on the local Y coordinate. Bottom left: Distribution of incident angle β for the end-caps modules. Bottom right: Distribution of incident track η for the end-caps modules.

4.12 Geometry optimisation

After the integration of the parametrisation in tkLayout, the tracking performance of different detector geometries can be estimated. As mentioned, two different pixel geometries are considered for the construction of the detector, 50 $\times$ 50 μm^2 and 100 $\times$ 25 μm^2. The

thickness of the sensor in both cases is 150 μm. Using the same geometry but different pixel sizes, it is possible to predict the strengths and weaknesses of the two configurations. In figure 4.17 the performance of the detector for the two pixel sizes is compared. As seen

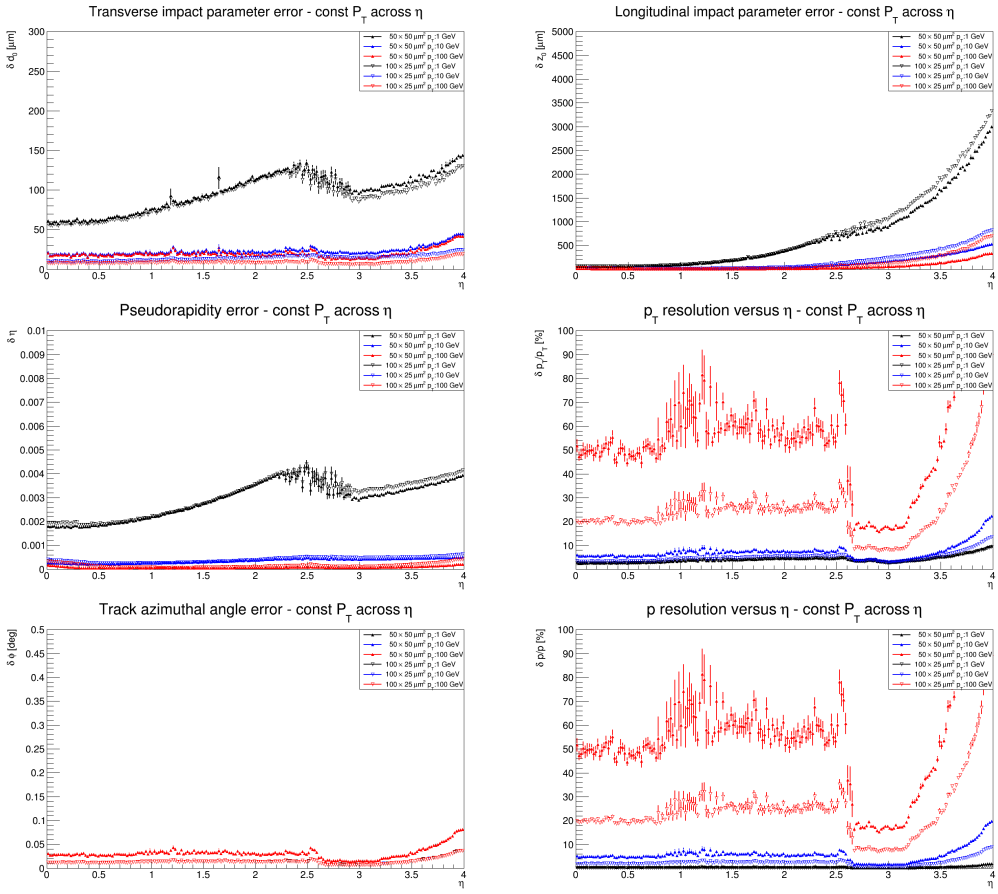


Figure 4.17: Performance of the detector for different value of the p_T and for two different pixel sizes, 50 $\times$ 50 μm^2 and 100 $\times$ 25 μm^2. The black, blue and red markers identify the performance for different p_T, respectively, 1, 10 and 100 GeV. The full triangles are used for the 50 $\times$ 50 μm^2 configuration, while the empty upside down triangles are used for the 100 $\times$ 25 μm^2 configuration.

from figures 4.17, the 100 $\times$ 25 μm^2 configuration outperforms the 50 $\times$ 50 μm^2 configuration except for the longitudinal impact parameter resolution at high η for low p_T tracks. One important advantage of using this configuration that is not visible from this plots is the bandwidth reduction. The 100 $\times$ 25 μm^2 configuration is especially useful in the periphery of the barrel, where the track incident angle is large and the cluster size is larger. As the best resolution performance is obtained for a cluster size of 2, and a larger cluster size requires more bandwidth to be read out, the 100 $\times$ 25 μm^2 configuration has the best performance. The pixel size that will be used for the construction of the detector has still to be decided, but physics performance aspects point toward the 100 $\times$ 25 μm^2 configuration. The tkLayout software has been used extensively to optimise the layout of the pixel and outer tracker, through several iterations, and the simulation of the sensor resolution has been an important consideration.

Chapter 5

Time reconstruction

5.1 Introduction

During my last years of the PhD I was lucky enough to work with low-gain avalanche diode (LGAD) detector technology. These interesting detectors achieve extremely good time resolution, of the order of 20-30 ps for charged particles and 5-10 ps for photons. I implemented a new lab setup at UZH with a radioactive strontium source that would allow us to measure the time resolution of these detectors. After characterising LGADs from different production batches and vendors, like FBK and CNM, I moved on to characterise a new promising type of detector, the AC-coupled LGAD. As I will discuss in detail in the following chapter, one of the setbacks of the LGAD is the relatively large region without multiplication (gain) between the pads, such that there is effectively a dead area with a width of 45-100 μm, depending on the different design choices. This translates into a decreased active area or fill factor, especially in the case of a pixel-like implementation. The AC-LGAD technology resolves these problems and has a theoretical fill factor of 100%, while keeping an excellent time resolution performance, and also achieving spatial resolutions that outperform standard devices. In the following chapter, I will share my journey into the picosecond time-resolution realm and the results obtained along the way.

5.2 Timing at CERN

The collision rate inside the CMS experiment is 40 MHz, or one bunch crossing every 25 ns [37]. The number of collisions per bunch crossing depends on how many protons are in each bunch, how collimated the bunches are, and the overall proton density distribution in the bunches. The number of simultaneous collisions inside a single bunch crossing is called pileup. Rare events are of special interest in HEP experiments, and so it is desirable to have as many collisions as possible. At the same time, high pileup can degrade the detector performance. The goal of LHC scientists is to run their experiments in a high pileup regime with the requirement of maintaining high detector performance.
The high pileup environment poses different challenges to the detectors [53]:

- The high number of hits inside the detector can quickly saturate the readout bandwidth, especially in the detector region closer to the interaction point where the gran-

ularity of the detector is higher. The CMS experiment overcomes this problem in an elegant way with a two-level trigger. The pixel detector is not part of the first-level trigger and the data are streamed out only if the level-one trigger generates a trigger signal. On the other hand, this architecture requires the readout chips to store pixel hits on the detector module while waiting for the trigger decision, which requires space and power.

- The reconstruction is more complicated and time consuming and grows with the number of tracks, and consequently with the number of interactions per bunch crossing. The time necessary to reconstruct an event with N tracks is proportional to the factorial of N, $N!$. Furthermore, the high number of tracks inside the detector creates a number of unwanted effects like track and vertex merging, and the creation of combinatorial tracks. These are collections of uncorrelated hits, that the reconstruction algorithm has erroneously recognised as being generated by the same particle.

- The radiation damage inside the silicon sensors is proportional to the luminosity of the beam, so an overall higher pileup, proportional to the luminosity, is detrimental to the sensor performance.

Summarising the requirements to work in a high luminosity/PU environment, a detector has to be designed with high bandwidth readout, a large buffer, and appropriate radiation hardness. All the sensors inside the CMS tracker are built to measure a combination of energy deposits and/or position measurements.
There is a fifth hit parameter that has not been used so far on the CMS experiment: the time of arrival (TOA). The time at which a particle reaches a sensor depends on two factors: when it was created and the path to reach the detector. The proton-proton collisions inside the same bunch crossing happen at different times and in different positions. Particles coming from the same collision vertex reach the detector layers at different times, depending on the direction (η, ϕ), and additionally the TOA for charged particle depends on the transverse momentum (p_T). The TOA is the sum of the time at which the particle was created and the time of flight (TOF). Correcting for the TOF, it is possible to time tag a particle and determine which vertex it originated from. Simulations [53] have shown that this approach is successful and can be used in a number of ways:

- Pileup reduction and improved MET measurement: Particles that have an arrival time at a detector layer consistent with originating from pileup vertices can be rejected from consideration in the missing transverse momentum measurement.

- Particle identification: especially at low p_T, the time of flight can be used to discriminate between kaons and pions.

- Object reconstruction: a technique used in a broad spectrum of analyses requires the definition of a so-called isolation cone. For example, in the identification of hadronic tau leptons, an isolation cone is defined around the decay products to assess if the particles originating from the tau decay are not part of a jet. This ensures that the object in question is a real tau, and not a combinatorial tau built from random

tracks inside a cone. The time tag is a new handle on top of the standard "geometric isolation".

These performance improvements in object reconstruction and pileup rejection have consequences on the projected sensitivity of analyses, which have been thoroughly investigated within the MIP Timing Layer Technical Design Report (MTD TDR) [53].

5.3 MIP Timing Detector (MTD)

The CMS experiment will implement a timing detector in order to improve measurements at the high instantaneous luminosities of the HL-LHC. The MIP Timing Layer (MTD) [53] will supply timing information for MIPs with 30-40 ps resolution and will be operational at the beginning of HL-LHC operation in 2026, degrading as a result of radiation damage to 50-60 ps by the end of HL-LHC operations. The MTD exploits the fact that the individual interactions within a bunch crossing do not all occur at precisely the same time but, because of the longitudinal extent of the beams (4.5 cm RMS), they are distributed over time with a root mean square value of 180-200 ps. The MTD will provide timing in the barrel and end-cap regions, with a hermetic coverage up to a pseudorapidity $|\eta| < 3.0$. The system will use two different approaches in the barrel and end-cap regions. The MTD barrel detector will be comprised of LYSO crystals optically connected to SiPMs, while in the forward region, the MTD end-cap detector will use LGAD silicon sensors. The difference in approaches is dictated by the available space to insert the new timing layer and by the requirements in granularity, which are quite different between the barrel and end-cap regions. While the material budget of this new sub-detector is still kept to a minimum, to avoid the deterioration of the track reconstruction, the new timing layer will sit in between the tracker and the electromagnetic calorimeter. For a future upgrade in LS4 or beyond, UZH and PSI are pursuing the novel idea of extending the timing layer to cover the full fiducial volume of the tracking volume, up to $|\eta| < 4.0$, by replacing some disks inside the TEPX pixel end-cap detector with disks that implement 4D pixel detectors using LGADs. The disks would retain the position resolution and hit efficiency of the baseline pixel detector, adding, on top of that, a time-tagging functionality for each track reconstructed in the TEPX.
This upgrade would represent the first truly 4-D system in CMS, merging together the tracking and timing functionality, and therefore reducing the combinatorics of the pattern recognition.

5.4 Timing sensors: LGAD, AC-LGAD

A number of silicon technologies are able to achieve a time resolution of the order of tens of picoseconds. In this work I will focus on low-gain avalanche diodes (LGADs) [54], and in particular the AC-coupled LGAD [55]. The LGAD sensors developed for the MTD have the following interesting features:

- a typical thickness of 50 μm or less,

- a single-pad area is usually around 1 mm^2. The geometry chosen for MTD is 1.3×1.3 mm^2, with the pads arranged in a 16×16 array.

- internal gain: the sensor uses the n-in-p technology with the insertion of a p-doped layer close to the readout cathode. When a bias voltage is applied, a strong electric field is created in this region, where charge multiplication (gain) occurs.

- Wide termination structure: the sensor pads are isolated by a junction termination structure, with a width varying from 50 to 100 μm, depending on the technology and design of the particular production.

The sensors have a relatively large area compared to their thickness, of the order of 1-2 mm^2 per pad, and have internal gain. The internal gain improves the signal over noise ratio, leading to cleaner and stronger signals. This geometry, coupled with the large termination structure, ensures a uniform electric field in the sensor, which translates into a fixed drift time for the charge carriers, irrespective of the hit position. The good time resolution is a consequence of the homogeneous electric field coupled with the internal gain of the sensor. As an example of the typical structure of an LGAD, the properties of the devices produced in the AIDA-2020 are here detailed [56]. The structure of the sensor is shown in figure 5.1. In this case, the structure includes two identical 1 mm^2 sensors. The figure is symmetrical, and the two pads are isolated by two Junction Termination Extensions (JTE) and one P-stop. The structures are 50 μm thick and are created on top of a 300 μm thick substrate. The substrate does not have any effect on the sensor performance, in fact thinning is possible, but contributes to the structural integrity of the sensor and allows for easier handling. The metal deposited on the top and bottom of the sensor creates contact pads used to bias the structure and connect to the readout electronics. The metal is reflective and opaque to UV light, so a small optical window is usually included on the top metallisation layer, allowing the structure to be tested using a laser. The doping concentration of the bulk, the gain layer and the size of the P-stop are more or less free parameters that have been optimised taking into account different factors. The aim of this kind of sensor is optimal timing, which requires a high gain and high signal-to-noise (S/N) ratio.

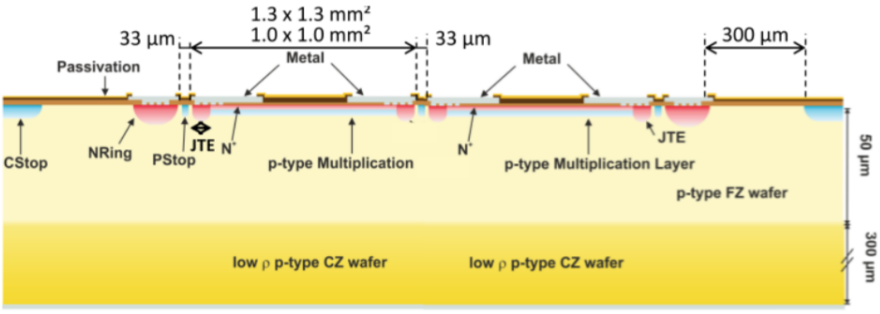


Figure 5.1: Cross section of an AIDA-2020 LGAD structure, with two identical sensors in the middle. The figure shows the different doping concentrations and structures inside the LGAD, like the multiplication p-type layer, the JTE, and the P-stop [56].

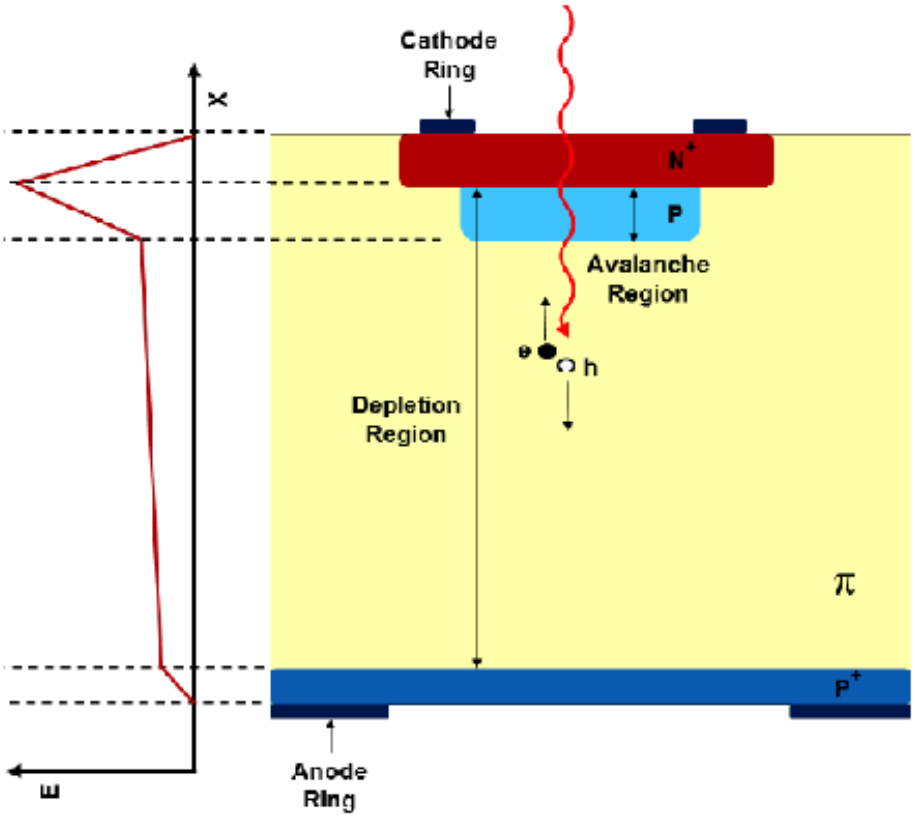


Figure 5.2: Electric field intensity as a function of the depth inside an LGAD sensor. The electric field is uniform in the depleted active region and rises in the gain layer. In that region, the charge is amplified before being collected [56].

5.5 Radiation hardness

The LGAD sensors will be installed inside the CMS detector, so a good radiation tolerance is required. In the first place, it is known that the performance of the sensor is very sensitive to small changes in the gain-layer implantation process. The loss of the multiplication layer due to radiation is caused by the acceptor removal mechanism that decreases the doping density of the gain layer, depending on the received fluence, until it cannot any more generate a high enough electric field. Different combinations of dopant have been tested to mitigate the gain layer removal, the result of these tests are summarised in figure 5.3. I will now discuss the different dopant combinations and their effect. One of the most common dopants used to co-implant together with boron is carbon. Recent studies [57] show that carbon co-implantation mitigates the loss of the gain layer after irradiation, while replacing boron by gallium does not improve the radiation hardness. Co-implantation with carbon attenuates the acceptor removal and has contributed to improving the radiation hardness up to fluence values of $2.5 \times 10^{15} n_{eq}/cm^2$. The position and doping profile of the gain layer play an important role. The shape of the gain layer profile, and its concentration and thickness, can be changed to improve the LGAD's performance. It has been proven that deep, narrow boron doping profiles, with a high concentration peak achieved by a low thermal diffusion process, are less prone to be unactivated after irradiation [58]. The deeper

implant also enhances the recovery capability of the bias voltage. Since the multiplication happens between the gain layer and the readout layer, a longer distance between the gain layer and the readout directly translates into a wider multiplication area. Of course the local electric field in this region will be lower for sensors with deeper implants, but the overall effect is an enhanced recovery capability bluefrom radiation damage. Another method to improve the radiation hardness is to reduce the bulk thickness, such that there is a higher electric field in the gain region: for example, a bias increase of 100 V in a 25 μm thick sensor increases the field by 4V/μm, while only by 1V/μm in a 100 μm thick sensor.
Having thinner detectors not only increases the radiation hardness, but also improves their timing performance: it provides a faster collection time and higher slew rates, which in the end, implies lower timing jitters. On the other hand, thinner sensors have higher capacitance and produce a smaller signal, so a good compromise between these figures of merit has to be found. A simulation [59] of the influence of the LGAD's thickness on the slew rate as a function of the gain is shown in figure 5.4.

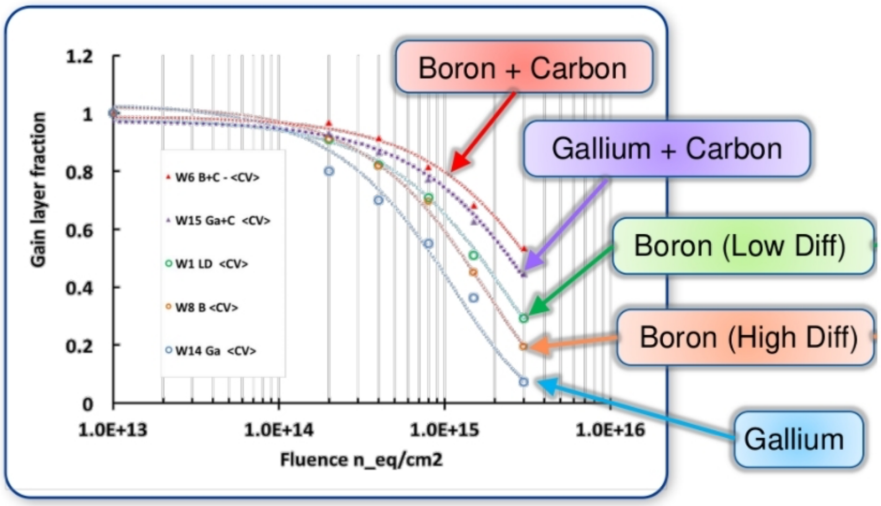


Figure 5.3: Gain layer degradation as a function of the fluence for different engineering techniques of the gain layer [60].

Finally, increasing the bias voltage of the sensor is a technique used to recover the performance. Because the resistivity of the sensors increases with the fluence, it is possible to operate them at higher voltages and enhance the gain. This recovery effect is shown in figure 5.5. The figure shows the timing performance as a function of different bias voltages for three irradiated HPK LGAD sensors measured at UZH. The sensors are from the same wafer. For larger radiation doses, the bias voltage has to be increased to recover the timing performance, nevertheless, it is encouraging that it is possible to recover the initial sensor performance by increasing the bias voltage.

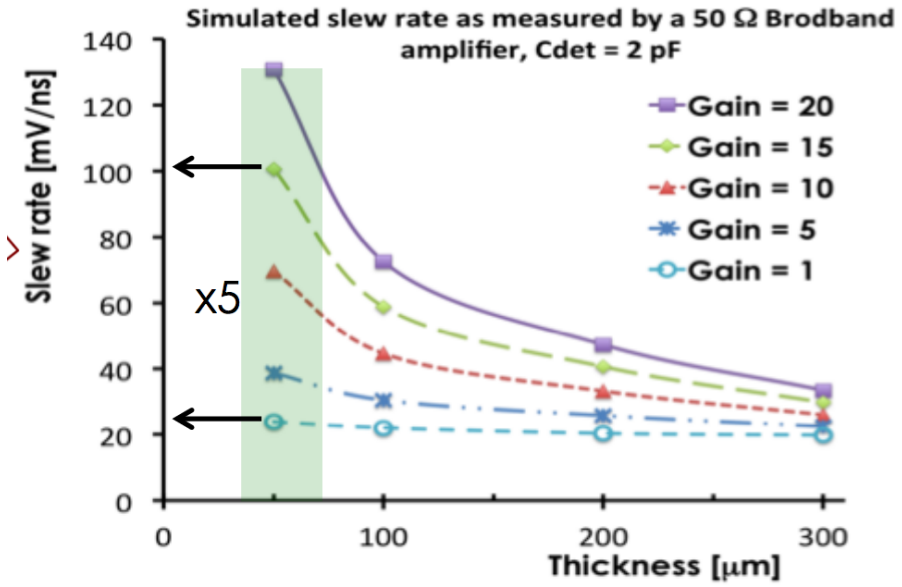


Figure 5.4: Simulation of the influence of the LGAD's thickness on the slew rate as a function of the gain.

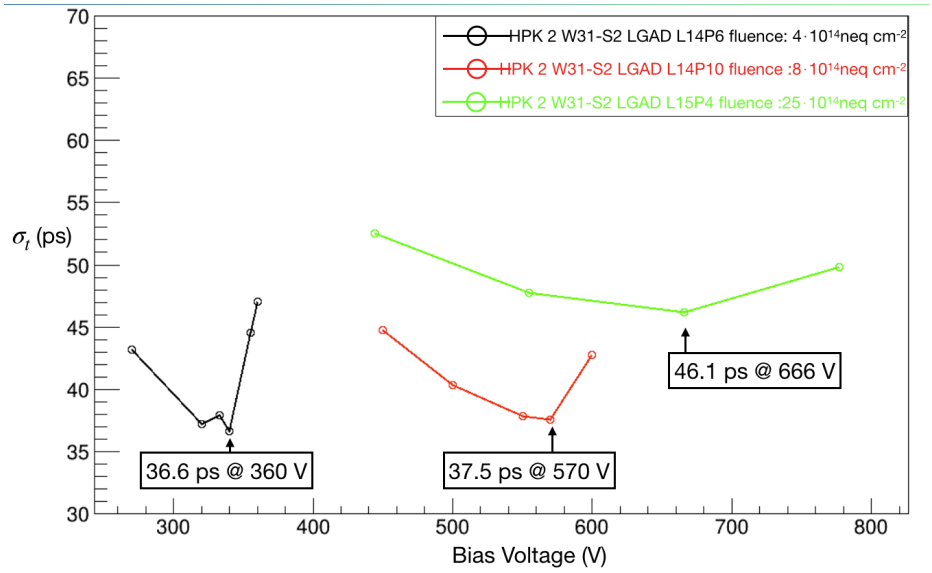


Figure 5.5: HPK LGAD's timing performance after irradiation as a function of the bias voltage. The sensors used are three identical sensors.

5.6 LGAD fill factor and AC-LGAD technology

The wide termination structure of the standard LGAD is an undesirable feature. It translates into a poor fill factor for smaller sensors, so much smaller no-gain areas are needed for pixel-like geometries. To overcome these LGAD limitations and achieve good time and space resolution, a new type of LGAD sensor has been developed, the AC-LGAD [61]. In this configuration the gain layer is not segmented and spreads throughout the entire sensor,

and the termination structures are not implemented. On top of the standard boron-doped gain layer, the n-type cathode is implemented. The readout segmentation is obtained at the level of the AC-coupled metal pads, which are capacitively coupled to the detector bulk via a dielectric spacer layer deposited between the silicon and the readout. The readout layer is segmented in smaller pads allowing for a pixel-like readout structure, while the continuous gain layer and the absence of a p-stop ensure a 100% fill factor. The pad size and shape are free parameters that have to be chosen carefully, together with the resistivity of the coupling layer. Furthermore, the resistivity of the n-doped layer is an additional free parameter that has to be optimised. Two different testing systems were used in the course of this thesis to study the performance of standard LGADs and AC-LGADs, the first one implementing a radioactive source and the second one a TCT laser. The experimental configuration, the readout electronics and preparation of the sensors, together with the results, will be discussed in the following sections.

5.7 Readout boards

In order to use and test the LGAD sensors, custom readout boards have to be used. The board connects the sensor to the high voltage and the ground potential, as well to the readout electronics. The sensor is glued to the readout board using electrically conductive tape, and connected to the power and readout pad using wire bonds. I used different kinds of boards developed by collaborators, the main differences being the number of readout channels and the presence or absence of a first stage amplifier on the board itself. Having a first stage amplifier close to the sensor is advantageous. The signal charge is usually of the order of 10-15 fC, and for such small signals any intermediate connection or cable in between the sensor and the first stage amplifier can introduce significant losses and distortions.
In this work three different boards were used:

- Torino board: a small board without amplification and with three readout channels.
- University of California at Santa Cruz (UCSC) board [62]: this board has one readout channel but a first stage amplifier built in the PCB.
- Zurich board: a development of the Torino board. This board hosts nine passive readout channels and features other minor improvements.

The boards are shown in figure 5.6. The UCSC readout board amplifier is based on a single transistor common-emitter design and acts as an inverting trans-impedance amplifier. Amplification is performed by an AC-coupled silicon-germanium bipolar transistor with a bandwidth of 75 GHz. At a bandwidth of 1.9 GHz, a gain of 29 dB is expected, while an integrated output noise of 260 μV is estimated. The feedback loop is designed for timing with small-capacitance sensors inducing typical rise times of the order of 800 ps, with a feedback resistor of 470 Ω.

Sensors are attached to the boards using double-sided conductive tape while the amplifier input is coupled to the front side metallization layer via multiple wire bonds to minimize inductance. A 1 MΩ resistor attached between input and ground serves for

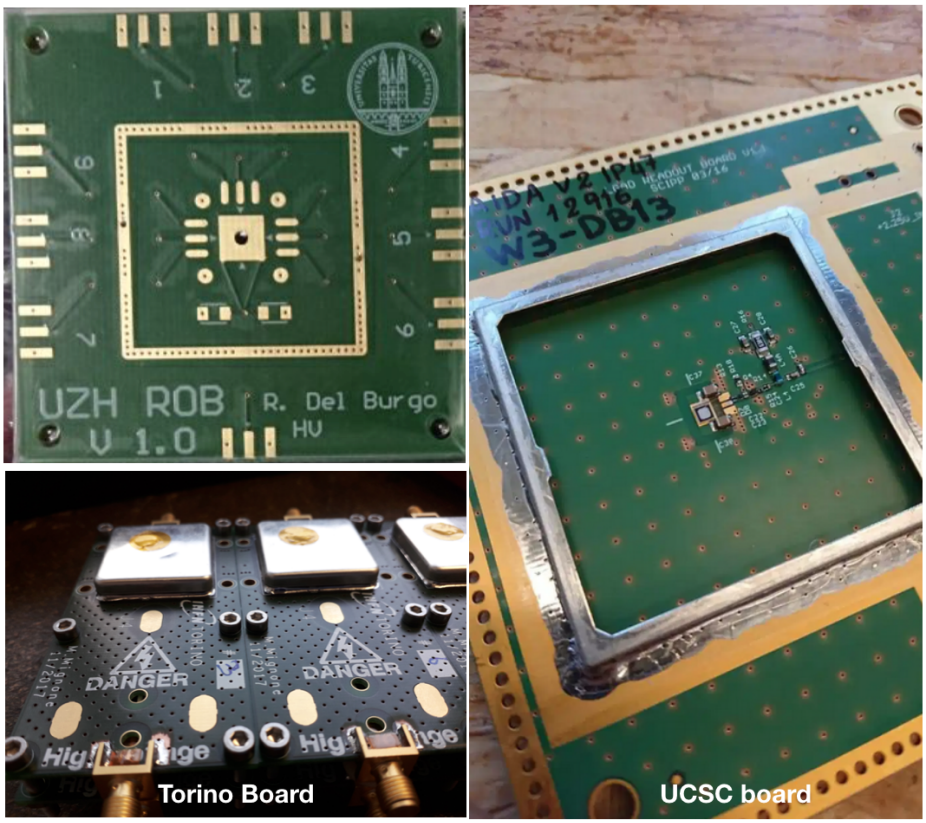


Figure 5.6: Top Left: UZH readout board. Bottom left: Torino readout board. Right: UCSC readout board.

detector biasing, followed by a pair of low forward-resistance silicon pin diodes. The latter, with a 50 V breakdown at 5 μA functions as a protection for the amplifier input.

All the boards feature an EM shielding to protect the signal from external noise.

5.8 Electrical characterisation of the LGAD sensors.

The timing sensors are characterized using different experimental configurations. First, the electrical properties of the sensors are tested performing IV (leakage current versus bias voltage) and CV (bulk capacitance versus bias voltage) scans. The IV scan is used to measure the breakdown voltage of the sensor and the leakage current, which changes with irradiation level and has a large influence on the performance and operational conditions of the sensors. Different probe stations have been used to measure the properties of the sensors during this thesis work, the latest one shown in figure 5.7 on the left side. On the right side, a collection of IV curves from sensors of the first AIDA-2020 run is shown. In this particular run of sensors, the leakage current was two orders of magnitude larger than expected. This is related to a defect in the JTE termination design. This defect does not degrade the timing performance of the sensors, and has been corrected in successive runs [56]. The CV scan measures the bulk capacitance of the sensor. The bulk capacitance depends on the applied bias voltage, or more precisely on the depletion depth of the sensor. Therefore, studying

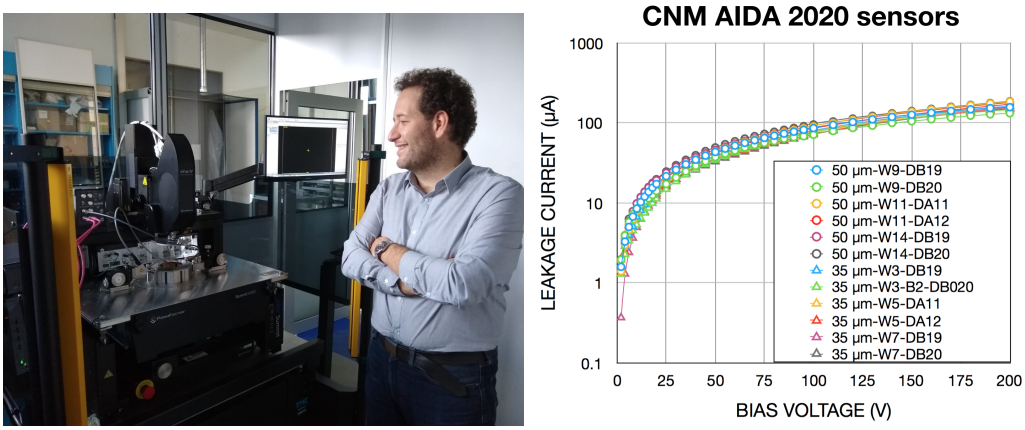


Figure 5.7: Left: probe station in the clean room at UZH. Right: IV curves of LGAD sensors from the first AIDA-2020 run.

the CV curve of the sensor, it is possible to derive the value of the depletion voltage of the bulk and of the multiplication layer. A CV scan for normal LGAD pads, produced at HPK, and AC-coupled LGAD from FBK [55] is shown in picture 5.8. Two separate flex points are visible in both curves. The one at lower voltage values is the voltage at which the gain layer is depleted. The second one shows the depletion voltage of the silicon bulk underneath. The AC-LGAD sensor has a higher depletion voltage because of the depletion of the additional resistive layer on top of the standard LGAD structure [55].

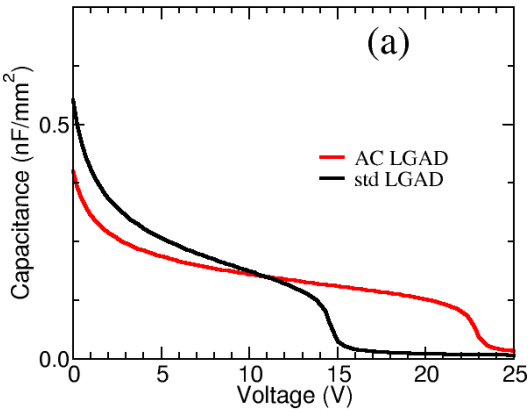


Figure 5.8: CV curves for a normal LGAD (black) and an AC-LGAD (red) [55].

Using the probe station, the electrical characteristics of different LGAD runs have been studied before testing the sensor performance using the Strontium (^{90}Sr) and the TCT systems.

5.9 Time resolution extraction

In the following sections I will present the time resolution performance of different sensors using two different setups. The resolution is extracted using two different configurations: two identical sensors, one sensor under test refereed as device under test (DUT) and one reference sensor. All the result are obtained without averaging signals, so the unfiltered original signal coming from the device is used. The time resolution of a single sensor for the two identical sensor configuration is extracted as:

$$\sigma_{sensor} = \sigma_{system}/\sqrt{2} \tag{5.1}$$

The time resolution for the sensor under test and one reference sensor configuration is extracted as:

$$\sigma_{DUT} = \sqrt{\sigma^2_{system} - \sigma^2_{ref}}. \tag{5.2}$$

5.10 Experimental setup with a radioactive source

Two different experimental approaches have been used to test the performance of the sensors. The first one used is the charge-collection system with a Strontium ^{90}Sr source. Two sensors, mounted on PCBs, are aligned under the radioactive source. When an emitted beta particle passes through the two sensors, the event is saved and the time resolution is then computed offline. The extraction of the time resolution will be explained in great detail in the following sections. The radioactive source has been chosen because of its energy spectrum, shown in figure 5.9 [63]. Strontium decays to Yttrium[64], emitting a low energy electron, with an energy of the order of half an MeV. The Yttrium decays afterwards to stable Zirconium, emitting an electron with a maximum energy of 2.3 MeV. The energy released in silicon by the 2.3 MeV electron is quite close to the energy released by a MIP in silicon. The use of a beta emitter source allows us to study an important contribution to the timing resolution of the sensor. The energy release of a photon and charged particle inside the sensor is quite different. The photon releases its energy in a point-like interaction, up to a maximum fixed depth depending on its energy. The charged particle, on the other hand, releases its energy along its path. Furthermore, the energy release is nonuniform, the so-called Landau effect [65]. Because the signal in the sensor starts with the first charged carriers generated, the nonuniform charge distribution degrades the timing performances. This effect is not present with photons, for example in the case when we test the structures using the TCT system. The setup with the radioactive source does not allow position tracking. It is therefore not possible to reconstruct an efficiency map of the sensor, but only to extract the signal characteristics and the time resolution. On the other hand, beta particles have enough energy to cross multiple sensors. This allows two sensors to be aligned in the same setup to extract the time resolution, with one device giving the time reference for the other. The Strontium set-up has been used in two different configurations, as shown in Figure 5.10. In the first one, one device under test (DUT) is measured in conjunction with a reference LGAD sensor, with a known time resolution. In the second configuration, two sensors of unknown but identical time resolution are under test.

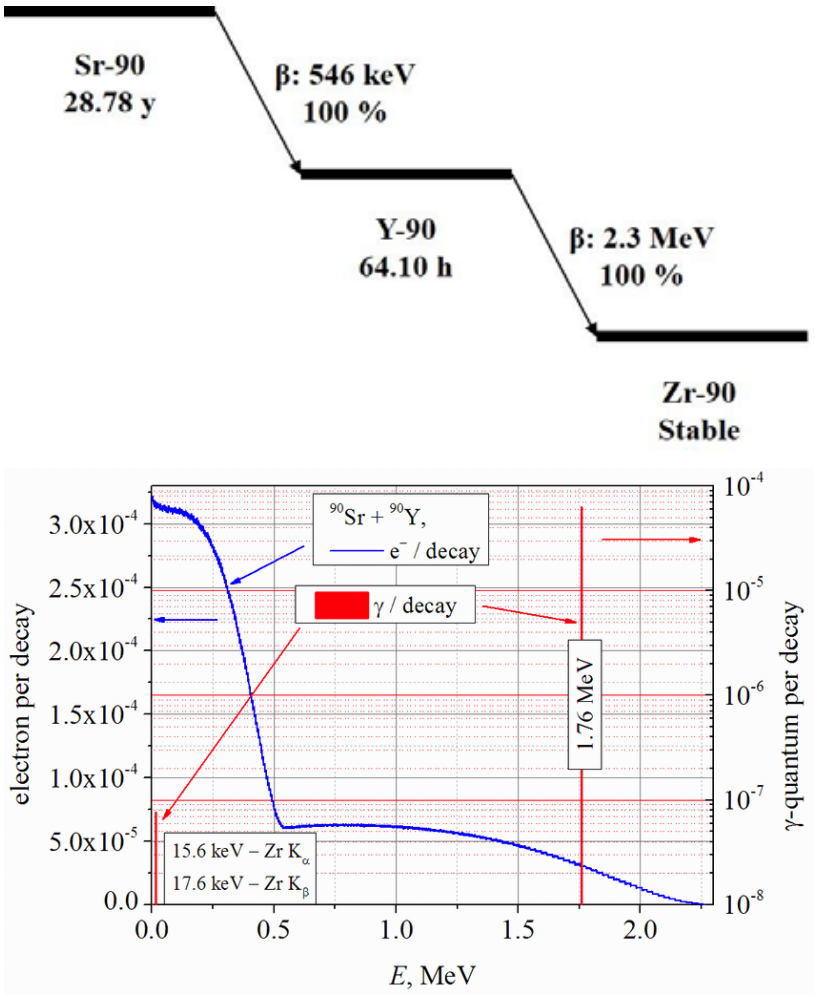


Figure 5.9: Sr-90 decays and energy spectrum. The Strontium decays to Yttrium, emitting a low energy electron, with an energy of the order of half an MeV. Subsequently, the Yttrium decays to stable Zirconium, emitting an electron with a maximum energy of 2.3 MeV.

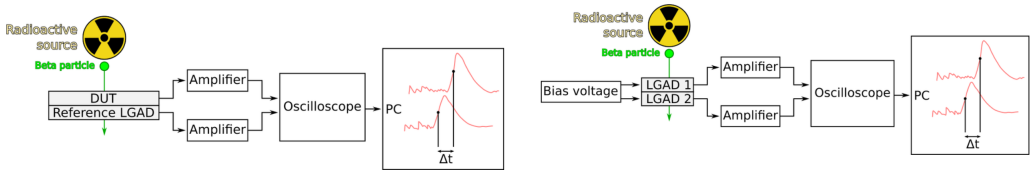


Figure 5.10: Two different configurations for the radioactive source setup, as used in this work.

Using two identical sensors does not require any prior knowledge of the device timing properties, while in the case of one device under study (DUT) and one reference sensor, the reference sensor features have to be well known and studied. From an implementation point of view, both configurations are identical, as shown in figure 5.11.

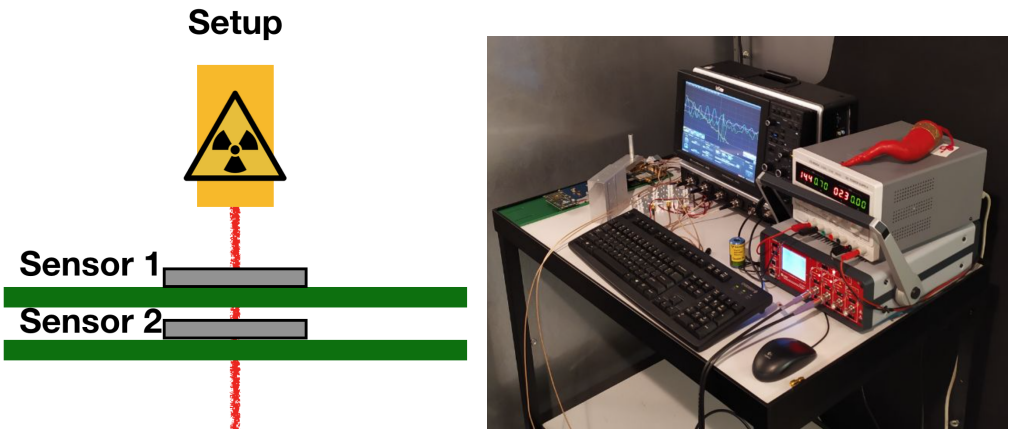


Figure 5.11: Left: Radioactive-source setup schematic. Right: Photo of the Radioactive-source setup used for studying the sensors in the UZH laboratory.

5.11 Laser TCT system

The TCT or Transient Charge Technique is a powerful approach for studying sensor characteristics. It is essentially composed of two parts: a pulsed laser and a movable chuck. In the course of this work, a large scanning TCT system from Particulars [66] has been used, as shown in figure 5.12.

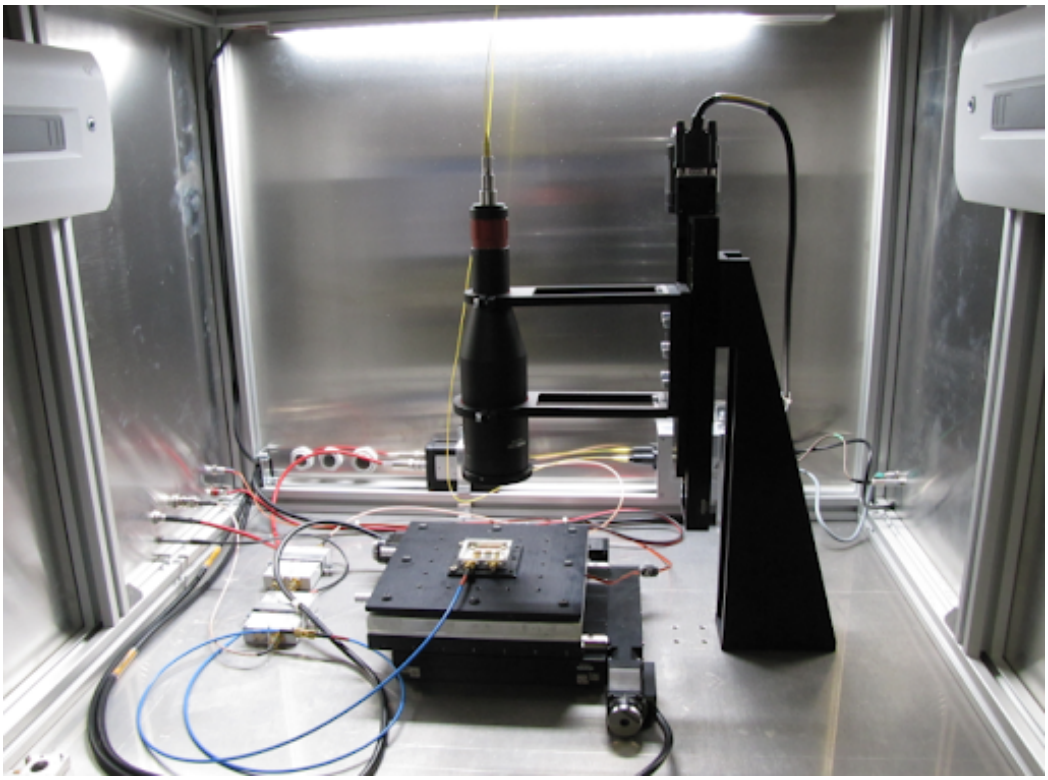


Figure 5.12: Large Scanning TCT system from Particulars.

The apparatus features an infrared laser with tunable power, with a beam-spot

smaller than 10 μm and a motorised chuck with a minimum step size of 2 μm. Using this setup, it is possible to shine the laser pulse on the sensor with high precision and study its response. The laser power can be tuned, for example to emulate the energy release of a MIP, and to reconstruct the charge collection efficiency map with high resolution.
The TCT is equipped with a set of Peltier cells, liquid cooling, and dry-air hose, so that low-temperature tests for irradiated sensors are possible.
The laser stops in one sensor, so it is not possible to use the same configuration as for the radioactive source. To study the time resolution of a device, a time reference is constructed by implementing a couple of beam splitters. The laser beam is split in two paths with an intensity ratio of 50% to 50%, and one of the two lines delayed by 50 ns. The two signals are then sent to the same DUT. The splitting of the signal is necessary only for the time-resolution studies, while for the measurements of all the other sensor characteristics, such as charge collection or position resolution, the standard setup can be utilized.
Before using the TCT system, it is important to find the laser focus. Finding the laser focus minimises the size of the beam spot and allows for the maximum accuracy while scanning the sensor. The setup is equipped with a beam spreader, which is used to focus the laser in the z-direction, perpendicular to the sensor's surface, using the following approach.
The laser is mechanically moved to a point at the border of the sensor until the signal decreases to half of the original value. This indicates that almost half of the laser beam-spot is outside of the active area if the sensor. A scan in the z direction is then performed. When the z position is closer to the focus position, the beam-spot is smaller, and most of it is hitting the sensor. On the other hand, when we are far away from the focal plane, the laser beam-spot is wide and a large fraction of it is not hitting the sensor. The z-scan can be represented as a plot containing the signal amplitude vs the z-position, with a typical parabola shape. The focal plane position is the parabola maximum. A scan is usually performed in multiple steps, starting from a coarse pitch to a finer pitch to find the precise focal plane position.
When the optimal position is found, the laser beam-spot size is very small with a typical Gaussian profile and a σ of 2 μm.
An example of the data extracted using the TCT setup is shown in figure 5.13. In this case, an AC-LGAD sensor from the RSD-1 (Resistive AC-Coupled Silicon Detectors) run has been used. The sensor is placed on the chuck, the laser's pulse is sent 50 times for each spatial point, and the signal charge as a function of the position is registered on the four metal pads connected to the readout electronics. Starting from this data it is possible to investigate the charge sharing, the spatial resolution, and the charge collection efficiency of the sensor.

5.12 Analysis of the timing performance

The standard radioactive-source setup, using two identical sensors, has been described in the previous section. In this section, the time resolution extraction will be discussed.
The timing resolution of the LGAD sensors is of the order of tens of picoseconds. The LGAD signal characteristics and the time of arrival reconstruction techniques were studied

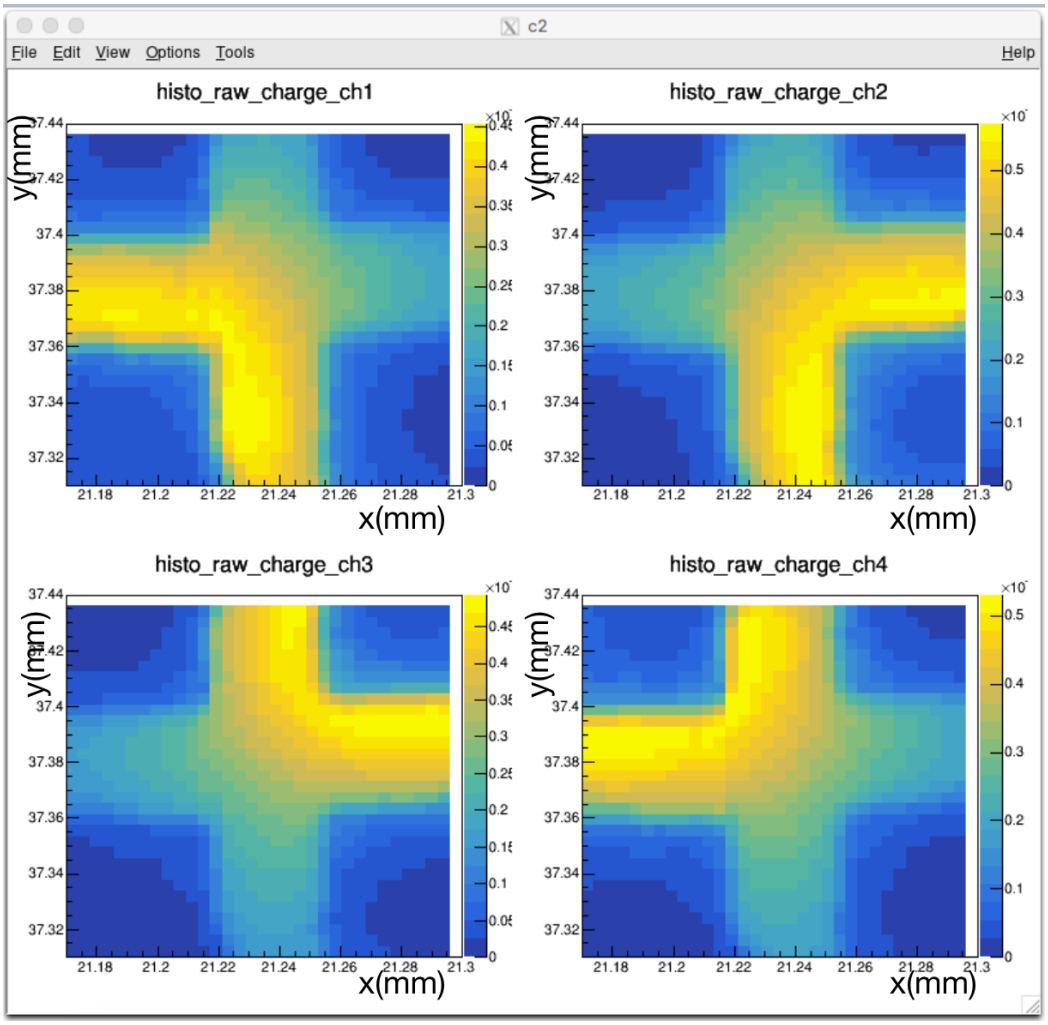


Figure 5.13: Example of the charge collected in four adjoining metal pads in an AC-LGAD sensor, obtained with a laser scan using a TCT system using an RSD-1 sensor from FBK. The x and y axes are the two spatial coordinates, while the colour map is the charge collected expressed in arbitrary units.

in depth before attempting to reconstruct the time resolution. In particular, two parameters are important for obtaining an optimal time resolution: the rise time and the signal over noise (S/N) ratio. The rise time is defined as the time the signal takes to rise from 10% to 90% of its amplitude. In the case of LGADs, this time is of the order of 600-800 ps. The S/N is the ratio between the signal amplitude and the noise, and depending on the gain-layer characteristics, can vary from 20 to 40. It depends also on the energy released by the particle, so in the case of the Strontium source, the S/N distribution will have the shape of its energy deposit distribution.

The charge collected from the setup, which is proportional to the energy deposit inside the sensor, is a convolution of a Gaussian distribution given by the charge resolution of the readout electronics and a Landau distribution, which depends on the physics behind the energy deposit inside the sensor. The time resolution of the setup can be subdivided in

different contributions [67]:

$$\sigma = \sigma_{Landau} \bigoplus \sigma_{Timewalk} \bigoplus \sigma_{Jitter} \bigoplus \sigma_{digi} \tag{5.3}$$

The Landau contribution, as already stated, depends on the fluctuation of the amplitude of the signal, due to the statistical nature of the carrier generation by the primary particle along its path. It can be reduced by using thinner sensors. The time-walk and jitter contributions depend on the sensor and readout electronics characteristics, in particular on the amplifier response. The two contributions are depicted in figure 5.14. The time-walk effect happens

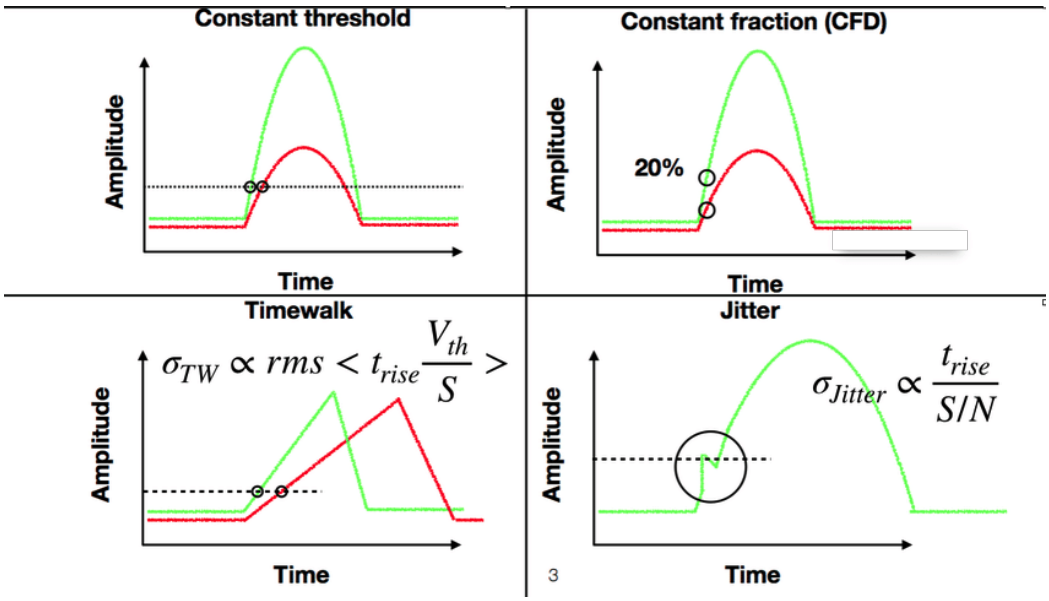


Figure 5.14: Top side: depiction of the constant-threshold and constant-fraction methods to determine the time of arrival. Bottom side: sketch and formula of the time-walk and jitter contributions to the time resolution.

when signals with different amplitudes are amplified. The amplifier response depends on the signal amplitude, and a signal with large amplitude rises faster than a smaller signal. The signals generated in two sensors often have a different amplitude because of the Landau-like energy deposit, therefore it is important to correct for this effect before computing the time resolution of the sensors. There are different ways to mitigate this noise contribution. The first one is to use sensors with short rise time and high gain, to deliver a fast, large signal. The analytical dependency of the time-walk contribution to the rise time and signal amplitude is:

$$\sigma_{Timewalk} =< t_{rise} \frac{V_{thr}}{S} >, \tag{5.4}$$

where V_{thr} is the voltage threshold used on the signal. The second way of minimizing the time-walk effect is during signal analysis. In this case, we have the freedom to define the time of arrival of the signal. The easier but less effective way is to use a constant threshold. A more efficient method consists in adopting a constant fraction discriminator or CFD. Instead of setting a fixed threshold and computing the corresponding time of arrival, in this case the threshold is a fixed percentage of the signal's amplitude, for example 20%. When

the signal crosses the CFD threshold, then the time of arrival is computed. In the case of a larger signal, the time of arrival is delayed, and this compensates for the faster rising edge, while for smaller signal the time of arrival arrives earlier, correcting for the smaller slope.

The second contribution is the jitter contribution, which has an analytical formula of:

$$\sigma_{Jitter} =< \frac{t_{rise}}{S/N} > \tag{5.5}$$

In this case, the problem arises when a noisy signal crosses the threshold. The noise induces an uncertainty on the exact level of the signal, and therefore on the time at which the signal crosses the threshold. To avoid this problem, it is important to have a large S/N ratio, so to use sensors with high gain, and to minimise the rise time. Furthermore the amplifiers can induce distortion, noise, and pick-up noise from different sources, like Wi-Fi and similar. The last contribution is the digitiser contribution, arising from the oscilloscope while sampling the signal and translating it into a digital format. In our particular case, this contribution is quite small because we use a fast oscilloscope with a 40 Giga-Samples per second rate. This term can be estimated to be of the order of 3 ps, one order of magnitude smaller than the time resolution that we expect to obtain from the sensors.

5.13 Performance of AIDA-2020 CNM LGADs

The AIDA-2020 run is an LGAD production performed by the CNM foundry, in the framework of the AIDA-2020 European project [68]. These sensors come with a fixed bulk-doping concentration and with a variety of doping concentrations for the multiplication layer, which is implanted with boron. The sensors have been produced with two different thicknesses, the standard 50 μm thickness and a thinner 35 μm version. The smaller thickness is more resilient to radiation damage, and results in sensors with very good timing performance. In this work the performance of these devices has been investigated, using three pairs of sensors from the first run, with different doping and thickness. The timing performance as a function of the CFD cut is summarised in figure 5.15. The sensors of this study are a pair of 50 μm thick sensors with low doping concentration in the gain layer, and two pairs of 35 μm thick sensors with low and medium doping concentrations in the gain layer. The doping concentrations are confidential and the foundry does not disclose the concentrations. The sensors have been tested using the radioactive-source setup and the time resolutions for different CFD cuts have been extracted for different bias voltages.
Looking at the time resolution as a function of the CFD cut, a typical trend is clear. For a low CFD cut, the signal is in the noise bandwidth, and the time resolution is degraded. For a high CFD cut, the time-walk effect starts to be relevant, so that the time resolution worsens. In general, the optimal CFD range for extracting the time resolution is between 20% and 80%.
All the sensors perform well, and with the 35 μm thick sensors, medium gain, and an outstanding time resolution of 27 ps is obtained. The measurements have been repeated one year later at room temperature and in a cold environment at -20° Celsius, and these sensors have been used as reference sensor for the study of the next iteration of the AIDA-2020 run.

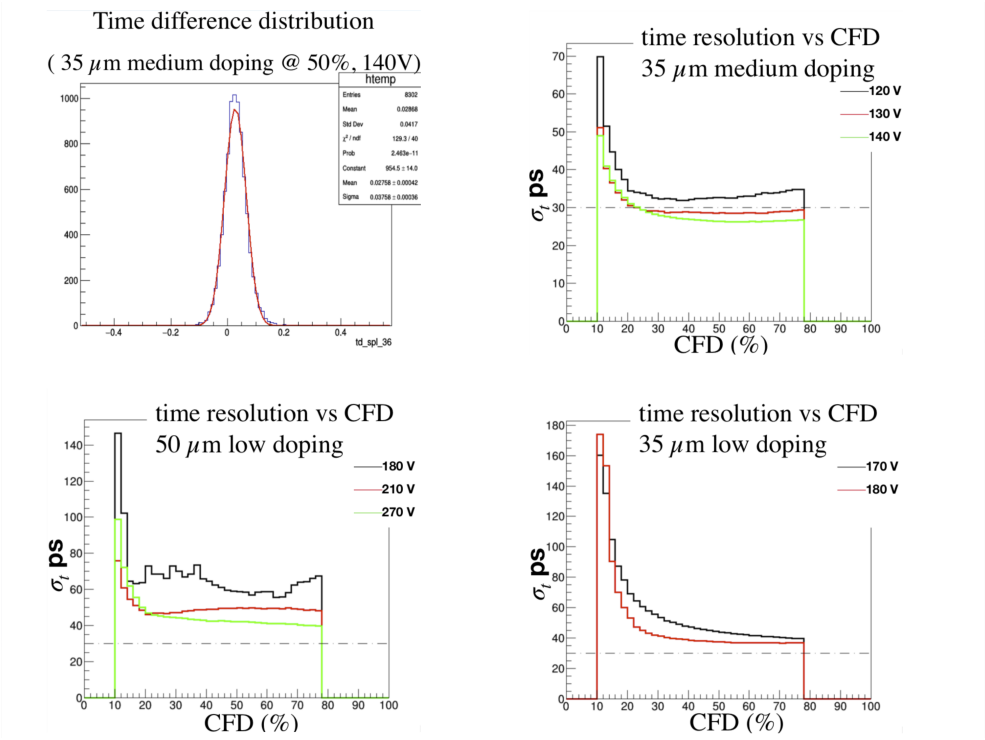


Figure 5.15: Timing performance for three pairs of sensors from the first AIDA2020 run produced from CNM. The red dotted line marks the 30 ps time resolution which is the target performance for this study.

The time resolution as a function of the CFD cut is shown in figure 5.16.

Figure 5.16: Timing performance at -20°C for the AIDA 2020 W5 sensors. These sensors have then been used as reference sensors to study the timing performance of irradiated devices.

5.14 Performance of LGAD sensors from the HPK2 production

As previously mentioned, the sensors from the first AIDA-2020 run had a problem with the JTE termination, inducing a high leakage current. Some months later, we received sensors from a new run in which the leakage current problem was purportedly solved, and these sensors were tested before and after irradiation, using as reference sensor the 35 μm sensor, with medium doping, from the first run.
Irradiated sensors from the LGAD run processed in 2020 at HPK (Japan) as prototypes for the MTD, and indicated here as HPK2 production, have been characterised both with measurements of the IV curves in the probe station and of the timing properties in the radioactive-source setup. In both cases, the sensors have been tested in a cold environment, at -20°C, to limit the leakage current and reproduce the operational environment. These sensors have an active thickness of 45 μm, and differ from the characteristics of the CNM and FBK productions mainly due to the deeper multiplication layer. They have been irradiated to a dose of $4 \times 10^{14} n_{eq}/cm^2$, $8 \times 10^{14} n_{eq}/cm^2$, $15 \times 10^{14} n_{eq}/cm^2$, and $25 \times 10^{14} n_{eq}/cm^2$. The first two doses are equivalent to a quarter and to a half of its expected total fluence that would be seen by the MTD during the operational lifetime.
The IV curves for these devices are shown in figure 5.17. It can be seen that the breakdown voltage increases with fluence and that the general behaviour of the LGAD leakage current becomes more similar to that of the PIN diode, as the acceptors in the multiplication layer are deactivated by the radiation damage. The timing performance of the sensors

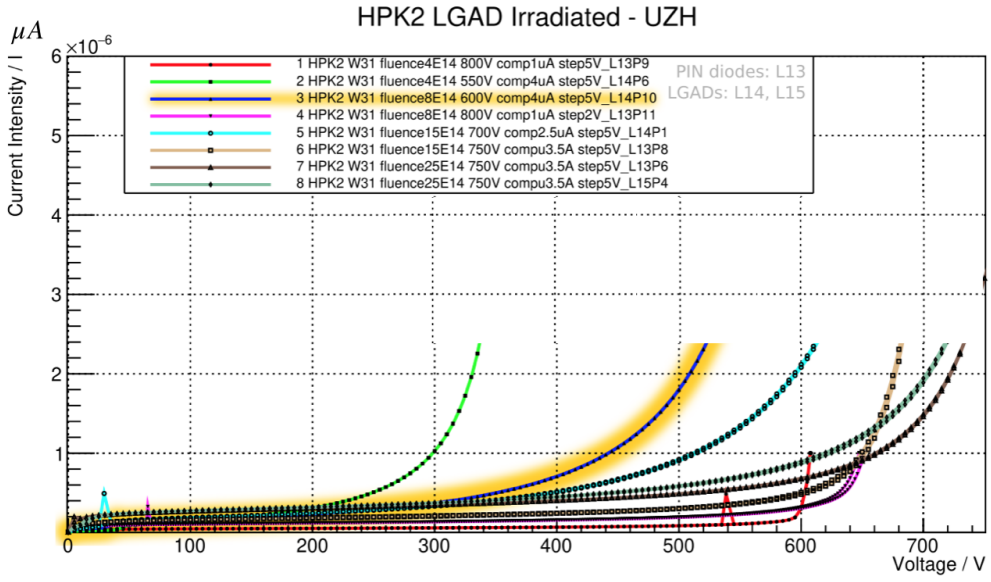


Figure 5.17: IV curves for HPK2 sensors after irradiation and annealing. The sensors have been exposed to four different radiation doses, $4 \times 10^{14} n_{eq}/cm^2$, $8 \times 10^{14} n_{eq}/cm^2$, $15 \times 10^{14} n_{eq}/cm^2$, and $25 \times 10^{14} n_{eq}/cm^2$.

after irradiation and annealing at room temperature is shown in figure 5.19. All the sensors

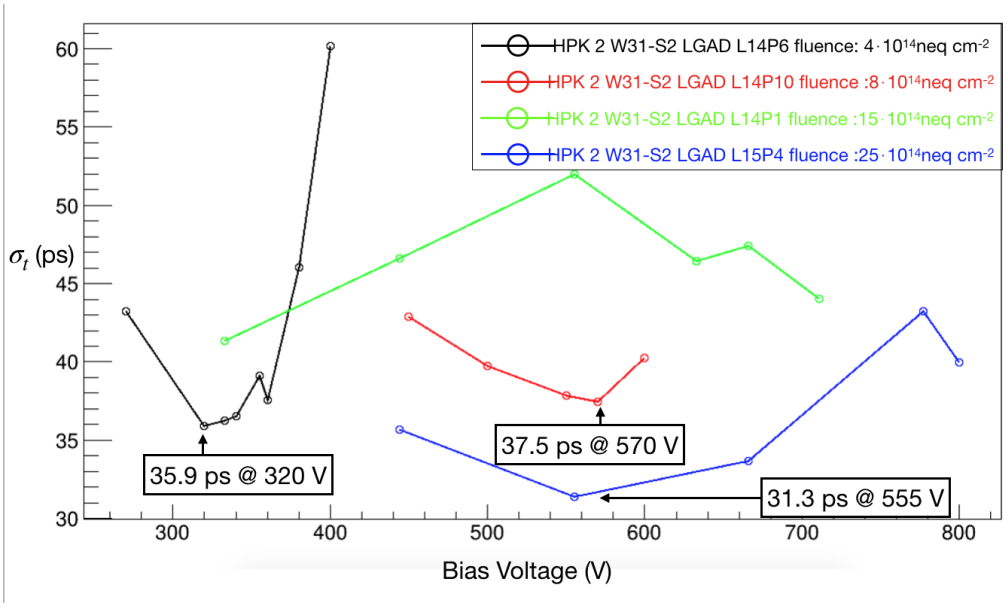


Figure 5.18: Timing resolution as a function of the bias voltage for sensors at -20°C irradiated at different fluences. All the sensors achieved timing resolutions under 40 ps except for one, which was discovered to be defective after testing.

achieved timing resolutions under 40 ps, except for one, which was discovered to be defective after testing. The resolution has been extracted by subtracting the time resolution of the reference device by the time resolution of the system following the formula:

$$\sigma_{system} = \sqrt{\sigma^2_{DUT} + \sigma^2_{ref}}, \tag{5.6}$$

where σ^2_{ref} is obtained before using two identical sensors. Then one of these sensors, after characterisation, is used as reference sensor in this setup. The performance is very impressive after irradiation and encouraging for the expectations of the performance of the MTD detector during its operational life.

5.15 Performance of the RSD1 AC-LGAD production

As mentioned in the previous chapters, a common problem of the LGAD technology is the fill factor for small pitch sensors. Each pad is isolated from the neighbouring one by the JTE structure and the P-stop, resulting in a minimum no-gain width of around 40 μm. This is a show-stopper for the creation of a pixelated timing sensor. To achieve real 4D tracking (position plus timing resolution), the LGAD technology has to be modified. One possible solution is represented by the AC-LGAD technology, developed by INFN Torino in collaboration with FBK [61]. This new technology introduces new challenges in the time and position reconstruction. The signal coming from a single charged particle is now spread out from the resistive layer onto multiple pads. The details of the charge sharing depend on the readout pad size, shape and pitch, and on the resistive layer thickness and doping concentration. The signal sharing is both beneficial and detrimental. Since the signal is spread on different pads, a smaller charge is collected by each of them. A sufficient signal is

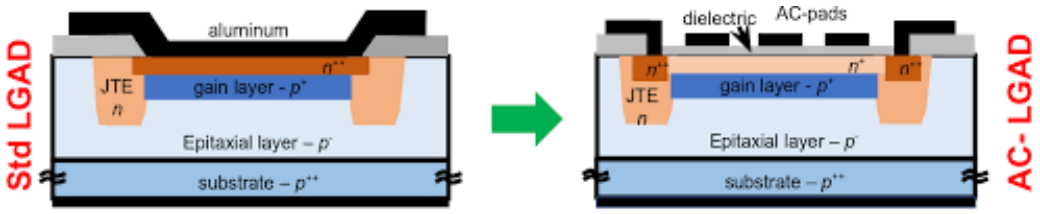


Figure 5.19: Left: standard LGAD structure. Right: AC-LGAD structure. The AC-LGAD structure is similar to the standard LGAD, with the addition of a resistive n^+ layer on top, and the segmentation of the readout pads, capacitively coupled to the silicon through a dielectric layer.

anyhow generated before irradiation, thanks to the multiplication layer, thus allowing both a good spatial resolution thanks to charge sharing, and with its small thickness, helping to maintain a low material budget.
To optimize the charge sharing level, sensors with different doping concentrations have been produced. The available flavours are shown in figure 5.20 [69]. Furthermore, to achieve an

wafer	*n*-plus dose	*p*-gain dose	dielectric thickness	*p*-stop dose	substrate
1	A	0.92	L	B	Si-Si
2	A	0.94	L	A	Si-Si
3	A	0.94	L	B	Epi
4	A	0.94	H	B	Si-Si
5	A	0.96	H	B	Si-Si
6	B	0.92	L	B	Epi
7	B	0.94	L	A	Si-Si
8	B	0.94	L	B	Si-Si
9	B	0.96	L	B	Si-Si
10	B	0.96	H	B	Si-Si
11	C	0.92	L	B	Si-Si
12	C	0.94	L	B	Epi
13	C	0.94	L	B	Si-Si
14	C	0.96	H	B	Epi
15	C	0.96	H	C	Si-Si

Figure 5.20: Wafers with different process splits for the AC-LGAD RSD-1 production [69]. The value for the different doses is not disclosed; there are three different *n*-plus doses, namely A, B, and C, and two different dielectric doses, high (H) and low (L).

optimal time resolution, the signals have to be recombined by taking into account the time delay induced by the signal traveling in the resistive layer. It is yet to be proven if the recombination can be done offline with just the TOA of each single pad in the cluster. On

the other hand, because the signal is spread on multiple pads even for perpendicular tracks, it is possible to use the center-of-charge to reconstruct the particle position, improving the spatial resolution for perpendicular tracks. The signal splits on different pads, with a sharing that depends strongly on the position in which the signal is created on the sensor. The TCT setup was used to study the AC-LGAD sensors, given the possibility with this system of knowing the position where the laser was impinging and the initial charge created.

5.16 Charge sharing and hit position resolution

The first step in characterising the AC-LGAD sensors is to study the signal and how it splits among different pads. The setup used is comprised of a TCT with a UV laser and a motorised chuck. The sensor is wire-bonded to a passive Torino readout board and four pads are read out by the oscilloscope. The signal is amplified using fast amplifiers before reaching the oscilloscope. Figure 5.21 shows the AC-LGAD sensor wire-bonded on the

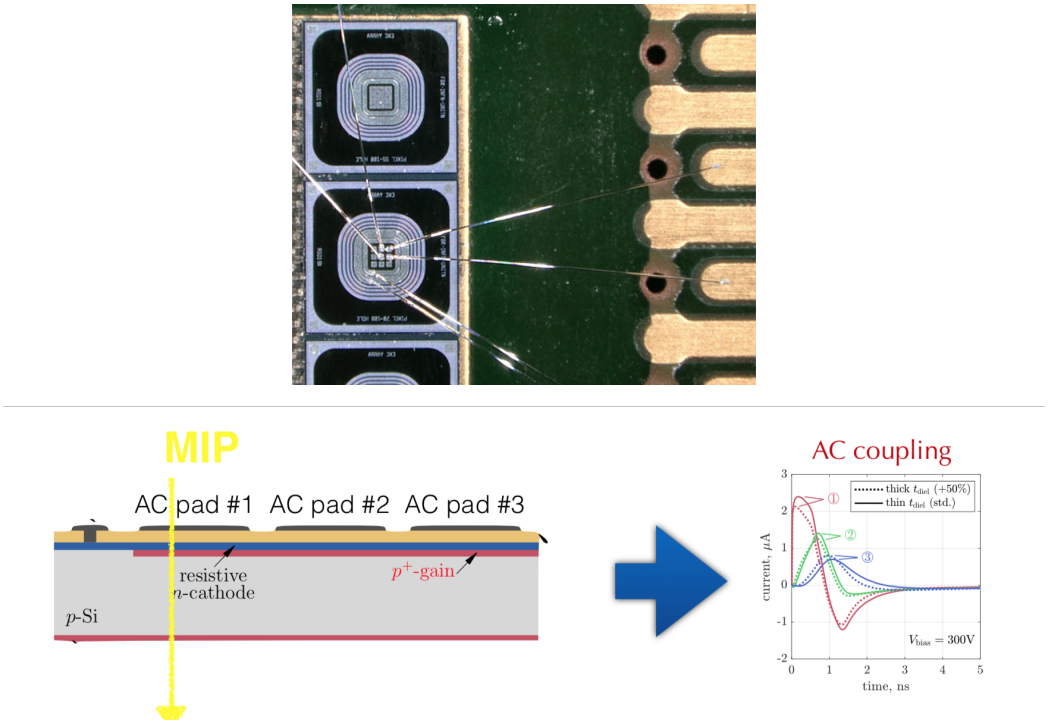


Figure 5.21: Left: AC-LGAD structure wire-bonded on top of the readout board. Right: signal spreading on different pads.

readout board on the left, and the signal collected from multiple pads for a passing MIP on the right. As mentioned before, the signal is shared by different pads. The signal shape is bi-polar, meaning that after the usual signal a slower under-shoot can be observed. This is due to the presence of the resistive layer that operates as a transmission line. It is important to understand how the signal spreads on different pads and how it is shared as a function of the position in which it is generated. Using this information it is possible to implement the right reconstruction algorithm for position reconstruction, which will be used also for timing studies. The easiest way to study the signal diffusion is to use the TCT, firing the

laser on the sensor in different positions, and plotting the charge collected by the pads as a function of the position. The first results obtained for a structure with a pitch of 200 μm and a metal pad size of 150 μm are shown in figure 5.22. On the left side of the figure, the

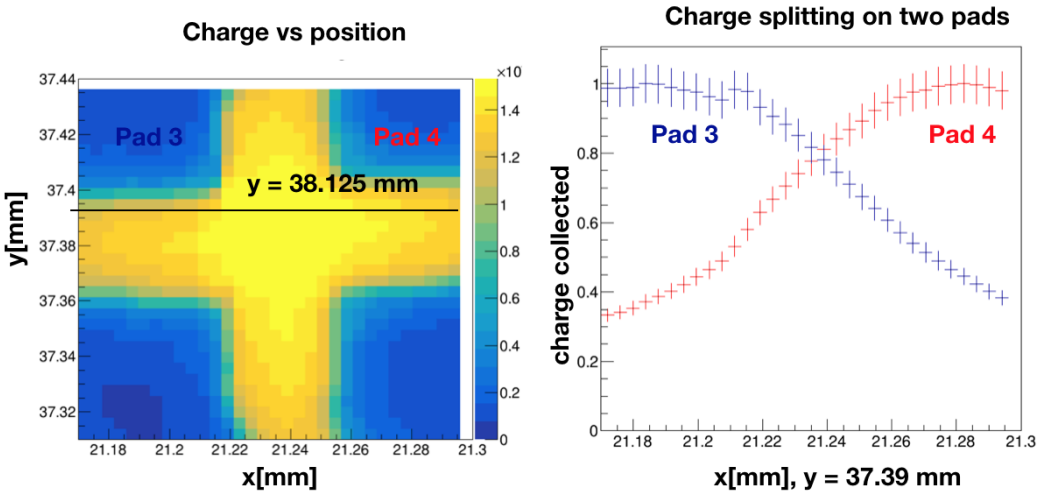


Figure 5.22: Left: AC-LGAD scan with the TCT. The plot shows the signal intensity as a function of the position for different pads. Right: relative signal charge collected on two neighbouring pads as a function of the laser position.

plot shows the signal charge as a function of the position for different pads using the charge collected from all the pads. On the right, the charge collected from Pad 3 and Pad 4 have been plotted as a function of x for a fixed value of y 5.22. The plot shows how the charge is collected depending on the hit position. If the hit is inside the pad or on its boundaries, 100% of the charge is collected by a single pad. If the hit is in between two pads, the charge is split between two pads according to a linear relation, which is good news for hit reconstruction, meaning that the normal charge-centroid-method works on these structures. Before applying this method, it is important to compare the charge collected from different amplifiers, to avoid the situation where a difference in the amplifier gain could degrade the final position resolution. Figure 5.23 shows the charge distribution collected from the four different pad's amplifiers. The last peak on the right, indicated by a red arrow in the four plots, is assumed to be the maximum charge collected by each pad, smeared by the charge resolution. The different position of these peaks in the four plots originates from the different amplifier gains. The maximum peak values are used to normalise the charge distribution. This procedure is no different from the one performed with the standard pixel sensors, where the readout of the single pixel cells are calibrated. After normalizing the charge, the estimated position is obtained as:

$$\begin{aligned} x_i &= \frac{\sum_i x_i \cdot ch_i}{\sum_i ch_i}, \text{and} \\ y_i &= \frac{\sum_i y_i \cdot ch_i}{\sum_i ch_i}, \end{aligned} \tag{5.7}$$

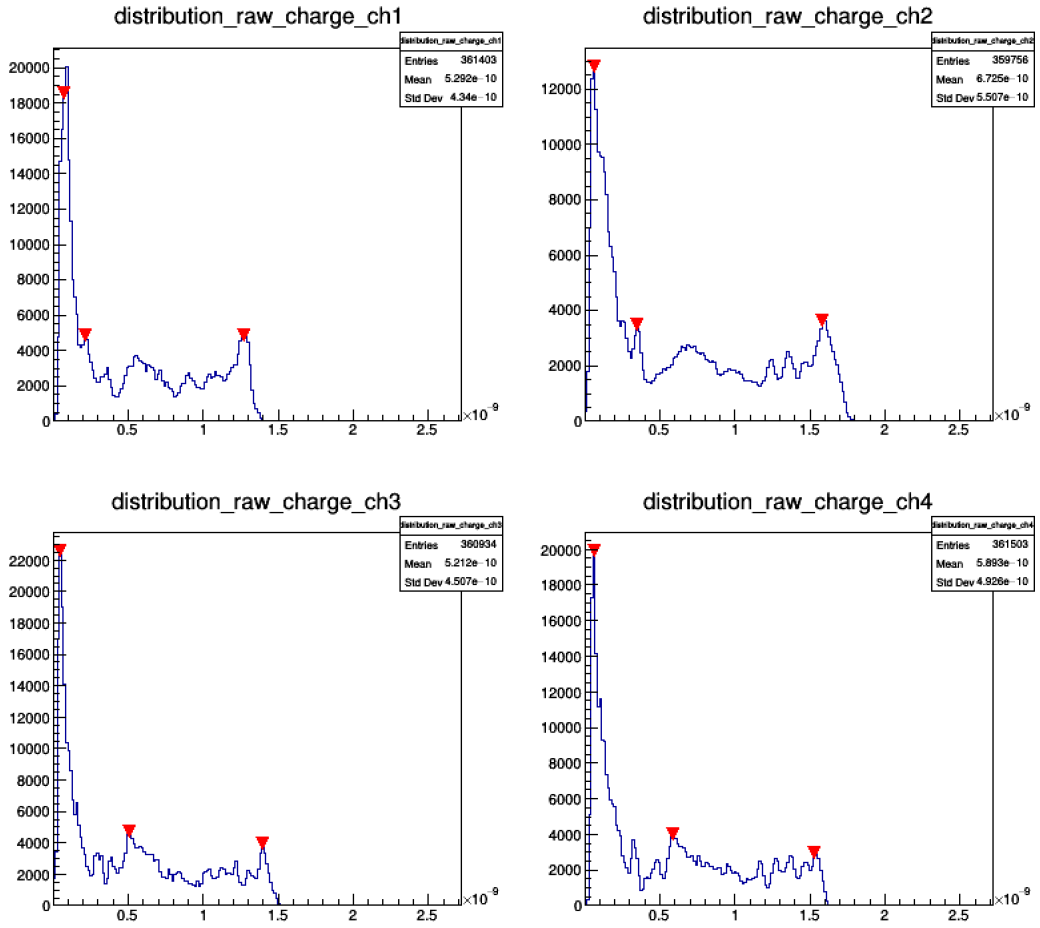


Figure 5.23: Normalised charge collection distributions for the four different readout pads and amplifiers. The red arrows indicate peaks in the distributions.

where x_i and y_i are the central position of the ith pad, and ch_i is the collected charge from the pad. The preliminary results for the position resolution are shown in figure 5.24. In this case, the true position is given by the TCT, while the reconstructed position is obtained by the charge-centroid method. The reconstructed position resolution is of the order of 3 μm, comparable to techniques used in test-beam studies [70]. The resolution is obtained using all the area of the scan, including the pad's reflective region in which most of the light from the laser is reflected. Using only the exposed area of the sensor, the resolution improves to 2 μm.

The spatial resolution obtained from the detector can be improved by reducing the pad size. As shown before, the charge sharing depends on the position in which the particle is hitting the detector, and if the hit is inside a single pad, most of the charge is collected by it. If the hit is in between different pads, the charge is collected by up to four different pads. This is not a problem in our setup where we can read out the charge with extreme precision and low threshold, but in a real detector, in which the charge readout buffer size is limited, it can pose a serious limitation.

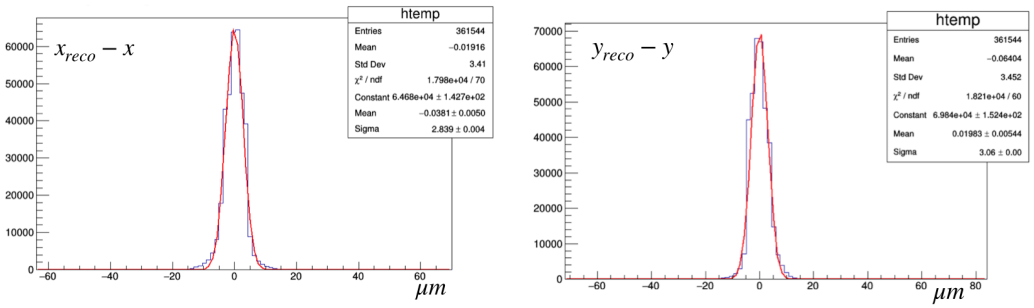


Figure 5.24: Position resolution distribution along the x and y directions for the AC-LGAD sensor UFSD-1 from FBK. The resolution is of the order of 3 μm.

5.17 Timing studies

To study the timing performance of this sensor, the TCT setup has been used with some modifications. The laser signal has been split using a beam-splitter. One of the signals is sent to the laser directly, while the second signal is delayed using a long optical fibre, and then sent to the laser. The system is optimised to minimise attenuation, and the two signals have similar amplitudes and shapes. Using this setup, it is possible to extract the time resolution of the sensor using just one sensor. Furthermore a precise scan of the sensor surface is possible. The signal analysis techniques are the same as used for the strontium setup, and the strontium source has been used to tune the laser power to create a number of electrons close to the one that would be created by a MIP when hitting the sensor. The results are shown in figure 5.25. The timing resolution is extracted reading out a single pad. The sensor shows a good timing resolution compatible with the timing resolution of a standard LGAD. Close to the readout pad the time resolution is 11 ps, degrading as the signal is shined further away. Combining the information from the other pads it would be possible to mitigate this effect and maintain an excellent time resolution on the whole surface of the sensor.

5.18 Towards the future

The ultimate goal for timing sensors is for them to replace the standard silicon sensors in HEP experiments. A mid-term more realistic goal for CMS is to extend the pseudorapidity coverage provided by the MTD timing layer from $\eta < 3.0$ to $\eta < 4.0$, the full fiducial tracking volume. Two cases are under evaluation: the replacement of existing TEPX disks or the addition of new timing layers in between the existing TEPX disks. Because of the exceptional timing and spatial performance of the AC-LGAD, this type of sensor is the natural candidate to replace disks inside the TEPX detector. This option would be the preferred one, because a new additional disk would require its own services, including power and data connection in an already crowded space.
The challenges to be addressed are:

- occupancy of the sensor,

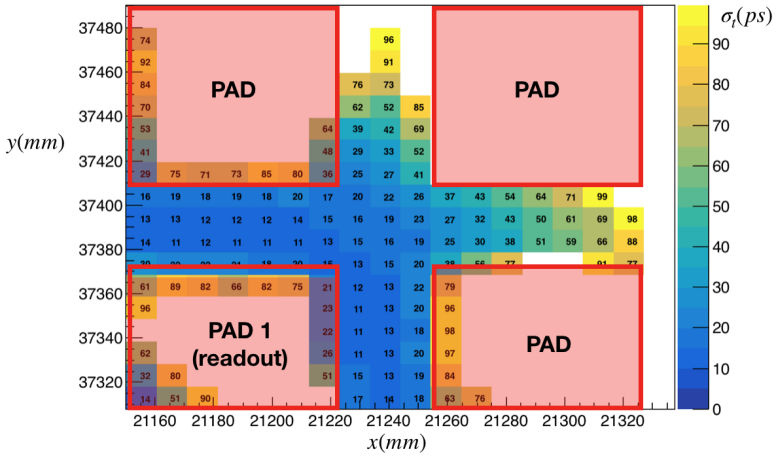


Figure 5.25: timing resolution of an AC-LGAD from the RSD1 production. The plot shows the timing resolution as a function of the position, and is obtained using a TCT laser setup. The laser power is tuned to deposit the same energy as a MIP in the sensor. The timing resolution quoted is obtained by reading the signal on a single pad, located in the bottom-right region. The sensor has a pitch of 100 μm and a pad size of 50 μm.

- radiation hardness,
- fast and low-power high granularity readout.

The expected occupancy in the TEPX is of the order of 10^{-4}, as shown in figure 5.26 [52]. It is calculated taking into account the nominal pixel area of 2500 μm^2.

The LGAD is a relatively recent technology, but quite well understood. On the other hand, the AC-LGAD is a five-year-old technology, and there are many efforts to optimise the performance of this new sensor family. R&D activities are currently carried out in the following areas:

- pad size and shape optimisation,
- radiation hardness improvement, and
- fast and low-power high granularity readout.

In the currently available AC-LGAD productions, the pad shape is a square. The only difference between different sensor pads is their size and their pitch, where the ratio between the pad area and the sensor area defines a sort of "fill factor" for the sensor. Not to confuse this figure with the usual definition of fill factor, it will be indicated from now on as "coverage factor". Sensors with large pads and high coverage factor show better timing performance compared to sensors with smaller pads and lower coverage. On the other hand, because the charge collection under the pad is flat, the resolution becomes binary when a particle is passing under a pad. This leads to the conclusion that to improve spatial resolution a low coverage factor is desirable. All these considerations are true for square pads, but it is possible that different pad shapes, as for example in the form of a cross, could be a better

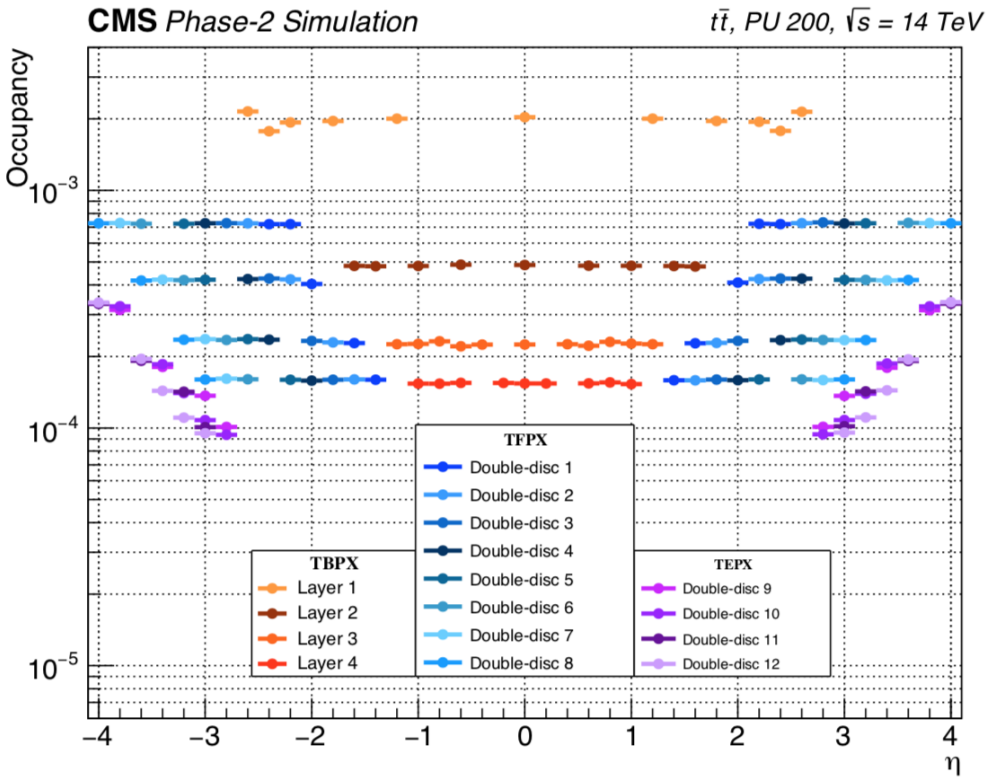


Figure 5.26: Hit occupancy, defined as the fraction of channels containing a digitized hit, as a function of η for all layers and double-discs of the Inner Tracker [52].

compromise to optimise timing and spatial resolution at the same time.
The radiation hardness of AC-LGADs has still to be investigated. These sensors feature an n-type resistive layer which behaves like a transmission line (plane in this case). This layer could be quite susceptible to radiation and dedicated studies are necessary to prove the radiation resistance of these kind of sensors.
At the moment, there are no available large-size readout chips with the timing requirements necessary to achieve 4D tracking.
To achieve a time resolution under 50 ps, there are a number of interesting challenges to overcome, like the:

- distribution of the clock signals,
- pixel unit (PU) footprint constraints,
- power cooling constraints.

The clock distribution to different modules and within a single module is challenging. The clock distribution to different sensors over a wide area with a jitter smaller than 30 ps is a problem that has been faced by the MTD project and seems solvable. The clock distribution within modules to the different TDCs is more difficult, due to the small area available for each pixel cell. In the present ETROC [71] used in the MTD end-cap region, the area available for each "pixel" is 1.3 $\times$ 1.3 mm^2, while in the projects in which the AC-LGAD

technology could be implemented, the available area is 10 to 100 times smaller. The wide space available in the ETROC design allows for the implementation of a capacitor array that is used to smooth and stabilise the clock and power distributions. This area would not be available for a smaller pixel pitch so a different approach will have to be used.
Nevertheless, the recent development of the Timespot1 chip is promising [72]. This readout chip is implemented in a 28-nm CMOS technology and features a 20 ps time resolution, with a pixel size of 50$\times$ 55 μm^2. Furthermore the nominal power consumption is of the order of 20 μW per channel, compared to the 10 μW per channel for the standard CMS phase-2 readout chip. This chip is quite promising, also in view of the current timeline, which will allow 10 years of development and construction before a timing disk would be inserted in the TEPX, around 2032.

Chapter 6

Conclusions

During my PhD I worked on three main tasks: the construction of the CMS phase-1 barrel pixel detector supply tubes, a simulation to extract the spatial resolution of the pixel sensor for the CMS phase-2 pixel detector upgrade, and the performance of the LGAD and AC-LGAD sensors.

The construction of the supply tubes, and their testing, commissioning, and merging was a success. The schedule was tight, but the detector performed as expected, with the exclusion of the DC-DC converter crisis period, which was mitigated and eventually solved. Multiple analyses benefited from the improved vertex resolution, and in a more general sense, all the analyses were able to benefit from the additional tracking layers and improved readout speed to improve efficiency with respect to fake rate, as well as track resolution parameters. Since the first insertion, the detector has been extracted and inserted multiple times for oridnary and extraordinary maintenance and the installation and commissioning procedure has proven to be reliable, while the detector system has proven its solid performance during its years of operation.
I simulated and predicted the the phase-2 sensor's resolution performance years before test beam results were available. The data obtained from the test beam confirm my predictions. Resolution as a function of the deep angle obtained from my simulation matches closely the result from the test beam, predicting the resolution values and the position of the minimum. The results have been used, and are still used today, inside the tkLayout software to optimise the performance of the future pixel upgrade for CMS.
I contributed to the understanding of the LGAD technology, collaborating with different groups and testing sensors from different vendors. I have characterised the electrical properties and the spatial and time resolution of those sensors, before and after irradiation. Thanks to my contribution UZH CMS group now is a solid player inside the timing community and RD50 collaboration. Thanks to my contribution, the UZH and PSI CMS groups are considering to replace some of the disks inside the future pixel detector upgrade with timing disks, using the AC-LGAD technology.

List of Figures

List of Tables

Bibliography

[1] Lhc accellarators website. https://home.cern/science/accelerators/large-hadron-collider, .

[2] Rf cavities lhc. https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.rf_cavities.

[3] Rf vacuum lhc. https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.high_vacuum, .

[4] V. Halyo, A. Hunt, P. Jindal, P. LeGresley, and P. Lujan. GPU Enhancement of the Trigger to Extend Physics Reach at the LHC. *JINST*, 8:P10005, 2013. doi: 10.1088/1748-0221/8/10/P10005.

[5] W. Herr and B. Muratori. Concept of luminosity. 2006. URL http://cds.cern.ch/record/941318.

[6] Tai Sakuma. Cutaway diagrams of CMS detector. May 2019. URL https://cds.cern.ch/record/2665537.

[7] URL https://wiki.physik.uzh.ch/cms/latex:example_spherical_coordinates.

[8] CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *Journal of Instrumentation*, 9(10):P10009, 2014. URL http://stacks.iop.org/1748-0221/9/i=10/a=P10009.

[9] V. Veszpremi. Performance verification of the cms phase-1 upgrade pixel detector. *Journal of Instrumentation*, 12(12):C12010?C12010, Dec 2017. ISSN 1748-0221. doi: 10.1088/1748-0221/12/12/c12010. URL http://dx.doi.org/10.1088/1748-0221/12/12/C12010.

[10] CMS Collaboration. *CMS Physics: Technical Design Report Volume 1: detector performance and software*. Technical Design Report. CERN, 2006. URL https://cds.cern.ch/record/922757.

[11] Cristina Biino. The CMS Electromagnetic Calorimeter: overview, lessons learned during Run 1 and future projections. *J. Phys.: Conf. Ser.*, 587:012001. 12 p, 2015. doi: 10.1088/1742-6596/587/1/012001. URL https://cds.cern.ch/record/2295063.

[12] Petar Adzic and all. Energy Resolution of the Barrel of the CMS Electromagnetic Calorimeter. Technical report, CERN, Geneva, Nov 2006. URL https://cds.cern.ch/record/1009081.

[13] CMS Collaboration. Performance of the CMS hadron calorimeter with cosmic ray muons and LHC beam data. *Journal of Instrumentation*, 5(03):T03012–T03012, mar 2010. doi: 10.1088/1748-0221/5/03/t03012. URL https://doi.org/10.1088/

1748-0221/5/03/t03012.

[14] Silvia López. Cms detector performance. *EPJ Web of Conferences*, 182:02076, 01 2018. doi: 10.1051/epjconf/201818202076.

[15] Jian Wang. Upgrade plans and ageing studies for the CMS muon system in preparation of HL-LHC. Technical report, Sep 2018. URL http://cds.cern.ch/record/2641471. * Temporary entry *.

[16] Muon Identification and Isolation efficiency on full 2016 dataset. Mar 2017. URL https://cds.cern.ch/record/2257968.

[17] Sirunyan. Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s}$ = 13 TeV. *JINST*, 15:P10017. 63 p, Jun 2020. doi: 10.1088/1748-0221/15/10/P10017. URL http://cds.cern.ch/record/2721198. Replaced with the published version. Added the journal reference and the DOI. All the figures and tables can be found at http://cms-results.web.cern.ch/cms-results/public-results/publications/TRG-17-001 (CMS Public Pages).

[18] A.M. Sirunyan, A. Tumasyan, W. Adam, E. Asilar, T. Bergauer, J. Brandstetter, E. Brondolin, M. Dragicevic, J. Erö, M. Flechl, and et al. Particle-flow reconstruction and global event description with the cms detector. *Journal of Instrumentation*, 12(10): P10003–P10003, Oct 2017. ISSN 1748-0221. doi: 10.1088/1748-0221/12/10/p10003. URL http://dx.doi.org/10.1088/1748-0221/12/10/P10003.

[19] Leonardo Rossi, Peter Fischer, Tilman Rohe, and Norbert Wermes. *Pixel detectors: from fundamentals to applications*. Particle acceleration and detection. Springer, Berlin, 2006. doi: 10.1007/3-540-28333-1. URL http://cds.cern.ch/record/976471.

[20] Gourab Sabui, Peter Parbrook, Miryam Arredondo, and Z. Shen. Modeling and simulation of bulk gallium nitride power semiconductor devices. *AIP Advances*, 6:055006, 05 2016. doi: 10.1063/1.4948794.

[21] C. Canali, G. Majni, R. Minder, and G. Ottaviani. Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature. *IEEE Transactions on Electron Devices*, 22(11):1045–1047, 1975. doi: 10.1109/T-ED.1975.18267.

[22] P.A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020. doi: 10.1093/ptep/ptaa104.

[23] Stefano Meroli, Daniele Passeri, and Leonello Servoli. Energy loss measurement for charged particles in very thin silicon layers. *Journal of Instrumentation*, 6:P06013, 06 2011. doi: 10.1088/1748-0221/6/06/P06013.

[24] G. Lutz. *Semiconductor radiation detectors*. Springer, 1999.

[25] N Cartiglia, R. Arcidiacono, G-F. Dalla Betta, G. Borghi, M. Boscardin, M. Costa, F. Fausti, F. Ficorella, M. Ferrero, M. Mandurrino, J. Olave, L. Pancheri, F. Siviero, V. Sola, M. Tornago, G. Paternoster, H. Sadrozinski, A. Seiden, and M. C. Vignale. Signal formation and designed optimization of Resistive AC-LGAD (RSD). Feb 2020. URL https://cds.cern.ch/record/2718069.

[26] M. Moll, E. Fretwurst, and G. Lindstrom. Leakage current of hadron irradiated silicon

detectors - material dependence. *Nucl. Instrum. Meth. A*, 426:87–93, 1999. doi: 10.1016/S0168-9002(98)01475-2.

[27] Javier Fernandez-Tejero. *Design and Optimization of Advanced Silicon Strip Detectors for High Energy Physics Experiments*. PhD thesis, 07 2020.

[28] Y. Allkofer, C. Amsler, D. Bortoletto, V. Chiochia, L. Cremaldi, S. Cucciarelli, A. Dorokhov, C. Hrmann, R. Horisberger, D. Kim, and et al. Design and performance of the silicon sensors for the cms barrel pixel detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 584(1):25?41, Jan 2008. ISSN 0168-9002. doi: 10.1016/j.nima.2007.08.151. URL http://dx.doi.org/10.1016/j.nima.2007.08.151.

[29] G Kramberger, V Cindro, I Mandi?, M Miku, M Milovanovi?, and M Zavrtanik. Modeling of electric field in silicon micro-strip detectors irradiated with neutrons and pions. *JINST*, 9:P10016. 19 p, 2014. doi: 10.1088/1748-0221/9/10/P10016. URL https://cds.cern.ch/record/2634242.

[30] Martin Lipinski. The Phase-1 Upgrade of the CMS Pixel Detector. Technical report, CERN, Geneva, May 2017. URL https://cds.cern.ch/record/2265423.

[31] Danek Kotlinski. The design of the CMS pixel detector system. *Nucl. Instrum. Methods Phys. Res., A*, 477:446–50, 2002. doi: 10.1016/S0168-9002(01)01827-7. URL https://cds.cern.ch/record/555503.

[32] A Dominguez, D Abbaneo, K Arndt, N Bacchetta, A Ball, E Bartz, W Bertl, G M Bilei, G Bolla, H W K Cheung, M Chertok, S Costa, N Demaria, Daniel Dominguez Vazquez, K Ecklund, W Erdmann, K Gill, G Hall, K Harder, F Hartmann, R Horisberger, W Johns, H C Kaestli, K Klein, D Kotlinski, S Kwan, M Pesaresi, H Postema, T Rohe, C Schäfer, A Starodumov, S Streuli, A Tricomi, P Tropea, J Troska, F Vasey, and W Zeuner. CMS Technical Design Report for the Pixel Detector Upgrade. Technical report, Sep 2012. URL http://cds.cern.ch/record/1481838. Additional contacts: Jeffrey Spalding, Fermilab, Jeffrey.Spalding@cern.ch Didier Contardo, Universite Claude Bernard-Lyon I, didier.claude.contardo@cern.ch.

[33] Cms twiki public pixel offline plots 2016. https://twiki.cern.ch/twiki/bin/view/CMSPublic/PixelOfflinePlots2016.

[34] The CMS Phase-1 Pixel Detector Upgrade. *JINST*, 16:P02027. 84 p, Dec 2020. doi: 10.1088/1748-0221/16/02/P02027. URL http://cds.cern.ch/record/2748381.

[35] Lutz Werner Feld, Christian Fimmers, Waclaw Karpinski, Katja Klein, Martin Lipinski, Marius Preuten, Max Rauch, David Michael Rittich, Jan Domenik Sammet, and Michael Wlochal. Development of a DC-DC conversion powering scheme for the CMS Phase-1 pixel upgrade. Technical report, CERN, Geneva, Dec 2013. URL https://cds.cern.ch/record/1637045.

[36] Melis Yurddaskal and Buket Baba. The effect of curvature on the impact response of foam-based sandwich composite panels. *Steel and Composite Structures*, 20:983–997, 04 2016. doi: 10.12989/scs.2016.20.5.983.

[37] Katja Klein. The phase-1 upgrade of the cms pixel detector. *Nuclear Instruments and*

Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 845:101–105, 2017. ISSN 0168-9002. doi: https://doi.org/10.1016/j.nima.2016.06.039. URL https://www.sciencedirect.com/science/article/pii/S0168900216305824. Proceedings of the Vienna Conference on Instrumentation 2016.

[38] J. Hoss, H.-C. Kastli, B. Meier, T. Rohe, and A. Starodumov. Radiation tolerance of the readout chip for the phase i upgrade of the CMS pixel detector. *Journal of Instrumentation*, 11(01):C01003–C01003, jan 2016. doi: 10.1088/1748-0221/11/01/c01003. URL https://doi.org/10.1088/1748-0221/11/01/c01003.

[39] Edward Bartz. The token bit manager chip for the CMS pixel readout. 2003. doi: 10.5170/CERN-2003-006.185. URL https://cds.cern.ch/record/720634.

[40] P Moreira, A Marchioro, and P Placidi. CMS Tracker PLL Reference Manual. Technical report, CERN, Geneva, 2000. URL http://cds.cern.ch/record/1069705.

[41] L Amaral, S Dris, A Gerardin, T Huffman, C Issever, A J Pacheco, M Jones, S Kwan, S C Lee, Z Liang, T Liu, Z Meng, A Prosser, S Padadopoulos, I Papakonstanstinou, C Sigaud, S Silva, C Soos, P Stejskal, J Troska, F Vasey, P Vichoudis, T Weidberg, A Xiang, and J Ye. The versatile link, a common project for super-LHC. *Journal of Instrumentation*, 4(12):P12003–P12003, dec 2009. doi: 10.1088/1748-0221/4/12/p12003. URL https://doi.org/10.1088/1748-0221/4/12/p12003.

[42] Giacomo Ripamonti, Stefano Michelis, Federico Faccio, Georges Blanchot, Stefano Saggini, Roberto Rizzolatti, Mario Ursino, Adil Koukab, and Maher Kayal. A 2.5V Step-Down DC-DC Converter for Two-Stages Power Distribution Systems. *PoS*, TWEPP-17:059. 5 p, 2018. doi: 10.22323/1.313.0059. URL https://cds.cern.ch/record/2315732.

[43] Tbm manual. https://www.physik.uzh.ch/~leac/CMSPixel/Docs/TBM/TBM08b_Documentation.pdf.

[44] presentation on st testing. https://www.physik.uzh.ch/~leac/CMSPixel/Docs/PresentacionJueves-DanielH.pdf.

[45] F. Scholz. Untersuchung der eigenschaften eines $co_2 khlsystemsfrdasup - gradedescms - spurdetektors. Master'sthesis$, 2012.

[46] M. Lipinski. *System Tests and Qualification of Pixel Modules and DC-DC Converters for the Phase-1 Upgrade of the CMS Pixel Detector*. PhD thesis, 2019.

[47] Failure of feast2.1-based modules in the cms pixel detector system (physics run 2017). https://project-dcdc.web.cern.ch/public/Reports.html.

[48] Armin Burgmeier. *Position Resolution and Upgrade of the CMS Pixel Detector and Search for the Higgs Boson in the tau+tau- Final State*. PhD thesis, 2014.

[49] Particle Data Group. Review of Particle Physics. *Progress of Theoretical and Experimental Physics*, 2020(8), 08 2020.

[50] Claudio Canali, Filippo Nava, and Lino Reggiani. *Drift velocity and diffusion coefficients from time-of-flight measurements*, pages 87–112. Springer Berlin Heidelberg, Berlin, Heidelberg, 1985. ISBN 978-3-540-38849-4. doi: 10.1007/3-540-13321-6_3. URL https:

//doi.org/10.1007/3-540-13321-6_3.

[51] Electron mobility vs doping concentration. http://www.ioffe.ru/SVA/NSM/Semicond/Si/electric.htm.

[52] The Phase-2 Upgrade of the CMS Tracker. Technical report, CERN, Geneva, Jun 2017. URL https://cds.cern.ch/record/2272264.

[53] Collaboration CMS. A MIP Timing Detector for the CMS Phase-2 Upgrade. Technical report, CERN, Geneva, Mar 2019. URL https://cds.cern.ch/record/2667167.

[54] N. Moffat, R. Bates, M. Bullough, L. Flores, D. Maneuski, L. Simon, N. Tartoni, F. Doherty, and J. Ashby. Low gain avalanche detectors (LGAD) for particle physics and synchrotron applications. *Journal of Instrumentation*, 13(03):C03014–C03014, mar 2018. doi: 10.1088/1748-0221/13/03/c03014. URL https://doi.org/10.1088/1748-0221/13/03/c03014.

[55] G. Giacomini, W. Chen, G. DAmen, and A. Tricoli. Fabrication and performance of AC-coupled LGADs. *Journal of Instrumentation*, 14(09):P09004–P09004, sep 2019. doi: 10.1088/1748-0221/14/09/p09004. URL https://doi.org/10.1088/1748-0221/14/09/p09004.

[56] Anna Macchiolo and Ivan Vila. Final LGAD Sensors Characterization, AIDA-2020 Deliverable report. Mar 2020. URL http://cds.cern.ch/record/2713903.

[57] M. Ferrero, R. Arcidiacono, M. Barozzi, M. Boscardin, N. Cartiglia, G.F. Dalla Betta, Z. Galloway, M. Mandurrino, S. Mazza, G. Paternoster, F. Ficorella, L. Pancheri, H-F W. Sadrozinski, F. Siviero, V. Sola, A. Staiano, A. Seiden, M. Tornago, and Y. Zhao. Radiation resistant lgad design. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 919:16–26, 2019. ISSN 0168-9002. doi: https://doi.org/10.1016/j.nima.2018.11.121. URL https://www.sciencedirect.com/science/article/pii/S0168900218317741.

[58] R. Padilla, C. Labitan, Z. Galloway, C. Gee, S.M. Mazza, F. McKinney-Martinez, H.F.-W. Sadrozinski, A. Seiden, B. Schumm, M. Wilder, and et al. Effect of deep gain layer and carbon infusion on lgad radiation hardness. *Journal of Instrumentation*, 15(10):P10003–P10003, Oct 2020. ISSN 1748-0221. doi: 10.1088/1748-0221/15/10/p10003. URL http://dx.doi.org/10.1088/1748-0221/15/10/P10003.

[59] Z. Galloway, V. Fadeyev, P. Freeman, E. Gkougkousis, B. Gruey, C. A. Labitan, Z. Luce, F. McKinney-Martinez, H. F. W. Sadrozinski, A. Seiden, E. Spencer, M. Wilder, N. Woods, A. Zatserklyaniy, Y. Zhao, N. Cartiglia, M. Ferrero, S. Giordanengo, M. Mandurrino, A. Staiano, V. Sola, F. Cenna, F. Fausti, R. Arcidiacono, F. Carnasecchi, V. Cindro, G. Kramberger, I. Mandi?, M. Miku, and M. Zavrtanik. Properties of hpk ufsd after neutron irradiation up to 6e15 n/cm2, 2020.

[60] Low gain avalanche detectors for 4-dimensional tracking applications in severe radiation environments. https://journals.jps.jp/doi/10.7566/JPSCP.34.010015.

[61] Marco Mandurrino, F. Siviero, M. Tornago, R. Arcidiacono, M. Boscardin, N. Cartiglia, G. F. Dalla Betta, M. Ferrero, F. Ficorella, L. Pancheri, and et al. Demonstration of 200-, 100-, and 50-μ m pitch resistive ac-coupled silicon detectors (rsd) with 100% fill-factor for 4d particle tracking. *IEEE Electron Device Letters*, 40(11):1780–1783, Nov 2019. ISSN

1558-0563. doi: 10.1109/led.2019.2943242. URL http://dx.doi.org/10.1109/LED.2019.2943242.

[62] Ucsc single channel board cern. https://twiki.cern.ch/twiki/bin/view/Main/UcscSingleChannel.

[63] Wikipedia page. https://en.wikipedia.org/wiki/Strontium-90.

[64] Strontium decay webpage. https://www.nndc.bnl.gov/nudat2/reCenter.jsp?z=38&n=52.

[65] 56 - on the energy loss of fast particles by ionisation. In D. TER HAAR, editor, *Collected Papers of L.D. Landau*, pages 417–424. Pergamon, 1965. ISBN 978-0-08-010586-4. doi: https://doi.org/10.1016/B978-0-08-010586-4.50061-4. URL https://www.sciencedirect.com/science/article/pii/B9780080105864500614.

[66] Tct @ particulars. http://particulars.si/products.php?prod=LargeScanTCT.html.

[67] Hartmut F. W. Sadrozinski, Abraham Seiden, and Nicolò Cartiglia. 4-dimensional tracking with ultra-fast silicon detectors, 2017.

[68] Aida2020 official webpage. https://aida2020.web.cern.ch/aida2020/.

[69] Marco Ferrero, Roberta Arcidiacono, Marco Mandurrino, Valentina Sola, and Nicoltiglia. *An Introduction to Ultra-Fast Silicon Detectors: Design, Tests, and Performances*. 06 2021. ISBN 9781003131946. doi: 10.1201/9781003131946.

[70] Marta Tornago, R. Arcidiacono, Nicolo Cartiglia, M. Costa, Maricel Ferrero, Marco Mandurrino, Federico Siviero, Valentina Sola, Amedeo Staiano, A. Apresyan, K. Petrillo, R. Heller, S. Los, G. Borghi, Maurizio Boscardin, Gian-Franco Dalla Betta, F. Ficorella, Lucio Pancheri, G. Paternoster, and A. Seiden. Resistive ac-coupled silicon detectors: principles of operation and first results from a combined analysis of beam test and laser data, 07 2020.

[71] Etroc presentation. https://indico.fnal.gov/event/19437/contributions/52267/attachments/32463/39763/MTD-BO-6_In-depth-ETL-ASIC_Liu.pdf.

[72] R. Mulargia. Preliminary beam test results for the TIMESPOT project. *Journal of Instrumentation*, 15(11):C11004–C11004, nov 2020. doi: 10.1088/1748-0221/15/11/c11004. URL https://doi.org/10.1088/1748-0221/15/11/c11004.