

Study of monolithic CMOS pixel sensors in the Belle II experiment upgrade

November 4, 2023

Contents

Introduction	6
1 Belle II and SuperKEKB (SKB) accelerator	9
1.1 Physics program of the B-factories	9
1.1.1 Open questions in SM	10
1.1.2 Peculiarity of asymmetric B factories	10
1.2 SuperKEKB accelerator	12
1.2.1 The facility	12
1.2.2 "Nano-beam" scheme	13
1.3 Belle II detector	14
1.3.1 Vertex Detector (VXD)	14
1.3.2 Central Drift Chamber (CDC)	16
1.3.3 Particle identification system (TOP e ARICH)	16
1.3.4 Electromagnetic calorimeter (ECL)	18
1.3.5 K_L muon detector (KLM)	18
1.3.6 Trigger system	18
1.4 Current state of data taking	19
2 Belle II Upgrade	21
2.1 Background sources and limitations in Belle II	21
2.1.1 Major background sources	21
2.1.2 Current background status and future predictions	22

2.2	Purposes of the upgrade	24
2.3	Summary of possible VXD upgrade	25
2.3.1	Depleted Field Effect Transistor (DEPFET)	25
2.3.2	Thin and Fine-Pitch SVD	26
2.3.3	Silicon On Insulator (SOI)	26
2.3.4	CMOS Monolithic Active Pixels Sensor	28
3	VTX detector	31
3.1	VTX Layout and mechanical structure	31
3.1.1	iVTX	32
3.1.2	oVTX	34
3.1.3	Thermomechanics and data transmission	36
3.2	Performance simulation	37
3.2.1	Potential VTX geometries	37
3.2.2	Tracking efficiency at low momentum and impact parameter resolution	37
3.2.3	Vertexing resolution	38
3.3	OBELIX chip design	39
3.3.1	Sensor specification	40
3.3.2	Sensor implementation	42
4	CMOS MAPS sensors	46
4.1	Semiconductor detectors	46
4.1.1	Movement of charge carriers and signal formation in semiconductors	47
4.1.2	The pn junctions as detector	48
4.1.3	Complementary Metal-Oxide-Semiconductor (CMOS)	49
4.2	Hybrid and monolithic pixel sensors	50
4.2.1	Hybrid	50
4.2.2	Pixel	50
4.3	CMOS Monolithic Active Pixel Sensors technology	50

4.4	History of Monopix developments	50
5	TJ-Monopix 2	51
5.1	Matrix and flavors	51
5.1.1	Flavors	52
5.1.2	Pixel design	53
5.1.2.1	Improved front-end circuit design	54
5.2	Threshold and noise	55
5.2.1	S-Curve method	55
5.2.1.1	Normal FE	56
5.2.1.2	Cascode FE	58
5.2.1.3	HV-Cascode FE	58
5.2.1.4	HV-Normal FE	60
5.2.1.5	Summary Table	61
5.2.2	Threshold dispersion and tuning	62
5.2.2.1	First results from fine tuning	63
5.3	ToT calibration with internal injection	64
5.3.1	Injection circuit issues	65
5.3.2	Time Over Threshold (TOT) curves and fit [CHECK]	65
5.4	Response to radioactive source and absolute calibration	66
5.4.1	^{55}Fe	67
5.4.2	^{241}Am	69
5.4.3	^{109}Cd	69
5.4.4	Injection capacitance calibration	69
5.4.5	Check on linearity of tot fit	73
5.5	Operation with low threshold	75
5.5.1	Register optimization	75
5.5.2	Comparison between data and simulation	76
5.5.2.1	some nice picture of the optimized thr and tuning	77
5.6	Cross talk issue and mitigation	78

5.6.1	Hot pixel issue	78
5.6.2	Hot pixel strategy (study)	80
5.6.3	Cross-talk (Results)	82
5.6.4	Mitigation	84
5.6.4.1	Final results?	86
5.7	Test Beam results	87
5.7.1	Experimental apparatus and DUTs	87
5.7.2	Hit detection efficiency measurements	88
6	Conclusions	91
	Bibliography	92

Introduction

Belle II is a particle physics experiment located at KEK laboratory in Tsukuba (Japan). The detector is a general-purpose spectrometer to study electron-positron collisions produced by the SuperKEKB accelerator, a second generation flavor-factory which operates at the luminosity frontier, holding the world record of instantaneous luminosity with $L_{ist} = 4.7 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$.

SuperKEKB is the upgrade of the preceding facility KEKB (operational from 1998 to 2016) and it consists in a 3 km-circumference asymmetric accelerator which collides electrons and positrons beams at a center-of-mass energy near the $\Upsilon(4S)$ resonance ($\sqrt{s} = 10.58 \text{ GeV}$). It started its data taking in March 2019.

In the next decade, the collider aims to collect an unrivaled dataset of 50 ab^{-1} (x50 Belle dataset, x100 BaBar dataset) and to reach a new peak of instantaneous luminosity. This will allow to study the charge-parity violation in B mesons system with more precision and to search for new hints of physics beyond the Standard Model.

To achieve these challenging targets, it would be necessary a significant upgrade of the accelerator and its main components (like the injection system and the equipment nearby the interaction region) and it could be an opportunity to install also a new detector. As a matter of fact, to the increase in luminosity corresponds not only large data collected and greater possibility to study rare processes, but also higher doses of radiation and greater backgrounds, which could undermine the integrity and the operation of the Belle II detector. In particular the subdetectors which are closest to the beam pipe are those more exposed to severe conditions, like the vertex detector (VXD), composed by the inner pixel detector (PXD, made by layers of pixels) and the outermost silicon vertex detector (SVD, made by layers of strips). They deal with the reconstruction of charged particle tracks and of decay vertices with high performance. Recent studies have shown that the current detector could operate efficiently up to a luminosity of $L_{ist} = 2 \times 10^{35} \text{ cm}^{-2}\text{s}^{-1}$, but safety margins are not so large. Consequently, in this context, different upgrade projects have been proposed, which intend to design a new vertex detector, making it more resistant even in harsher working conditions, while the luminosity will be gradually increased.

This work focuses especially on the VerTeX Detector (VTX) proposal (the one chosen for the final upgrade), whose program provides for replacing the whole VXD with fully-pixelated five layers at different distances from the beam pipe, equipped with the same type of sensor technology, which is the CMOS Depleted Monolithic Active Pixel Sensors (DMAPS).

The good results achieved by the ALICE experiment (LHC, CERN), which employed the same technology, have addressed to this solution which has proven

to be reliable and promising in maintaining low occupancy, despite the worse expected background environment, and good radiation hardness even after irradiation.

In order to fulfil the physics requirements of Belle II experiment, a new silicon sensor is being designed, called OBELIX, exploiting the 180 nm TowerJazz Semiconductor process. Developments will ensure a faster, lighter and highly granular chip, reducing the material budget and improving tracks and vertices reconstruction.

OBELIX planning is based on studies done on previous prototypes, among which TJ-Monopix 2, whose characterization is the main topic of this thesis. Continuous laboratory experiments and beam tests have been conducted and are still in progress, in order to study the efficiency of the chip before and after irradiation, its power consumption and also to fully characterize its electrical features. In particular, we have completely characterized the response of the pixel matrix, returning important results that have allowed to interpret data taken during the Test Beam at Desy (July 2022), and that are being used in the design of the OBELIX chip. In more details, the threshold distributions for all the different types of front-end circuits implemented in the matrix have been studied, together with their dispersion and noise distributions. The calibration of the Time Over Threshold curves (which is a time width signal processing method used in this prototype) have been done by internal injection tests. The absolute calibration of the whole matrix have been achieved, employing a ^{55}Fe radioactive source. Other radioactive sources have been used too, in order to check the trend of the ToT curves for charge values not accessible by internal injection. Additionally, different register settings have been examined in the interest of operating the matrix at low threshold, that is crucial to keep high efficiency even after irradiation. For this reason, several tests have been conducted to tune the threshold, in order to reduce the dispersion and make the threshold on the matrix as uniform as possible. During this investigation, a cross-talk issue has been discovered and therefore studied to understand its causes and possible solutions to mitigate this effect.

Chapter 1 briefly introduces some of the open questions in the Standard Model, in order to depict the background of the Belle II physics program. Then a short description of the SuperKEKB accelerator and Belle II detector is given, too.

Chapter 2 presents the fundamental reasons behind the choice of an upgrade. The primary sources of the experiment background are described in a few words, to understand the limitation of the detector and the accelerator, for increasingly higher luminosity values. Eventually a summary of the four main upgrade proposals for the vertex detector is presented, which are distinguished by the different type of sensors employed: Depleted Field Effect Transistor (DEPFET), Thin and Fine-Pitch SVD, Silicon On Insulator (SOI) and CMOS Monolithic Active Pixels Sensors.

Chapter 3 examines in depth the VerTeX detector (VTX) upgrade program, which involves the CMOS Monolithic Active Pixel Sensors as fundamental components of the five layers of the final vertex detector. Continuous studies and simulations are ongoing to test the performance, and some of them are shown here. Moreover the specifications and the implementation of the new chip (OBELIX) thought for this proposal are described. The innovative sensor has to fulfill the requirements of Belle II experiment, even in extreme environment

due to higher doses of radiation and backgrounds.

Chapter 4 describes the principles underlying the operation of semiconductor detectors and some different type of sensors which use this technology, like the hybrid and monolithic pixel sensors. In particular the CMOS Monolithic Active Pixel Sensors technology is presented, on which the entire developments of the OBELIX chip is based. In the end, the history of the developments that led to the TJ-Monopix chip series is retraced, in order to better understand the main features of the last one, TJ-Monopix 2, which represents the starting point for OBELIX design, and whose characterization is the work of this thesis.

Chapter 5 lastly shows the results obtained from laboratory measurements and tests conducted on the TJ-Monopix 2 chip. The response of the matrix has been studied in different working conditions, in order to analyze the behaviour at high and low threshold. The absolute calibration of the all front-end circuits implemented in the chip, has been done too. Moreover a cross-talk issue have been discovered and analyzed, in order to understand its causes and a possible mitigation of this effect since it prevented from using the matrix at low threshold.

1. Belle II and SuperKEKB (SKB) accelerator

The first chapter introduces some of the main unexplained aspects of the Standard Model (SM), on which the Belle II physics program is focused. A short description of the SuperKEKB accelerator and the Belle II detector's structure is also presented and in conclusion some highlights on the current state of measurements.

1.1 Physics program of the B-factories

The SM is a very successful theory describing three of the fundamental forces involving elementary particles: the strong, weak and electromagnetic interactions, with the exclusion of the gravitational force. It classifies all the elementary constituents of matter in 4 main groups, as shown in Figure 1.1: quarks and leptons, constituting the matter fields; the gauge bosons, representing the interactions; and the Higgs boson, whose non-zero vacuum expectation value is needed to give mass to the otherwise massless matter fields.



Figure 1.1: Particle classification in the Standard Model.

1.1.1 Open questions in SM

Despite the undeniable success of the SM in providing predictions for all known physics phenomena, which have been experimentally verified with high precision over the years, there are many fundamental aspects of nature on which it is unable to give answers [1]. Some of them are listed in the following.

- Three generations of quark and leptons have been discovered, but it is not known whether they are the only ones and why.
- Moreover, the reasons behind quark and leptons mass hierarchy is unknown.
- Although the Higgs mechanism is able to explain the cause of elementary particles' masses through spontaneous electro-weak symmetry breaking, it is not clear whether neutrinos can get their non-zero but very small masses through the interaction with the Higgs boson.
- Another open question is the matter-antimatter asymmetry in the Universe. Charge-Parity (CP) violation is necessary to explain the asymmetry, but the SM mechanism would predict a value several orders of magnitude smaller than what is needed to explain the matter domination over antimatter, which allowed the evolution of the universe as we know it today.
- The flavour-changing currents are phenomenologically described with two mixing matrices: the Cabibbo-Kobayashi-Maskawa (CKM) matrix for the quarks and the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix for the neutrinos. The SM fails to provide any explanation neither for the structure of the matrices (only charged for the quarks, only neutral for the neutrinos), nor for the values of their elements (nearly diagonal for the quarks, highly non diagonal for the neutrinos).
- Astrophysical observations indicate the existence of dark matter, but its origin and nature have not been explained yet.

All these open questions stimulate the research for new particles and processes that could provide more fundamental explanations.

At the energy frontier, experiments at the Large Hadron Collider (LHC) at CERN (CH) are looking for new particles created from the proton-proton collision with a center mass energy up to 14 TeV.

At the luminosity frontier, instead, new particles and interactions are searched through precision measurements of suppressed reactions or deviations from SM. The discrepancies could be interpreted as a clue of new physics beyond SM. The Belle II experiment at the SuperKEKB is following this last approach .

In particular, the experiment investigates the CP violation in the B mesons system and searches for new physics evidence in the decays of B and D mesons, in τ leptons and in the dark matter sector (DM).

1.1.2 Peculiarity of asymmetric B factories

The SuperKEKB e^+e^- collider operates with a center-of-mass energy of $\sqrt{s} = 10.58$ GeV at the $\Upsilon(4S)$ resonance, which decays almost instantaneously into a pair of B - anti B mesons in nearly 96% of cases.

In SuperKEKB the beam energies are different, leading to a center-of-mass boost in the laboratory that allows the reconstruction of the B-mesons decay vertices, their lifetimes, and the time-dependent decay rate asymmetries.

In a beams symmetric situation, the B-mesons would be produced almost at rest, decaying roughly at the same point with undetectable decay length. The investigation of CP violating processes instead, requires measuring the decay time difference of the two B mesons and its uncertainty is dominated by that of the decay vertex measurement (order of hundreds microns). Let us look at this in more details.

SuperKEKB collides an electrons beam of 7 GeV (High Energy Ring, HER) with a positrons beam of 4 GeV (Low Energy Ring, LER) with a Lorentz boost factor of $\Upsilon(4S)$ of $(\beta\gamma)_{\Upsilon(4S)} \approx 0.28$.

The same boost is also acquired by the B-mesons, because they are produced almost at rest ($m_{\Upsilon(4S)} - m_{B_0} \approx 19$ MeV). Moreover knowing that $\tau_B \simeq 1.5 \times 10^{-12}$ s and so $c\tau_B \simeq 450$ μm , we can compute the average flight distance travelled before decaying:

$$l = (\beta\gamma)_{\Upsilon(4S)} c\tau_B \approx 126 \mu\text{m} \quad (1.1)$$

This value must be within the vertex detector sensitivity in order to distinguish the vertex decay and as consequence to make precision measurements of lifetimes, mixing parameters and CP violation. The main task of the Vertex Detector (VXD) is to reconstruct the production and decay vertices of the particles originated from the beam collisions. This aspect is crucial to perform time-dependent measurements, core of the Belle II physics program. The six-layer VXD can determine the position of the vertices with a precision better than 100 μm , allowing to reconstruct secondary vertices, i.e. the decay position of the particles coming from B decays, and also from τ leptons and D mesons.

The event kinematics is illustrated in Figure 1.2. The two B mesons are produced in an entangled quantum state, so when from the decay products of one of the two it is possible to assign its flavor (for example B^0 , identified as B_{tag}^0) accordingly can be assigned that of the second, which will be the opposite (\bar{B}^0 , called \bar{B}_{phys}^0).

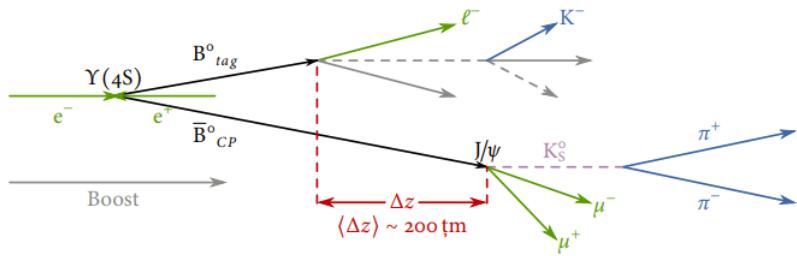


Figure 1.2: Example of the kinematics of the golden channel of Belle II experiment.

After this reconstruction, both B decay vertex positions in the longitudinal direction z_1 and z_2 are evaluated, in order to compute their difference:

$$\Delta z = z_1 - z_2 = (\beta\gamma)_{\Upsilon(4S)} c \Delta t \quad (1.2)$$

where Δt is the proper time decay difference.

Therefore this topology allows to transform a temporal information in a spatial one that we are able to measure using a high precision vertex detector. Without the boosted center of mass none of it could be possible, and this is an essential feature for an asymmetric B-factory.

1.2 SuperKEKB accelerator

Belle II sensitivity in the precision measurements is feasible especially thanks to the extraordinary performance of the SuperKEKB accelerator which host the (almost) hermetic detector. This complex facility is the result of efforts and efficient collaboration between the researches of KEK laboratory and all the international working groups that participate to the experiment.

1.2.1 The facility

SuperKEKB (Figure 1.3) is an asymmetric e^+e^- collider with a circumference of 3 km and a center of mass energy peak equal to $\sqrt{s} = 10.58$ GeV, which corresponds to the mass of the $\Upsilon(4S)$ resonance. Compared to its predecessor KEKB (which started its operation in 1998 and concluded it in 2010), the current accelerator has allowed to obtain the highest luminosity ever achieved, equal to $4.7 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ in July 2022. This target was possible using a new scheme to accelerate and collide the beams, the so called *nano-beam scheme* (section 1.2.2).

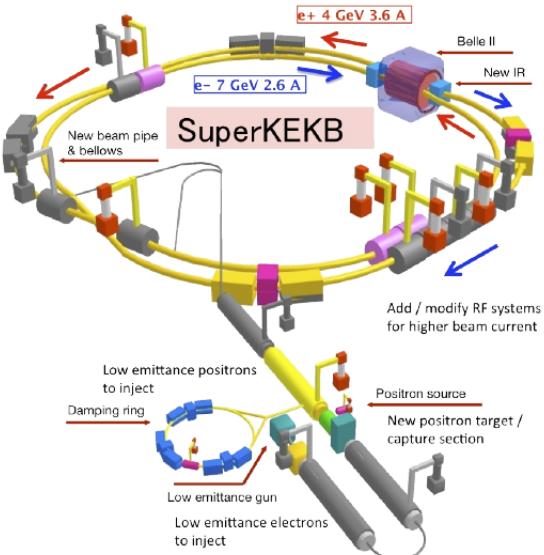


Figure 1.3: SuperKEKB accelerator structure.

Luminosity

Istantaneous luminosity is one of the key parameters of any accelerator and it represents the interaction rate per unit of cross section between colliding particles. Inverting this equation is possible to obtain N, namely the number of the physical events produced in the interaction with a given luminosity:

$$L = \frac{1}{\sigma} \frac{dN}{dt} \quad \Rightarrow \quad N = \int_0^T L \sigma dt \quad (1.3)$$

where T is the duration of the experiment and σ the cross section of the physical process of interest. Luminosity is dependent from both machine and beam parameters. With respect to this, it can be expressed as:

$$L = \frac{\gamma_{\pm}}{2er_e} \left(1 + \frac{\sigma_y^*}{\sigma_x^*}\right) \left(\frac{I_{\pm} \xi_{y\pm}}{\beta_y^*}\right) \left(\frac{R_L}{R_{\xi_{y\pm}}}\right) \quad (1.4)$$

where "±" denotes respectively positrons and electrons beam, $\sigma_{x,y}^*$ is the beam size at the Interaction Point (IP) in the horizontal and vertical plane, I is the beam current, β_y^* the vertical beta function at the IP. $\xi_{y\pm}$ is the vertical beam parameter which include the horizontal beta function at the IP, the horizontal emittance, the bunch lenght and the crossing angle between the beams. R_L and $R_{\xi_{y\pm}}$ are the reduction factors due to geometrical loss such as the hour-glass effect and finite crossing.

As already mentioned, SuperKEKB holds the current world record in luminosity (with $\beta_y^* = 1.0$ mm) and in the future the target will be to reach $6 \times 10^{35} \text{ cm}^{-2} \text{s}^{-1}$ (by the 2030s), by increasing currents beam and reducing their size at the IP, through the decrease of the betatron function down to $\beta_y^* = 0.3$ mm. But this process makes the beam-induced background increase significantly, risking deterioration and poor functioning of the detector.

For these reasons the control of the beam backgrounds becomes crucial: right now it has been estimated that the background should remain acceptable up to a luminosity value equal to $2.8 \times 10^{35} \text{ cm}^{-2} \text{s}^{-1}$ with $\beta_y^* = 0.6$ mm. So the chance to achieve higher luminosity is closely related to an upgrade plan of both the whole detector and the accelerator.

1.2.2 "Nano-beam" scheme

We have seen that the *beta function* β at the IP (β^*) is a decisive factor to define the luminosity. To be able to ramp the luminosity up, it is necessary to reduce the value of β depending also, but not only, on the variation of the other machine parameters that appear in the Equation 1.4.

The mechanism used in SuperKEKB is called *nano-beam scheme*, and it allows to obtain luminosity 40 times greater than that of KEKB, managing to decrease the β^* by a factor 20 compared to KEKB.

This new scheme, originally designed by P. Raimondi[2], dictates that the beam bunches have to collide with sufficiently small σ_x^* and at large angle. In case of SuperKEKB the angle is equal to 83 mrad at the IP (larger with respect to the crossing angle used in KEKB) with the beam size of 50 nm in the vertical

direction and $100 \mu\text{m}$ in the horizontal direction (in Figure 1.4 a simplified representation of the differences).

This strategy also helps to reduce the *hourglass effect*, which happens when the β^* is comparable or smaller than the bunch length, causing a decrease in luminosity. With a large crossing angle, the overlap length, which is the effective bunch length, is much shorter than the bunch length along the beam axis.

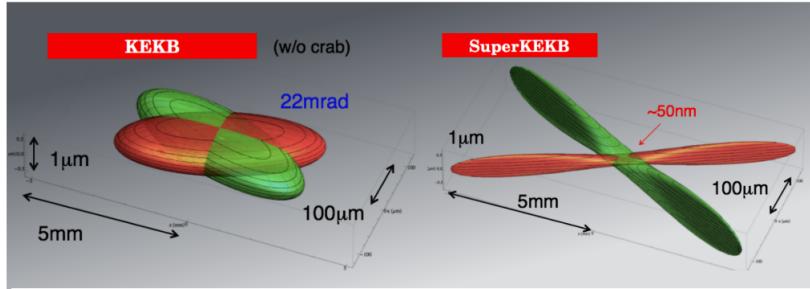


Figure 1.4: Comparison between the beam schemes used in KEKB and SuperKEKB.

Using a crossing angle large enough has other positive implications on the operation of the accelerator and its further improvements, like allowing the placement of a new focusing system at the IP (which may require more space), considering a future redesign of the interaction region.

1.3 Belle II detector

The Belle II detector is a general-purpose spectrometer which consists of a concentric subdetectors sequence placed around the berillium beam pipe of 10 mm radius, nearby the IP of the two beams. Here we will go through a brief description of the subdetectors (Figure 1.5) going in order from the beam pipe outwards: the Vertex Detector, the Central Drift Chamber, the TOP and the ARICH, the Electromagnetic Calorimeter and the K_L Muon detector.

In Table 1.1 a summary of the main characteristics of all subdetectors.

1.3.1 Vertex Detector (VXD)

The **VerteX Detector (VXD)** is composed by two devices divided into layers, the silicon Pixel Detector (PXD) and the Silicon Vertex Detector (SVD), for a total of six layers around the beam pipe.

The inner two layers of PXD (L12) consist of pixelated sensors based on the depleted field effect transistor (DEPFET) technology, realised with very thin ($< 100 \mu\text{m}$) sensors which allows to minimise multiple scattering, thus improving the tracking resolution for low-momentum particles. They are at a radius of 14 mm and 22 mm , respectively.

The remaining four layers of SVD (L3456) instead, are equipped with double-sided silicon strip (DSSD) sensors (at 39 mm , 80 mm , 104 mm and 135 mm

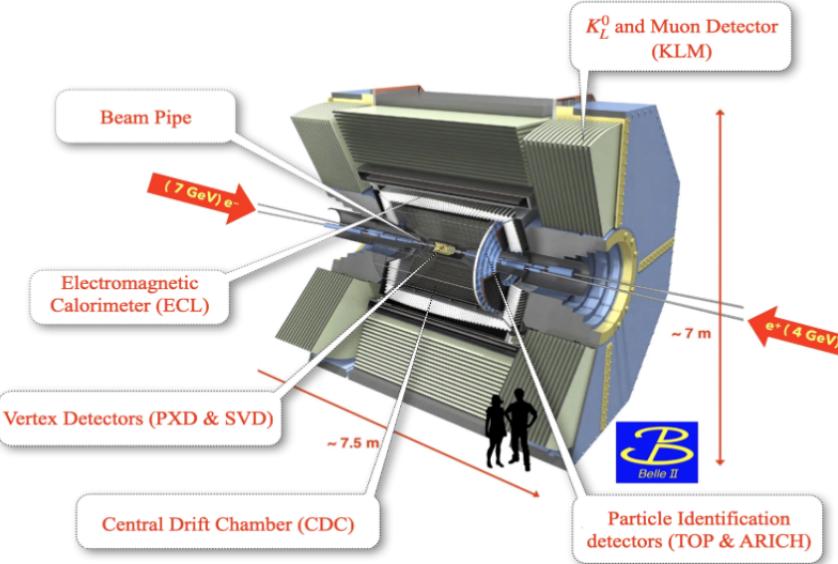


Figure 1.5: Belle II detector.

Purpose	Name	Component	Configuration	Readout channels	θ coverage
Beam pipe	Beryllium		Cylindrical, inner radius 10 mm, 10 μm Au, 0.6 mm Be, 1 mm paraffin, 0.4 mm Be		
Tracking	PXD	Silicon pixel (DEPFET)	Sensor size: 15 \times (L1 136, L2 170) mm^2 , pixel size: 50 \times (L1a 50, L1b 60, L2a 75, L2b 85) μm^2 ; two layers at radii: 14, 22 mm	10M	[17°; 150°]
	SVD	Silicon strip	Rectangular and trapezoidal, strip pitch: 50(p)/160(n)–75(p)/240(n) μm , with one floating intermediate strip; four layers at radii: 39, 80, 104, 135 mm	224k	[17°; 150°]
	CDC	Drift chamber with He-C ₂ H ₆ gas	14 336 wires in 56 layers, inner radius 160 mm, outer radius 1130 mm	14k	[17°; 150°]
Particle ID	TOP	RICH with quartz radiator	16 segments in ϕ at $r \approx 120$ cm, 275 cm long, 2 cm-thick quartz bars with 4 \times 4-channel MCP PMTs	8k	[31°; 128°]
	ARICH	RICH with aerogel radiator	2 \times 2 cm-thick focusing radiators with different n , HAPD photodetectors	78k	[14°; 30°]
Calorimetry	ECL	CsI(Tl)	Barrel: $r = 125$ –162 cm, endcap: $z = -102$ –+196 cm	6624 (barrel), 1152 (FWD), 960 (BWD)	[12.4°; 31.4°], [32.2°; 128.7°], [130.7°; 155.1°]
Muon ID	KLM	barrel: RPCs and scintillator strips	2 layers with scintillator strips and 12 layers with 2 RPCs	0.16k, ϕ 16k	[40°; 129°]
	KLM	endcap: scintillator strips	12 layers of (7–10) \times 40 mm^2 strips	17k	[25°; 40°], [129°; 155°]

Table 1.1: Summary of the main characteristics of all subdetectors.

respectively). Since a lower background rate is expected with respect to PXD, DSSD allow to achieve similar performance with a much smaller number of readout channels. These layers are mainly used for tracking/vertexing and also for particle identification (PID), through the measurement of the energy loss (dE/dx).

We can notice in Figure 1.6 that because of the essential asymmetric configuration of the beam energies and the consequent boost of the particles produced

in the collisions (section 1.1.2), the structure of the vertex detectors is also asymmetric along the longitudinal axis.

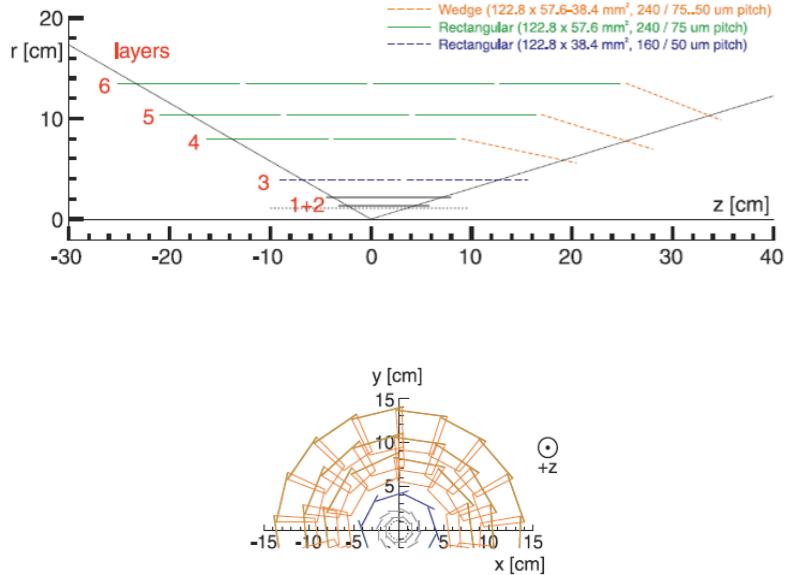


Figure 1.6: A schematic view of the Belle II vertex detector with a Be beam pipe and the six layers of PXD and SVD.

1.3.2 Central Drift Chamber (CDC)

This is the central tracking device, with a large-volume drift chamber and small drift cells. The chamber gas is composed of a He–C₂H₆ (50:50) mixture with an average drift velocity of 3.3 cm μs⁻¹ and a maximum drift time of about 350 ns for a 17 mm cell size.

The CDC contains 14336 wires arranged in 56 layers either in *axial* (aligned with the solenoidal magnetic field) or *stereo* (skewed with respect to the axial wires) orientation (Figure 1.7). In fact by combining information from both the axial and the stereo layers it is possible to reconstruct full three-dimensional helix charged tracks and measure their momenta. It also provides information for PID by measuring ionization energy loss, which is particularly useful for low-momentum particles.

1.3.3 Particle identification system (TOP e ARICH)

The **TOP** (**T**ime **O**f **P**ropagation) is a special kind of Cherenkov detector used for PID in the barrel region. It employs the two-dimensional information of a Cherenkov ring image, obtained from the time of arrival and the impact position of Cherenkov photons at the photodetector at one end of a 2.6 m quartz bar. It is composed by 16 detector modules, each one consisted in a 45 × 2 cm quartz bar (Cherenkov radiator) with a small expansion volume (about 10 cm long) at the sensor end of the bar (Figure 1.8).

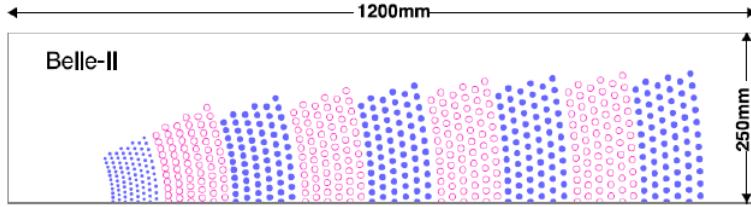


Figure 1.7: Schematic view of the CDC drift cells: blue dots represent the axial wires and the pink empty ones the stereo wires.

In order to achieve a single-photon time resolution of about 100 ps (required for a good PID), 16-channel microchannel plate photomultiplier tubes (MCP-PMT) are employed, specially developed for this purpose.

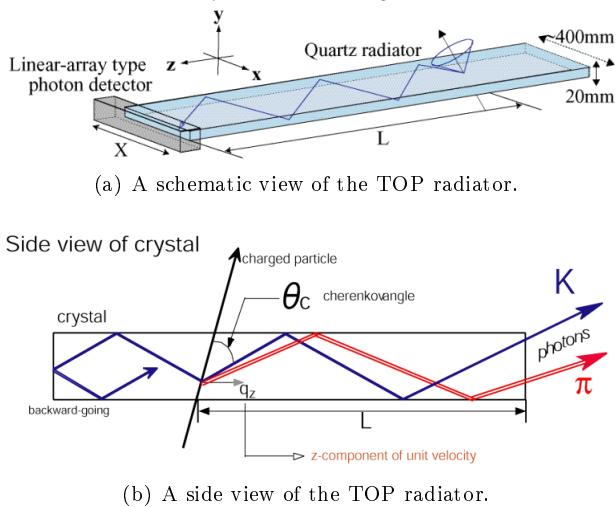


Figure 1.8: TOP detector.

The **ARICH (Aerogel Ring Imaging CHerenkov)** is used to identify charged particles and it is placed in the forward endcap region. It is a proximity focusing Cherenkov ring-imaging detector which adopts aerogel as Cherenkov radiator. In particular this detector employs a novel method to increase the number of detected Cherenkov photons: two 2 cm-thick layers of aerogel with different refractive indices ($n_1 = 1.045$ upstream, $n_2 = 1.055$ downstream) that increase the yield without degrading the Cherenkov angle resolution (Figure 1.9).

Hybrid avalanche photon detectors (HAPD) are exploited as single-photon-sensitive high-granularity sensors. Here photo-electrons are accelerated over a potential difference of about 8 KV and are detected in avalanches photodiodes (APD).

The main task of these detectors is to improve the K/π separation from

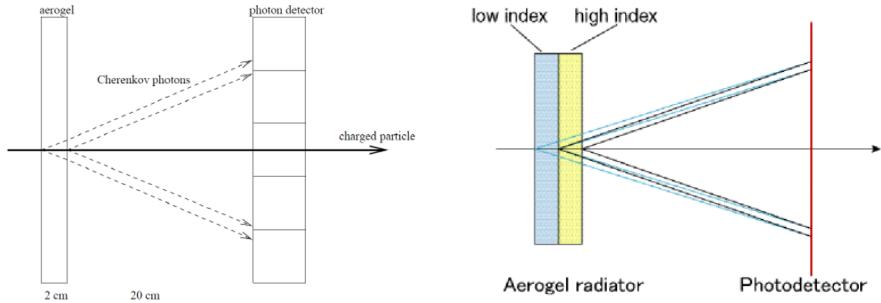


Figure 1.9: ARICH detector.

about ≈ 1 GeV until 3.5 and 4 GeV/c of momentum, respectively.

1.3.4 Electromagnetic calorimeter (ECL)

The **ECL** is a highly segmented array of tellurium-doped caesium iodide CsI(Tl) crystals assembled in a 3 m long barrel section with a radius of 1.25m, and two endcaps discs located at 2 m (forward) and 1 m (backward). All of them are instrumented with a total of 8736 crystal, covering about 90% of the solid angle in center-of-mass system.

This detector is used to detect gamma rays and to identify electrons in order to separate the latter from hadrons, especially pions.

1.3.5 K_L muon detector (KLM)

It consists of an alternating sandwich of 4.7 cm-thick iron plates and active detector elements located outside the volume of the superconducting solenoid that provides a 1.5 T magnetic field. The iron plates serve as the magnetic flux return joke for the solenoid. They also provide 3.9 interaction lengths or more of material, beyond the 0.8 interaction lengths of the calorimeter in which K_L^0 mesons can shower hadronically. The active detector elements have been chosen in order to cope with the reduction of the detector efficiency under the SuperKEKB background rates: resistive plate chambers (RPCs) for the outermost active layers and in the two innermost layers of the barrel and endcaps regions, scintillator strips with wavelength-shifting fibers are used, readout by silicon photomultipliers (SiPMs).

1.3.6 Trigger system

The trigger system of Belle II has a non-trivial role to identify events of interest during data-taking at SuperKEKB, where high background rates are expected. This system is divided into two levels: a hardware-based low-level trigger (L1) and a software-based high-level trigger (HLT), implemented in the data acquisition (DAQ) system.

- **L1**: based mainly on fast track reconstruction in the CDC and on ECL energy, has a latency of $5 \mu\text{s}$ and a maximum trigger output rate of 30 kHz, limited by the read-in rate of the DAQ.
- **HLT**: is a key component of the DAQ, used to fully reconstruct events that pass the L1 trigger selection. It has to reduce online event rates to 10 kHz for offline storage and it must identify track regions of interest for PXD readout in order to reduce data flux. It fully recreates events with offline reconstruction algorithms, using all detectors information except for the PXD.

1.4 Current state of data taking

SuperKEKB accelerator reached a peak luminosity of $4.7 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ and Belle II accumulated almost 428 fb^{-1} before the beginning of Long Shutdown 1 (LS1) in July 2022. (Figure 1.10).

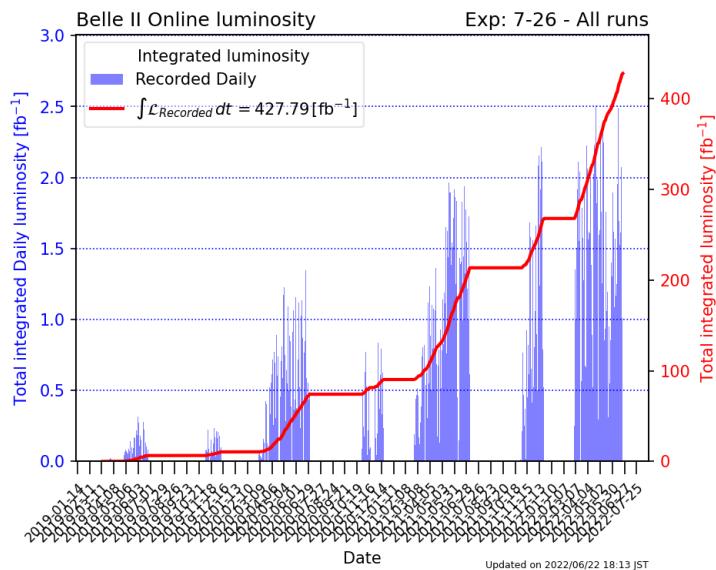


Figure 1.10: Total recorded integrated luminosity before Long Shutdown 1.

The target of SuperKEKB is to achieve a $L_{ist} = 6 \times 10^{35} \text{ cm}^{-2}\text{s}^{-1}$ and to increase the integrated luminosity from 428 fb^{-1} (current value, starting in 2019) to 50 ab^{-1} (as shown in the projection plot in Figure 1.11).

A three-phase program has been drawn up:

- **short term**: year 2023. Long Shutdown 1 (LS1) started in July 2022, and will be concluded at the end of 2023. The main purpose of the shutdown was the installation of a complete PXD, since in the 2019 installation only two ladders were included in layer 2. In addition, significant maintenance and improvement work has been carried out both for Belle II and SuperKEKB.

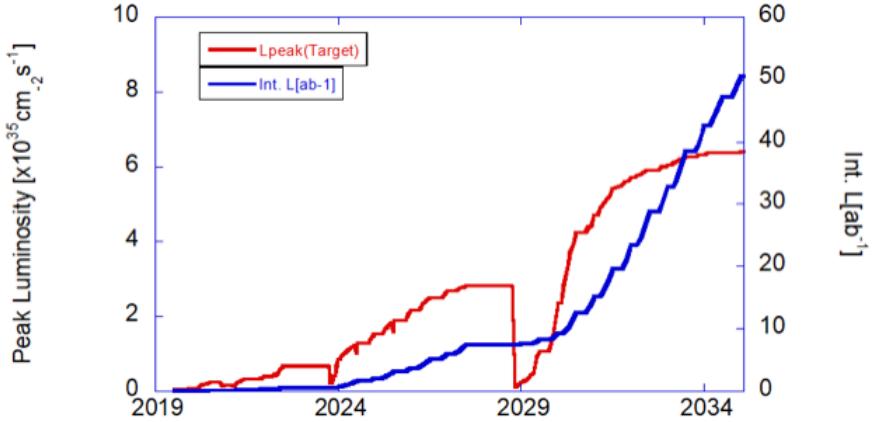


Figure 1.11: Luminosity projection plot (plan for the coming years).

- **medium term:** approximately year 2028-29. Long Shutdown 2 (LS2) will probably be needed for the upgrade of the Interaction Region (IR) to reach a new luminosity target $L_{peak} = 6 \times 10^{35} \text{ cm}^{-2} \text{s}^{-1}$. Several open questions and difficulties have triggered many studies and discussions about a possible redesign of the machine lattice during this phase. In particular it would be necessary to deal with the limitation of the optics of the machine, concerning the further increasing of the luminosity and accordingly of the backgrounds rates. A new Vertex Detector might be also required, to accommodate the new IR design, and other sub-detector upgrades are possible.
- **long term:** years > 2032 . Studies have started to explore upgrades beyond the currently planned program, such as beam polarization and ultra-high luminosity and so possibly L_{peak} in excess of $1 \times 10^{36} \text{ cm}^{-2} \text{s}^{-1}$. While the beam polarization has a concrete proposal, for ultra-high luminosity studies have just started.

At time of writing we are in the period of a long shut-down (LS1), last since June 2022 and the installation of a complete pixel detector (PXD) is almost done. The restart of data taking is planned at the beginning of 2024.

2. Belle II Upgrade

This second chapter wants to address some of the main reasons in favor of the upgrade of Belle II. We will give an overview of the primary background sources in the experiment to understand how to mitigate them in order to achieve a better performance of the whole detector, even ramping up the luminosity. Eventually we will also introduce some of the proposals made for the enhancement of the vertex detector.

[INTRODUCTION FROM CDR - OVERVIEW]

2.1 Background sources and limitations in Belle II

SuperKEKB is already the world's highest-luminosity collider and it aims to reach a new peak in the near future and also to increase the collected statistics, to become more sensitive to rare processes and precise measurements of Belle II physics program. But to be able to do this without loosing the good functionality of the entire detector, it is necessary to understand how to reduce the beam background where possible and how to cope with the consequent challenges.

Several simulations and measurements of beam background are still being done in order to guess possible future machine scenarios, under new luminosity conditions. This is necessary to study the vulnerability of the subdetectors (and more generally of the machine) and so to design the countermeasures to adopt against the deterioration of performance and materials.

2.1.1 Major background sources

In the following some of the primary *beam-induced* and *luminosity-dependent* background sources are listed.

Touschek effect : It is an intra-bunches scattering process, where the Coulomb scattering of two particles in the same beam bunch causes a variation of their energies, increasing the value of one of them and lowering that of the other from the nominal value. This interaction among the bunch particles is the first beam background source at SuperKEKB.

Beam-gas scattering : this represents the collision of beam particles with residual gas molecules in the beam pipe. It is the second beam background source and it can occur via two processes: *Coulomb interaction*, which changes the direction of the beam particles and *bremssstrahlung scattering*, which instead decreases their energy.

Because of these two processes, the scattered particles fall out the stable orbit and hit the beam pipe while they move around the ring. This mechanism causes electromagnetic showers that could reach the detector if their origin (loss position) is close to it.

Radiative Bhabha scattering and two-photon processes : There are several undesirable collision processes at the IP which have very high cross sections but only little interest for the physics studied in the experiment. Two of them are **Bhabha scattering** ($e^+e^- \rightarrow e^+e^-\lambda$) and **two-photon processes** ($e^+e^- \rightarrow e^+e^-e^+e^-$). In the first channel the emitted photon interacts with the iron magnets and produces a very large amounts of neutrons via the photo-nuclear resonance mechanism (such neutrons are the main background source for the outermost Belle II detector). The electrons-positrons pairs of the latter instead, can spiral around the solenoid field lines and leave multiple hits in the inner layers of the detector.

These processes increase the Belle II occupancy and radiation dose and they are referred as *Luminosity background* because their strength is proportional to the luminosity. The future upgrade intends to deal with this problem in order to keep occupancy low, even with higher radiation levels.

Synchrotron Radiation (SR) : X-rays emitted from the beam when electrons and positrons pass through the strong magnetic field near the IP. The HER beam is the main source of this type of background, because SR power is proportional to the square of beam energy and magnetic field. SR can potentially damage the inner layers of the vertex detector due to an higher radiation dose in that region. As a matter of fact, many current studies aim to enhance radiation hardness of the detector.

Making clear that even the intentional interaction of the beams is a source of noise for the measurements, there are also other background sources beyond those mentioned above and during the last decade a well-structured set of countermeasures have been developed trying to ease each one of them.

2.1.2 Current background status and future predictions

Several monitoring devices are located all along the accelerator to keep under control radiation doses on both detector and delicate regions of the ring, in such a way to intervene as soon as possible in case too high levels are reached. Indeed large doses of radiation could cause accidental damages on the detector, decreasing its performance.

A lot of dedicated beam background studies have been performed and for example, in Figure 2.1 are shown rates measured in June 2021 with a betatron function at the IP of $\beta_y^* = 1$ mm and a collision luminosity of $L = 2.6 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. Current background rates in Belle II are acceptable and in most cases, well below the limits listed in table 2.2.

Event though the current level is of no concern in terms of occupancy for the innermost layers of the vertex detector, in the case of a larger amount of SR,

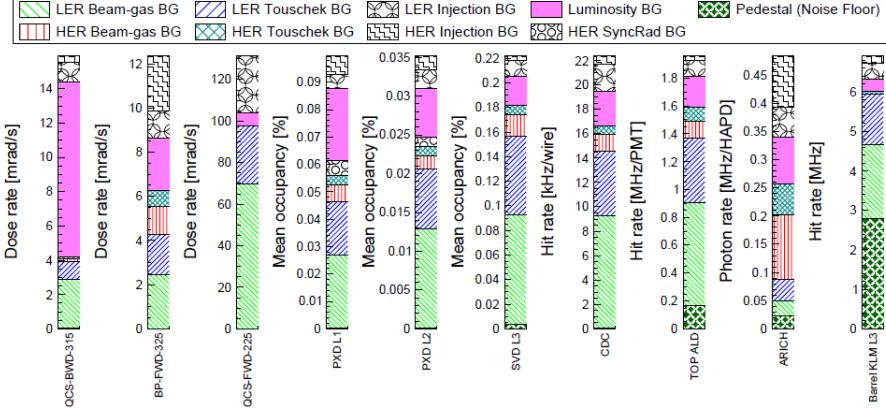


Figure 2.1: Measured Belle II background in June 2021. Each column shows different background sources for Belle II sub-detectors and also for superconducting quadrupole magnet (backward and forward) and the beam pipe.

Detector	BG rate limit	Measured BG
Diamonds	1–2 rad/s	< 132 mrad/s
PXD	3 %	0.1 %
SVD L3, L4, L5, L6	4.7 %, 2.4 %, 1.8 %, 1.2 %	< 0.22 %
CDC	200 kHz/wire	22.3 kHz/wire
ARICH	10 MHz/HAPD	0.5 MHz/HAPD
Barrel KLM L3	50 MHz non-luminosity BG	4 MHz
Before LS1 After LS1		
TOP ALD	3 MHz/PMT + luminosity BG	1.8 MHz/PMT

Figure 2.2: Background rate limits for Belle II detector sub-systems. The third column shows the total measured background rate in June 2021.

for example, it may cause inhomogeneities in PXD module, which would make more difficult to compensate them by adjusting the operation voltages of the affected ones.

Until now it can be said that SuperKEKB and Belle II are operating stably. Beam-induced background rates are well below the limits of the detector and do not prevent from increasing further the current and hence the luminosity, as demonstrated by the predictions for the background rates *before LS2* with a known machine configuration. For what concern the predictions at $L_{ist} = 3 \times 10^{35} \text{ cm}^{-2}\text{s}^{-1}$ instead, called *after LS2* operation, there are several uncertainties tied to the machine configuration. In fact at the moment, the *working machine lattice* to reach the target luminosity is not known and in addition the final design of the IR and beam pipes is not concluded yet.

Therefore an alternative solution is employed to roughly estimate the back-

ground rates. The background predicted Before LS2 phase is considered as a starting point and then different scaling factors are applied for single-beam background component, considering three different possible scenarios:

- **x2 - optimistic Scenario-1(v1)**
- **x5 - nominal Scenario-2(v2)**
- **x10 - conservative Scenario-3(v3)**, an arbitrary factor assuming that all single-beam backgrounds will be increased by a order of magnitude After LS2.

These are then used to simulate the behaviour of the whole detector in future perspectives, as we will see in the following.

2.2 Purposes of the upgrade

Current studies foresee that SuperKEKB may reach higher luminosity targets with the existing accelerator complex, but in order to achieve the established final value of $6 \times 10^{35} \text{ cm}^{-2}\text{s}^{-1}$, an enhancement of the interaction region are under consideration.

Belle II detector is also designed to operate efficiently under the high levels of backgrounds extrapolated to luminosity target, but safety margins are not so large. Moreover in the case of a redesign of the interaction region large uncertainties in the background extrapolations are unavoidable.

Therefore the global upgrade program is justified by many considerations, among them:

- improve detector's resistance to higher levelss of background;
- make each subdetectors long-lived against radiation damage;
- push forward safety margins for running at higher luminosity;
- develop the technology to cope with different future paths;
- improve overall physics performance.

In particular all different upgrade ideas of the whole Belle II detector intend to ensure its proper functioning at the higher level of luminosity, condering also further improvements of the lattice machine and so of the colliding beams. Indeed current detector configuration is not expected to maintain its performance level when facing higher beam background level or higher rates.

In regards to the Vertex Detector, all proposed improvements aim to:

- reduce occupancy level by employing fully pixelated and fast detector (nowadays CMOS technology has been chosen);
- increase robustness against tracking efficiency and resolution losses from beam background;

- improve radiation hardness for delaying detector ageing effects and so performance degradation;
- reduce the inserted material budget between subdetectors in order to achieve good resolution by lessening the multiple scattering, above all at lower momenta.

2.3 Summary of possible VXD upgrade

The Vertex Detector is particularly sensitive to machine background because it is the closest to the beam pipe and therefore subject to high doses of radiation. As we have already seen, current studies are trying to extrapolate how it could be affected by reaching the future luminosity target, but there are a lot of uncertainties due to models and still not well defined design of the interaction region. Moreover a completely new detector might be required, in event of a considerable redesign of the IR. However in this case, also the physics performance could be improved, taking advantage of the more recent technology developments.

In the following we will present in a few words the four main proposal for future upgrade: Depleted Field Effect Transistor (DEPFET) pixel, thin sensor, CMOS Monolithic Active Pixel Sensor and SOI technology. The first two are more conservative and try to exploit as much as possible of the existing detector, making some appropriate adjustments to the sensor type, readout or mechanical structure. The last ones instead, plan to build an entirely new detector.

During the design of all of them, some reference radiation levels have been considered to ensure that the background robustness of the innermost layers is suitable:

- Hit rate capability: $120 \text{ MHz}/\text{cm}^2$;
- Total Ionizing Dose: 10 Mrad/year ;
- NIEL fluence: $5 \times 10^{13} n_{eq}/\text{cm}^2/\text{year}$.

2.3.1 Depleted Field Effect Transistor (DEPFET)

This first proposal intend to minimize risks and costs of the project, preserving the general layout of the PXD system. The upgrade consists to improve the sensor above all, in order to provide higher safety factor for the allowed occupancy and to prevent some issues that at the moment weaken the good functionality of the detector.

Some of the main improvements are listed below:

- improve signal transmission on the pixel matrix and the signal processing in the read-out, in order to reduce the read-out time per row from the current 100 ns to 50 ns. In this way the frame time and the background occupancy might be reduced by factor 2, while leaving unchanged the optimized size and number of PXD pixels as it stands;

- increase the robustness against beam losses which could make inefficient or even inoperative gate lines on almost all PXD modules. This reaction seems to be due to a high photocurrent on the chip because of the high instantaneous dose. It could be mitigated by adding protection circuits on-chip;
- Total Ionizing Dose (TID) effect on the chip provokes an unexpected avalanche current that does not compromise the sensor performance but requires more power supply to provide enough current. This issue might be solved by bringing some changes in the DEPFET pixel layout.

2.3.2 Thin and Fine-Pitch SVD

The Thin and Fine-Pitch SVD (**TFP-SVD**) is a new detector concept that aims to improve not only SVD, but also the inner part of the CDC, whose functionality could be threatened by future beam background condition. This proposal takes into account the Double-sided Silicon Strip Detector (DSSD) as a prime candidate for a tracking device in the inner and middle detector volume since a single sensor can cover a large dimension of about $100 \times 100 \text{ mm}^2$. In the current detector the DSSD technology is already used in the SVD, which deals with vertex reconstruction and low momentum tracking, together with PXD.

One of the major improvements of this technology is the reduction of the material budget. Currently SVD has about $0.7\%X_0$ material budget per layer. TFP-SVD instead, decreasing the sensor thickness to $140 \mu\text{m}$, intend to reduce it to $0.41\%X_0$.

Moreover small sensor thickness is expected to reduce the voltage needed to reach the full depletion, even after radiation damage. The front-end thought for TFP-SVD (SNAP128) has 128 input channels and a 127 MHz clock in each of them, to generate the binary hit information sampled. It also offers a reduction of the amount cables.

Some concerns about TFP-SVD are the feasibility and efficiency of the final sensor production and the small signal charge due to the short path length of the particles through the sensor.

In any case a first prototype has been produced by Micron-Semiconductor Ltd (UK), with a size of $52.6 \times 59.0 \text{ mm}^2$. The characterization studies are in agreement with the expectation and also a lower full depletion voltage is confirmed. It is planned to increase the dimensions to $100 \times 100 \text{ mm}^2$ in the further prototype.

2.3.3 Silicon On Insulator (SOI)

The basic idea of the proposal is to replace the whole VXD detector employing a new design of pixel, called Dual Timer Pixel (DuTiP), based on SOI technology. This new sensor concept has been invented to cope with the expected higher background accordingly to higher value of luminosity ever achieved.

Concept

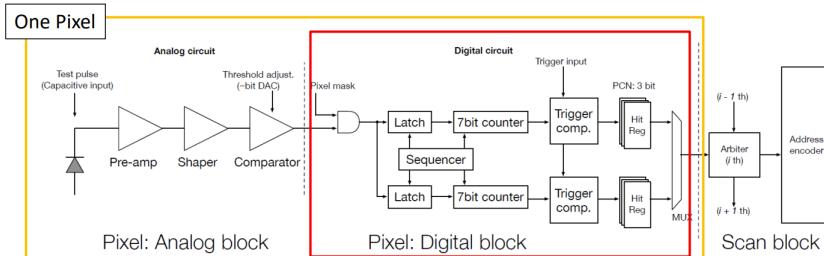
SOI technology has been chosen as baseline for the new pixel design thanks to its monolithic structure, thinness, low power consumption and low parasitic

capacitance. In addition it is resistant against neutron and single event upset (SEU¹), even though an important issue is TID effect on which efficient solutions have been studied.

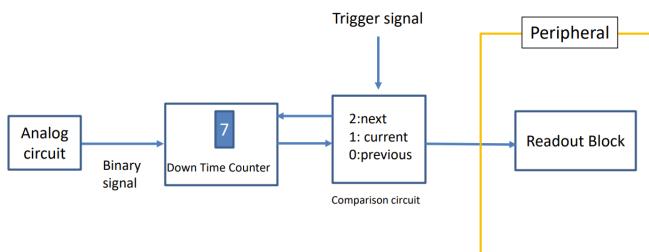
DuTiP has been invented to fulfil the requirements of a new vertex detector with faster readout, lower occupancy, smaller data size and smaller data transfer. In particular its concept rest on the concern to store at least two hits during Belle II trigger latency ($5 \mu\text{s}$), to avoid loss information of the inner part of the detector in higher background environment.

The analog part is quite standard for a binary detector and consists of a sequence of preamplifier, shaper and comparator. Digital part is equipped with two timers (7 bit counters) to store at least two hits. When a processed hit signal arrive to the digital part it is stored and one of the timers start to counting down from a starting time set to trigger latency plus one clock, waiting for trigger signal. If the trigger signal is received when the time is 1 (it could be also 2 or 0), the signal is readout as *Current* (*Next* or *Previous* respectively) timing (*PCN timings*). If the trigger is not received at the PCN timings in the pixel, the timer is reset.

This complicated digital circuit has to be assembled on each pixel and Lapis semiconductor $2.0 \mu\text{m}$ FD-SOI CMOS technology has been chosen, based on the experience gained in the successful development of other detectors like pixel detectors for the future ILC (SOFIST) and CEPC.



(a) Analog, Digital and Scan blocks for DuTiP detector.



(b) Operational sketch.

Figure 2.3: Schematic of DuTiP circuits.

¹A Single Event Upset (SEU or SEE, Single Event Error) occurs when a ionizing particle deposits charge close to a storage node (e.g. RAM cell, register) causing a bit value to flip leading to corrupt information. This effect could be mitigated by using, for example, redundant memory. It is not a permanent damage.

Sensor design and features

The size of the new designed pixel is $45 \mu\text{m}$ and the sensor layer thickness of $50 \mu\text{m}$, which gives an intrinsic resolution better than $15 \mu\text{m}$ on z direction averaging over incident polar angle. ALPIDE was chosen as analog circuit with some modification to adapt it to SOI technology.

DuTiP pixel detector is designed to cover the current VXD acceptance with 7 layers: 1-3 with S (smaller size chip) type sensors, 4-7 with L(larger size) type (Figure 2.4).

sensor type	layer	pitch [μm]	row \times column [pixels]	array $r\phi \times z$ [mm^2]	array area [cm^2]	chip $r\phi \times z$ [mm^2]
S	1-3	45	320×640	14.4×28.8	4.15	17.2×29.6
L	4-7	45	480×640	21.6×28.8	6.22	24.4×29.6

Figure 2.4: The size of Small (S) and Large (L) DuTiP chips.

Stitching technique allows to produce longer chips in the z direction, but the structure of the ladders has not been decided yet. Anyway the target is to minimize the dead region between chips in the ladder. For the inner layer of the detector might be possible the cooling with airflow at room temperature; for the outer layers instead, a combination of air and water flows.

For layer 1, which is expected to work in more severe condition of background, the pixel occupancy has been estimated with the trigger latency of $8.0 \mu\text{s}$ and for both L and S type it is small enough, $O(10^{14})$ or less, thus stable tracking and vertexing are contemplated. Moreover without using two timers for layer 1, the signal loss probability with the trigger latency of 4.5 (8.0) μs is about 0.2 (0.4)% and so not negligible. In fact if the background rate is higher and the latency is longer, the signal loss probability increases.

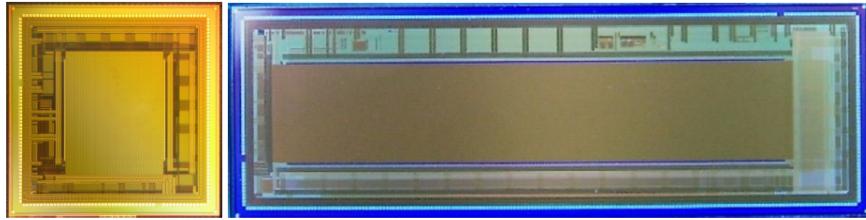
The first prototype of this new chip has been delivered in June 2021, with all in-pixel functionalities except for the scan block and the fast readout system. The chip is a matrix of 64×64 pixels and size of $6 \times 6 \text{ mm}^2$. Its characterization is ongoing and it seems to work fine, also with radioactive sources and red laser tests.

The second prototype DuTiP2 had been delivered in 2022, with 32×320 pixels and size of $17.2 \times 6 \text{ mm}^2$. It has all functionalities except for fast hit data collection to periphery. The phase of testing is ongoing. A third prototype has been submitted in 2023 and will be delivered in 2024 (all prototypes shown in Figure 2.5).

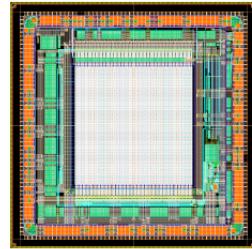
2.3.4 CMOS Monolithic Active Pixels Sensor

The last proposal is the one that we will analyze in more details in the next chapter. Like the previous one, it aims to replace the entire current VXD detector using a new technology, in this case the Monolithic Active Pixel Sensor CMOS (Complementary Metal-Oxide Semiconductor).

The program hopes to solve some of the issues discussed in the previous chapters, with a new system of two inner layers and three outermost, for a total of 5 stages equipped with a single sensor type, called VTX (layout in Figure 2.6). Also the



(a) DuTip1 prototype on the left and DuTiP2 on the right.



(b) DuTip3 third prototype.

Figure 2.5: DuTiP prototypes.

mechanical structure has been redesigned but it is expected that the all system could work at room temperature, so as consequence an important reduction of services is also contemplated.

The new pixel design is called OBELIX (Optimized BELle II pIXel sensor), based on the pixel matrix of the TJ-Monopix 2 chip, whose characterization is the main topic of this work (Chapter 5).

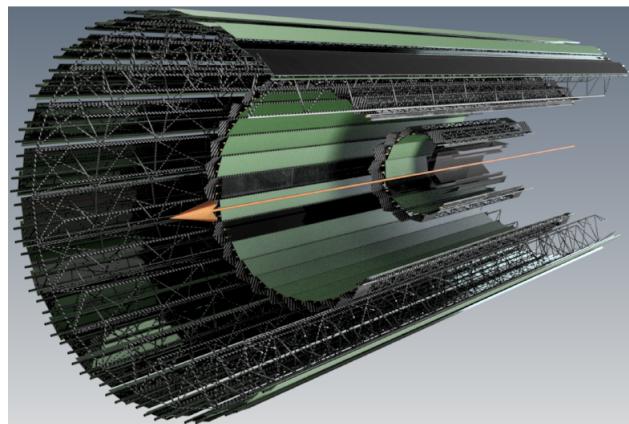


Figure 2.6: Overall VTX layout.

The VTX solution intends to reduce the current PXD integration time by at least two order of magnitude and also other advantages are expected:

- lower detector occupancy which allows to cope with higher background

and to mitigate any data-transmission bottleneck.

- Better tracking efficiency and improved momentum resolution and impact parameter resolution at low transverse momentum.
- Smaller data cables cross sections and less complex cooling system which might lighten the needed services.
- Control and power systems become unique due to the employment of a single sensor chip for all layers.

At the current state of art intense R&D (Research and Development) is being carried out, taking advantage from other experiments experiences like ALICE, with the same type of sensor.

Among all these options that we sift through, the last one has been chosen for the future upgrade of the vertex detector. The first two proposals have now been abandoned, the one that contemplates the use of the SOI technology instead, is considered as a backup option.

After this briefly review of the main upgrade proposals, we can now deepen into the VTX program in the following chapter.

3. VTX detector

This chapter focuses on one of the four proposals for the vertex detector upgrade of Belle II, that is VTX. After a brief reference to the reasons behind the vertex detector upgrade, we will go through VTX concept and layout, designed with a new geometry with respect VXD and also with a different mechanical structure and new pixel sensors (OBELIX), needed to fulfill the new requirements dictated by new environment conditions. Moreover all ongoing studies are supported by continual tests and simulations that we will also take a look at.

3.1 VTX Layout and mechanical structure

In section 1.2.2 we have introduced in a few words the concept of the *nano-beam* scheme, which could allow to achieve the new target of instantaneous luminosity. This new strategy requires a strong focusing of the beams in particular at the IP, resulting in a large amount of beam induced background and as consequence in a higher dose of radiation in the innermost detector layers. Therefore they have to be robust enough to keep good performance against these new hard conditions. Furthermore to be able to increase the luminosity, SuperKEKB might have to consider an improvement of the final focusing magnets and so a potentially re-design of the interaction region, including the detector but independently of its hit rates and radiation hardness issues.

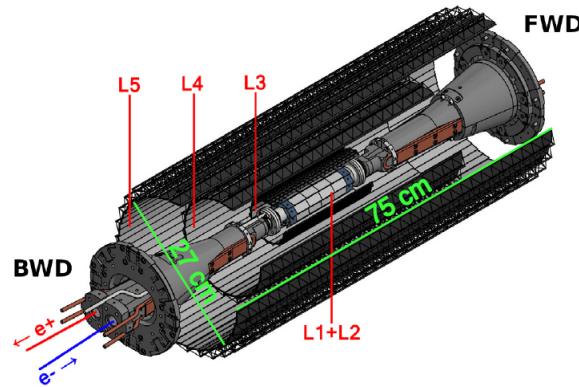


Figure 3.1: Concept of VTX layout with 5 barrel layers, filling the current VXD volume.

VTX project aims to replace the all VXD with a fully pixelated detector based on Depleted Monolithic Active Pixel Sensors (DMAPS) arranged on five layers at different distance from the beam pipe (Figure 3.1). Actually the radii and the number of the layers are currently subject to several studies and simulations, in order to achieve an optimized arrangement. For now two layers are planned in the innermost part (*i*VTX) and three in the outermost (*o*VTX). The active length of the ladders is expected to vary from 12 to 70 cm to cover the required acceptance of $17^\circ < \theta < 150^\circ$.

As already discussed for the other upgrade proposals, it is important to try to reduce the material budget, in order to minimize the multiple Coulomb scattering which particularly affects the very soft particles produced in Belle II collisions. By using a single sensor type, it is expected a reduction of the overall material budget up to 2% of radiation length, against the present 3% of VXD, which uses two different sensors such as pixels and strips.

3.1.1 iVTX

The *internal VTX* consists of the first two detector layers devised together with a self-supported air-cooled all-silicon ladder concept, where four contiguous sensors are diced out of a wafer, thinned and interconnected with post-processed redistribution layers. They are designed to be at 14 and 22 mm respectively from the beam pipe, and target an individual material budget of about 0.1% radiation length. This is actually achievable because the overall surface of these layers is moderate, below 400 cm^2 , as well as the sensor power dissipation is expected to be low and the connections needed for the operation to be a few. Precisely for these reasons, air cooling could be a workable system to avoid overheating.

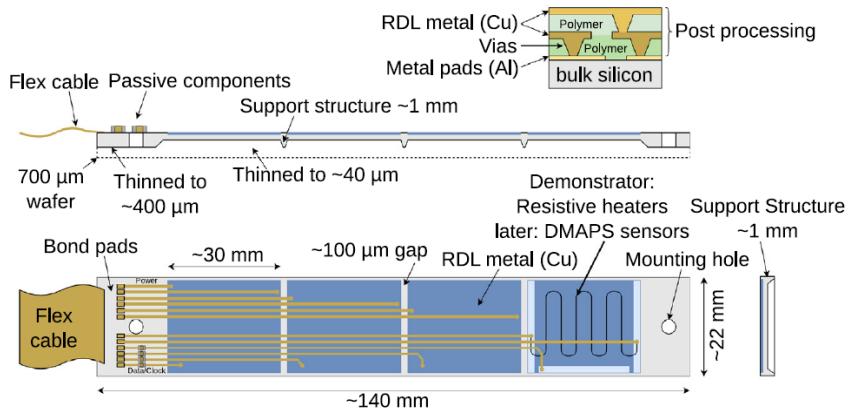


Figure 3.2: Sketch of the all-silicon ladder concept of the iVTX. Four dummy sensors are shown in blue on the silicon support in grey. The yellowish lines instead, indicate power and data transmission lines. Power is delivered to the ladder by a flex cable, which also transmits data to and from the chips in the final one.

The ladder has to be equipped with four of the aforementioned OBELIX chips and thinned to $50\text{ }\mu\text{m}$ except in some border regions, where a few hundreds

of μm are necessary to ensure mechanical stability. In order to interconnect the sensors along the ladder and provide a unique connector at the backward end, during the post-processing metal strips are etched on the redistribution layer (RDL). The latter has the main purpose to route power and data via impedance-controlled transmission lines to a flex cable, added at the end of the ladder. After the RDL processing, the backside of the ladder has to be thinned in accordance with what was previously mentioned. Mounting holes will be added via laser-cutting.

In Figure 3.2 is displayed a sketch of the iVTX demonstrator ladder (currently under production), 140 mm long and 22 mm wide (grey). Instead of the actual sensors, it is equipped with four dummy chips with a lenght of about 30 mm (blue), which are used as resistors to mimic the estimated heat load in order to test the air cooling system and more generally to characterize the electrical, mechanical and thermal performance of the ladder. A redistribution layer for power and data is also added to the demonstrator, in order to connect the chips with a flex cable at the end of the ladder (yellowish lines). In addition the wafer is thinned to 400 μm and the sensitive areas down to 40 μm with the purpose to test the mechanical integrity of the whole structure.

Research and Development (R&D)

The R&D is ongoing and the full-silicon ladder concept is currently being assessed with industrial partners. First thinned ladders have been produced and characterised with different thickness and geometry.

Furthermore several tests are focused on evaluating power delivery efficiency, the quality of the signals which travel through the ladder and also the process used to fully assembly it. In Figure 3.3 are shown eye diagrams from simulation with a transfer rate of 640 Mbps, which may imply that 320 Mbps of data rate will be possible.

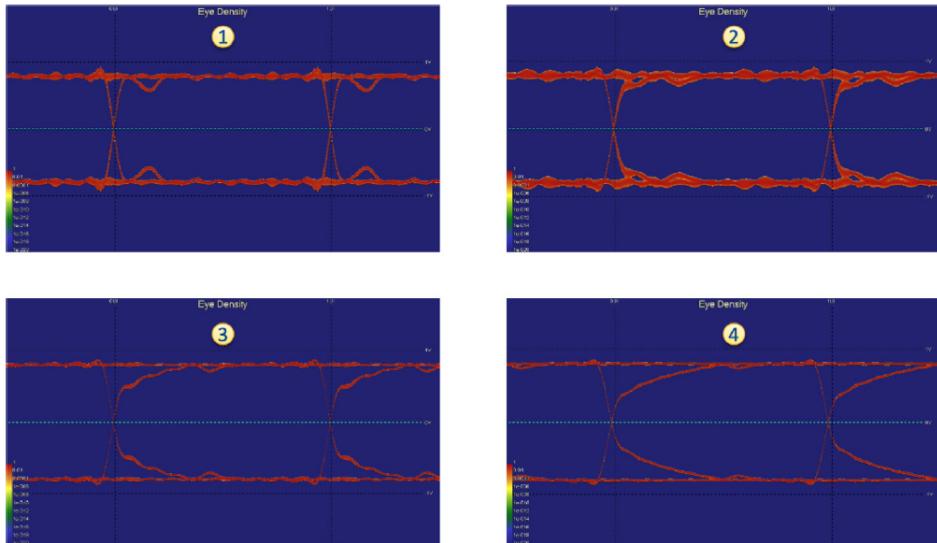


Figure 3.3: Eye diagrams of the iVTX data transmission lines at four different locations on the ladder.

Moreover it has been demonstrated that air at the temperature of 15° and flowing with a speed of 10 m/s succeeds to cool a single inner module, assuming power is uniformly dissipated on the sensor surface. The maximum temperature reached is 20°C.

Through very first estimates it is expected that an equivalent section of 6 tubes with 10 mm of diameter is necessary to expel the heat from the inner layers, roughly equal to 65 W. So it is essential to design a mechanical structure which provides for the space needed to the tubes in order to bring the air at the IP and also compatible with the new interaction region.

3.1.2 oVTX

The **outerVTX** consists of three layers respectively at radii of 39 or 69, 89 and 140 mm from the beam pipe and because of the larger distance required to cover the acceptance, they are not self-support. They follow a more traditional approach, strongly inspired by the design developed for the ALICE ITS2. Each ladder is water cooled and made of a light carbon fiber support structure, called ***truss***, which provides the mechanical integrity. Its structural design is shown in Figure 3.4 : 70 cm long and 5.8 g of weight, it is able to support more than 40 sensors in two rows next to each other with a small overlap, earning a material budget of 0.3% X_0 for the first two layers and 0.8% X_0 for the outermost one.

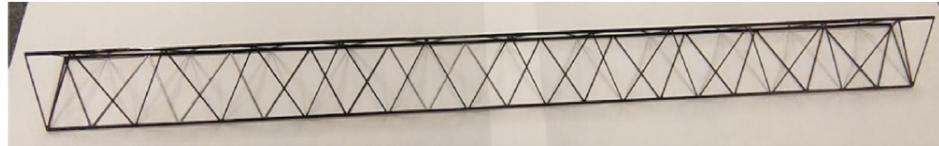


Figure 3.4: Prototype of the layer 5, called *truss*, which is the longest, made from thin carbon fibre structures.

For the cooling of the ladder a cold-plate concept is under development (Figure 3.5), on which the sensors are glued and that in turn is installed on the *truss*. For each row there is a polyimide cooling tube that runs over all the sensors and turns back at the other end, so that the heated coolant leaves on the same side. Then two flex print cables connect the two halves of the ladder to the connector.



Figure 3.5: A prototype of the cold-plate for cooling. One coolant tube(golden) is connected to the cold plate(black) and turns 180° on the other end (not shown) so that the coolant flows in both directions and thus leaves on the same side it starts.

For layer 3 instead, only one flex print cable in the backward side is considered, in order to leave more space in the forward for other possible services and accelerator components. As mentioned above, for the third layer two different solutions are under study: at radius of 39 mm e 69 mm respectively. In Figure 3.6 are displayed schematic examples of some hypotheses.

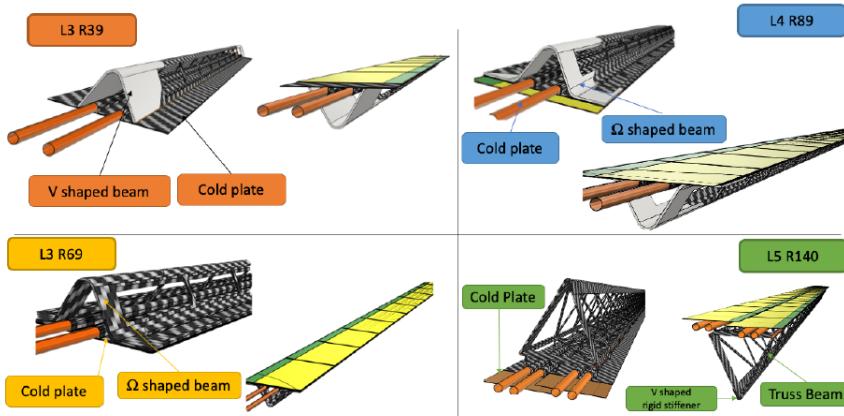


Figure 3.6: Schematic view of possible solutions for the three outermost layers.

In Figure 3.7 are shown the several substructures described before, that shape a ladder of the outermost layer 5. From bottom to top come in succession the carbon fibre structure, two cold-plates for the two neighbouring sensor rows (indicated as "Chips", in grey) and the flex cables for power and data transmission (green).

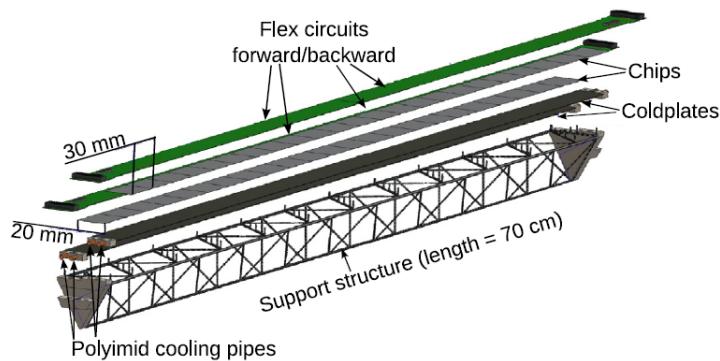


Figure 3.7: An exploding drawing of a fully assembled layer 5 ladder.

Research and Development

The carbon fiber support structure and flex cable have been designed and fabricated for the last ladder, which is also the longest. Services for the last two ladders, like electrical connections and cooling, can be provided both on forward

and backward sides. A Multiline Power Bus has been realized in order to power each OBELIX chip along the ladder by a dedicated VDD and GND pair.

After the assembly described in the previous, first thermo-mechanical tests have been performed and they show that the first resonance frequency is at 200 Hz, which is safely far from the one of the typical earthquakes in Japan, and also that the thermal properties are good.

Trying to reduce as much as possible the material budget, the transmission lines and the flex cables have to be as thin as possible, but they also have to ensure safe data transmission. Trace widths are trimmed to fulfill the same maximum voltage drop requirement (200 mV) for all the chips.

For this reason, the outermost ladders (70 cm long) are equipped with two flex cables, one from each side of the *truss*. In Figure 3.8 the resulting eye diagram from testing the signal integrity of one of the 35 cm long transmission lines for data transmission rate of 500 Mbps. This result demonstrates that the bandwidth is large enough to allow more than needed 160 Mbps for data transmission.

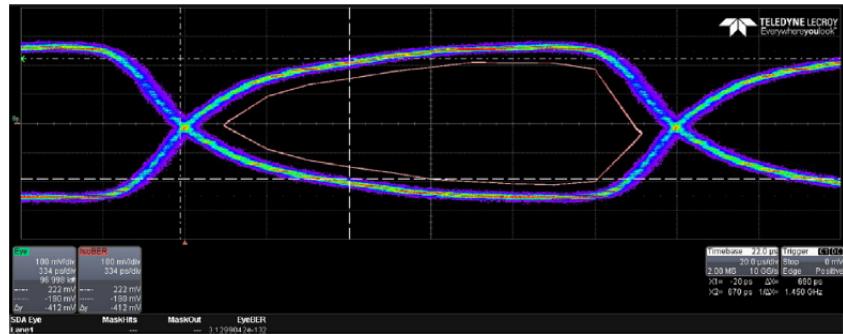


Figure 3.8: Eye diagram for the oVTX transmission line signal integrity of the layer 5 flex cable.

In addition, thermal tests have been performed for the last layer prototypes using kapton heaters to emulate the power dissipation of the chips. The coolant (demineralized water) has been set to a temperature of 10° at the beginning, the environment at 20° with a negative pressure of 0.2 bar. Results have been demonstrated that for three different configurations of the flow (such as monodirectional, bi-directional and with an U-turn at one end) the average temperature stands at 24° with a maximum gradient of $\Delta T \approx 4^\circ$ along the full length of the ladder.

All these investigations validate the design of the longest ladder, which is indeed the most challenging, and therefore the possibility to operate the chips safely.

3.1.3 Thermomechanics and data transmission

The proposed VTX detector intends to employ the same sensor type for all the layers in order to use a unique control and power supply system. It is expected

to operate at room temperature and for what we have seen in the previous, the smaller cross section of data cables, the usage of optical fibers and the less complex cooling system might allow a considerable reduction of services with respect to the current VXD. This allows more room for maneuver in the design of the new IR, needed for ramping up the instantaneous luminosity in the future.

As consequence also the design of the mechanical support system, data cables and acquisition system required could be more simple and in particular, the standard PCIe40 acquisition boards used in Belle II match well the data throughput requirement.

3.2 Performance simulation

As we have seen in Chapter 2, increasing the luminosity implicates higher level of machine related background and so larger doses of radiation, especially in the inner layers of the whole detector. For these reasons a lot of simulations and studies are focusing on ensuring that the main physics goals of the experiment will be achieved despite the more severe working conditions.

The VTX upgrade program which provides for a new silicon vertex detector with high granularity in both space and time, could bring significant improvements in tracking efficiency especially at low momentum, in the impact parameter resolution and in the robustness against backgrounds. Moreover, better vertexing performance entails not only improved time-dependent analyses of B and D mesons, but also an enhanced capability to distinguish among different decay topologies.

3.2.1 Potential VTX geometries

As we have already mentioned in the description of the oVTX layout (section 3.1.2), two different VTX geometries are currently under study, which differ only in the position of the third layer (Figure 3.6).

The *nominal* geometry is expected to maximize the track impact parameter resolution and it places the third layer at 39 mm from the IP. The *alternative* geometry instead, aims to improve the K_S^0 reconstruction efficiency and the third layer is located at 69 mm from the IP.

Several simulations and investigations are ongoing and they are comparing performance of these two different layouts with that of the current Belle II detector (utilizing a full Geant4 simulation of the detector in the study of specific decay modes of interest). Moreover, the different machine background predictions are also examined in order to consider all effects and correlations.

3.2.2 Tracking efficiency at low momentum and impact parameter resolution

Tracking efficiency at low momentum is one of the areas where the VTX upgrade scenarios show more promising results, particularly for the *soft pions* originated from the decays of $D^{*\pm}$ mesons.

Studies are based on the reconstruction of the decay chain $B^0 \rightarrow D^{*-} l^+ \nu$, with $D^{*-} \rightarrow \bar{D}^0 \pi_{soft}^-$ and $\bar{D}^0 \rightarrow K^+ \pi^-$ or $K^+ \pi^- \pi^+ \pi^-$. All background sce-

narios metioned in section 2.1.2 are considered in the evaluation of the *nominal* VTX performance and they are compared with the nominal Belle II geometry in the intermediate (**v2**) background hypothesis. In Table 3.1 is shown what has been obtained.

	Belle II (v2)	VTX (v1)	VTX (v2)	VTX (v3)
Generated events	32533	32559	32559	30255
Correctly reconstructed signal	10059	16913	16848	15583
Combinatorial	28495	51375	51826	47527
Efficiency	30.9%	51.9%	51.7%	51.5%
Purity	26.1%	24.8%	24.5%	24.7%

Table 3.1: Reconstruction efficiency and purity for the the decay chain $B^0 \rightarrow D^{*-} l^+ \nu$, with $D^{*-} \rightarrow \bar{D}^0 \pi_{soft}^-$ and $\bar{D}^0 \rightarrow K^+ \pi^-$, for the nominal Belle II detector at the intermediate background conditions (**v2**) and the nominal configuration of VTX in all three background scenarios.

We can see that the VTX reconstruction efficiency¹ in all three background hypotheses, results to be improved of almost a factor 1.7 with respect to the nominal Belle II, with comparable purity². Moreover efficiency remains practically stable in all background conditions, even in the most severe one.

This enhancement in tracking efficiency relies in particular on improved tracking efficiency for the π_{soft}^- mesons, as we can see in Figure 3.9.

For all scenarios with nominal VTX, there is a powerful improvement for transverse momenta below 0.05 GeV/c, with respect to the nominal Belle II. Only for p_T greater than 0.2 GeV/c the reconstruction efficiency of Belle II approach those of the VTX.

For p_T higher than 0.3 GeV/c instead, the momentum resolution is dominated by the CDC and there is not a significant enhancement in track momentum resolution considering the preceding example.

3.2.3 Vertexing resolution

Studies on vertexing performance have been conducted using samples of one million $B^0 \rightarrow J/\psi K_S^0$ events generated and reconstructed with all the aforementioned combinations. The distributions of the decay vertex resolution σ_z (i.e. the width of the distribution obtained considering the differences between the measured and the true simulated positions) along the z axis of the B decay signal are shown in Figure 3.10 .

In Table 3.2 a summary of the results that shows that the new geometries achieve a better resolution on the B decay vertex of about 35% on average and they also do not suffer of any significant degradation as the background conditions varies, unlike the nominal Belle II configuration.

The VTX geometries have been demonstrated to achieve better performance also for the vertexing resolution along the x and y axes. Also in these direc-

¹Efficiency is defined as the ration between the number of correctly reconstructed signal events and the total number of candidates.

²In a few words, the probability that a correctly reconstructed signal is a "signal event".

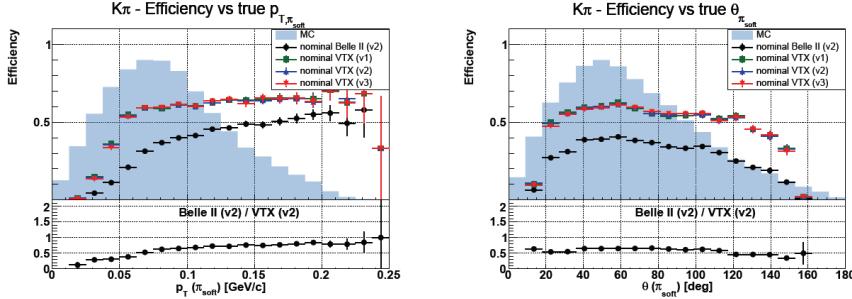


Figure 3.9: Reconstruction efficiency of $B^0 \rightarrow D^{*-} l^+ \nu$ as a function of the transverse momentum of the π_{soft}^- (from $D^{*-} \rightarrow \bar{D}^0 \pi_{soft}^-$) in the plot on the left and of the polar angle of the π_{soft}^- on the right. The shaded blue histograms represents the momentum spectrum of the π_{soft}^- . The nominal Belle II geometry efficiency in the intermediate background scenario (**v2**) is represented by black dots and it is compared with the nominal VTX configuration in the optimistic (**v1**, green squares), medium (**v2**, blue upward pointing triangles) and pessimistic (**v3**, red downward pointing triangles) background hypotheses. The bottom plots show the ratio between nominal Belle II and nominal VTX in the **v2** background scenario.

B_{sig} z vertex resolution (μm)	Bkg (v1)	Bkg (v2)	Bkg (v3)
Belle II	21.9	23.0	24.9
Nominal VTX	14.5	14.4	14.1
Alternative VTX	14.4	14.3	14.0

Table 3.2: B_{sig} vertex resolution along the z axis for the three detector layouts and the three background scenarios.

tions they turn out to be almost insensitive to the different levels of background.

Similar studies for the K_S^0 decay vertex resolution are displayed in Figure 3.11 and in the same way, the upgraded geometries reach a better vertexing resolution with respect to the nominal Belle II detector without any significant degradation as the backgrounds increase. It is important to notice that in the right plot there is a spurious effect that cause an apparent improvement of performance in the worst case scenario among those considered. This is due to the loss of reconstruction efficiency for candidates with large flight distance (thus affected by poorer vertex resolution).

3.3 OBELIX chip design

The VTX detector is designed with a single type sensor taylored to the specific needs of Belle II, called OBELIX (Optimized BELle II pIXel sensor) and currently under development, based on fast and high granular Depleted Monolithic Active Pixel Sensor (DMAPS). This new sensor design comes from an evolution

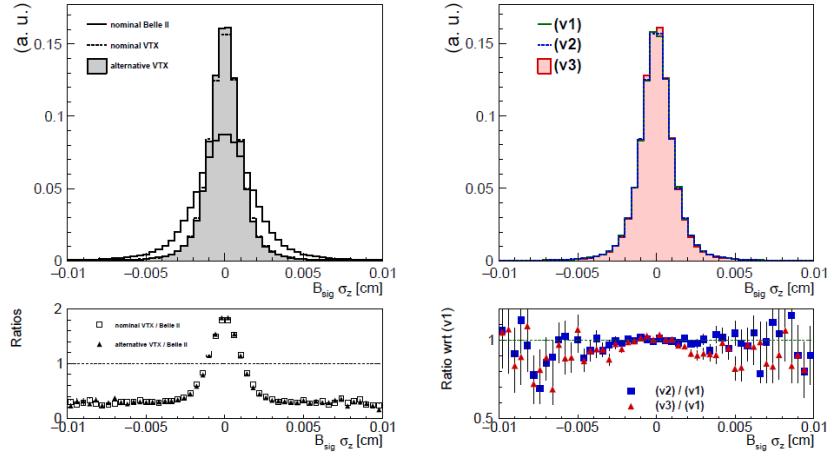


Figure 3.10: On the left: comparison of the B decay vertex resolution along the z axis in $B^0 \rightarrow J/\psi K_S^0$ events for the nominal Belle II (solid line), nominal VTX (dotted line) and alternative VTX geometry (filled grey histogram). The bottom plot shows the ratio between the VTX geometries (empty squares the nominal one and filled triangles the alternative) and nominal Belle II. On the right: B decay vertex resolution along the z axis for the nominal VTX geometry in the three background scenarios: optimistic **v1** (green solid line), intermediate **v2** (blue dotted line) and pessimistic **v3** (red filled histogram). The bottom plot represents the ratio between the two higher background scenarios and the optimistic one.

of TJ-Monopix 2, whose characterization is the main topic of this thesis, and which will be discussed in Chapter 5. Both of them are fabricated in a modified TowerJazz Semiconductor 180 μm CMOS process, that matches particularly well the Belle II requirements.

In particular TJ-Monopix 2 is equipped with four different flavors (section 5.1.1), which stand out for different type of collection electrode coupling and some small differences in the circuit design. For now their characterization are ongoing and the final decision on which to use for OBELIX has not been made.

3.3.1 Sensor specification

A schematic layout of the chip is shown in Figure 3.12 . The size of the sensor is expected to be $3 \times 1.9 \text{ cm}^2$, with an active area of $3 \times 1.6 \text{ cm}^2$ and an additional part in the periphery of about $3 \times 0.3 \text{ cm}^2$, dedicated to data pre-processing and triggering. The pixel pitches³ are designed to be from 30 μm to 40 μm in both directions. As a matter of fact staying in this range is necessary in order to achieve a spatial resolution below 15 μm , which is a requirement of the VTX upgrade program.

Moreover the sensor thickness has to be below 100 μm to respect the material budget constraint of 0.2% X_0 and as consequence the depleted sensitive region

³The distance between the centers of two contiguous pixels.

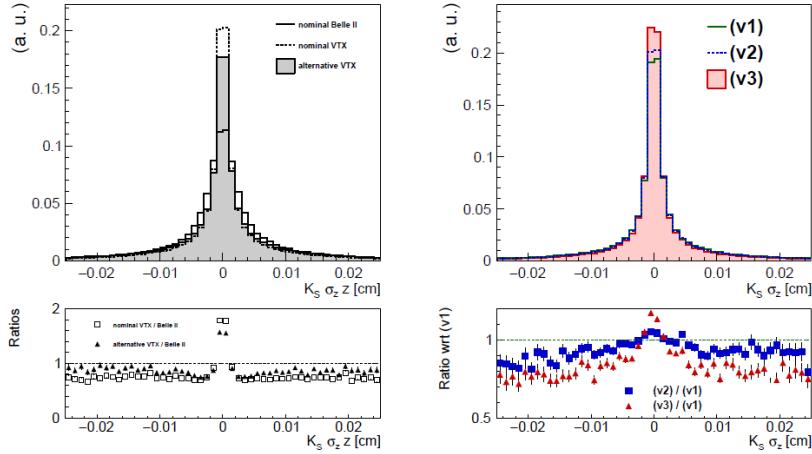


Figure 3.11: On the left: comparison of the K_S^0 decay vertex resolution along the z axis in $B^0 \rightarrow J/\psi K_S^0$ events for the nominal Belle II (solid line), nominal VTX (dotted line) and alternative VTX (filled grey histogram). The bottom plot shows the ratio between the VTX geometries (empty squares for the nominal and filled triangles for the alternative) and nominal Belle II detector. On the right: K_S^0 decay vertex resolution along the z axis for the nominal VTX in the three background scenarios: optimistic **v1** (green solid line), intermediate **v2** (blue dotted line) and pessimistic **v3** (red filled histogram). The bottom plot represents the ratio between the two higher background scenarios and the optimistic one.

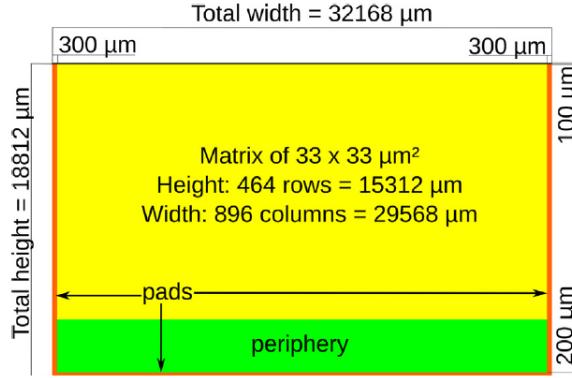


Figure 3.12: OBELIX chip design.

should be lower than 50 μm , perfectly in agreement with the available thickness of MAPS technology. To deal with the target hit rate of 120 MHz/cm^2 , the timestamp clock signal can reach down to 25 ns, even if studies have demonstrated that a window of 100 ns (*integration time*) is enough to limit to 320 Mbps the data throughput at the same expected hit rate. All characteristics inspected above allow to realize a sensor with high granularity in both space and time.

With respect to TJ-Monopix 2, which is equipped with a triggerless column-drain readout without memory at the periphery, OBELIX must have a triggered readout architecture, in order to satisfy the needs of Belle II. Moreover the latency is fixed to 5 (10) μ s and it might operate up to 30 KHz trigger rate.

Single Event Upset section 1 is a concern for the operation of the future detector but its size has not yet been quantified. Therefore an important feature of the chip must be to ensure that the control system is able to reset the sensor registers to default operational values at least every minutes. The reset frequency will be chosen after the measurement of the SEU cross section with OBELIX and the comparison to the occurrence distribution of large energy loss in the experiment.

The expected power consumption instead, is expected to be about 200 mW/cm⁻², a value which should allow air-cooling for the small areas corresponding to the two inner layers and liquid coolant for the outer ones.

Its main design features are summarised in Table 3.3.

Pixel pitch	30 to 40 μ m
Matrix size	512 rows \times 928 to 752 columns
Time stamping	25 to 100 ns precision over 7 bits
Signal Time over threshold	7 bits
Output bandwidth	320 to 640 Mbps
Power dissipation	100 to 200 mW/cm ²
Radiation tolerance	100 MRad and 10^{14} n _{eq} /cm ²

Table 3.3: Designed features of the OBELIX sensor.

3.3.2 Sensor implementation

As mentioned above, this new sensor is the development of TJ-Monopix 2, whose characteristics fit already the Belle II requirements (Table 3.4).

	Specification	TJ-Monopix2
Pixel pitch	< 40 μ m	< 33 μ m
Sensitive layer thickness	< 50 μ m	30 μ m and 100 μ m
Sensor thickness	< 100 μ m	-
Hit rate capability in the matrix	> 600 MHz cm ⁻²	> 600 MHz cm ⁻²
Hit rate capability at the sensor output	> 120 MHz cm ⁻²	\gg 100 MHz cm ⁻²
Trigger delay	> 10 μ s	-
Trigger rate	30 kHz	-
Overall integration time	< 100 ns	-
(optional) Time precision	< 50 ns	-
Total ionizing dose tolerance	100 kGy/year	1 MGy/year
NIEL fluence tolerance	5×10^{13} n _{eq} /cm ² /year	1.5×10^{15} n _{eq} /cm ² /year
SEU tolerance	frequently (min^{-1}) flash configuration	-
Matrix dimensions	around 30 \times 16 mm ²	19 \times 19 mm ²
Overall sensor dimensions	around 19 \times 19 mm ²	20 \times 19 mm ²
Powering	through voltage regulators	-
Outputs	one at < 200MHz	one at 160 MHz

Table 3.4: Comparison between OBELIX requirements and TJ-Monopix 2 features.

From TJ-Monopix 2 design, the pixel size of $33 \times 33 \mu\text{m}^2$ is maintained, as well as the layout of both digital and analogue parts (REFERENCE TO

FOURTH). Also the Time-Over-Threshold method to digitize the signal is preserved, with a bus width of 7 bit, together with the column-drain readout architecture implemented for pairs of columns. Other features, which will be explained in depth in Chapter 5, have been conserved in the new design like the 3-bit register dedicated to the threshold tuning, but with a larger range of correction for the last bit. Moreover to aim at the integration time of 100 ns, the clock frequency which defines the precision of ToT and BCID (that is the timestamp), has been decreased from 40 to 20 MHz. So the current baseline for OBELIX timestamp precision is 50 ns.

Additionally two new modules have been added to the implementation, related to the Belle II trigger: the Trigger Logic Unit (TRU) and the Track Trigger Transmitter (TTT).

Trigger Logic Unit (TRU)

The TRU has the task to select the fired pixel information from the matrix which are in-time with the triggers sent by the Belle II system. In more details, this module employs two stages of memory in order to manage the data coming from the pixel matrix (Figure 3.13). These components are designed in order to minimize power dissipation and to optimize the efficiency even in severe operating conditions: maximum hit rate of 120 MHz cm^{-2} , 30 KHz of trigger rate and $10 \mu\text{s}$ of trigger delay.

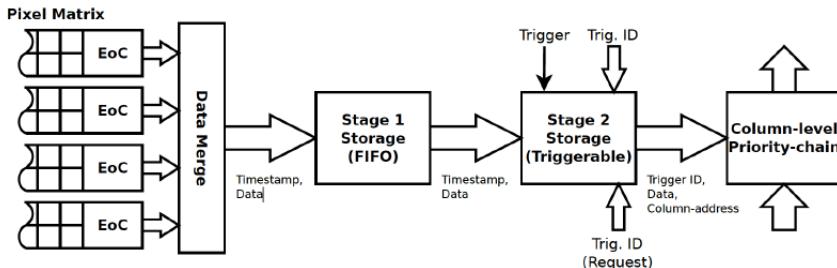


Figure 3.13: Schematic of the Trigger Logic Unit.

The first stage has to store the pixel information during the trigger delay. The second memory instead, has the function to compare the BCID of the fired pixel with each trigger time information buffered in a dedicated global memory which keeps track of received triggers. When they have a match, the pixel data is transferred to the Transmission Unit (TXU). In this way, the physics hits associated to a trigger but timestamped with a later BCID, for example due to timewalk effect, are also considered for further analysis. Considering the BCID precision time, the time integration of the OBELIX sensor becomes 100 ns.

Track Trigger Transmitter (TTT)

The TTT module divides the matrix in 8 logic regions (this value is still under study) and generates a one-byte word depending on the region that is fired. It is expected that this information could be transmitted to trigger system within 100 ns and along a line of transmission parallel to the main data output of the

sensor. This component behaviour is still under study and it needs of further simulations in correlation with the whole VTX system.

Control Unit (CRU) and power dissipation

The OBELIX sensor, as well as TJ-Monopix 2, is configured by several registers which allow to set important features for its operation like threshold settings, masked pixels, time response of the pixels, but they also define its power consumption. The Control Unit is responsible for receiving these instructions about the configuration and the trigger information and at the same time sending out data coming from TXU module.

For what concern power dissipation, there are three main features which have the greatest impact: the biasing current flowing into the in-pixel amplifier (I_{BIAS}), the BCID clock frequency (on which depends the timestamping precision) and the hit rate. In Figure 3.14 is shown the estimations of power dissipation as these parameters vary.

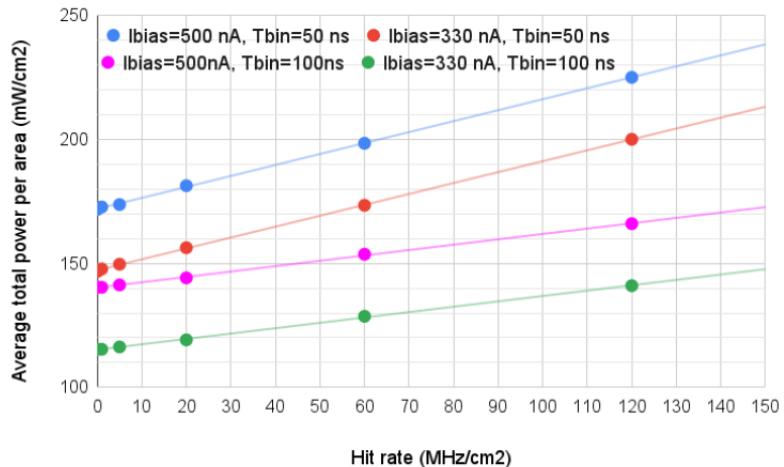


Figure 3.14: OBELIX sensor power dissipation depending on the front end current (I_{bias} , the BCID frequency ($tbin$) and the hit rate).

As we can see, the power consumption at the maximum hit rate of 120 MHz cm^{-2} exceeds by little more than 10% the power budget of 200 mW cm^{-2} , considering the higher precision for the timestamp of 50 ns. Therefore to stay within the power budget it is necessary to find a compromise: reducing timing precision by worsening the BCID precision to 100 ns or decreasing the pre-amplifier biasing current causing a degradation of the time walk.

The first version of the sensor, called OBELIX-1, is being designed and the submission for fabrication is planned in the last month of 2023. A second improved version, OBELIX-2, will be designed based on performance studies on the first version and it is expected that it will be the final sensor needed for the experiment.

In this chapter we have introduced some of the fundamental aspects of the proposed VTX upgrade, supported by continuous studies and simulations. In the following we will see in more details the technology on which the whole proposal is based: the CMOS Monolithic Active Pixel Sensor.

4. CMOS MAPS sensors

The fourth chapter aims to introduce the essential features of the semiconductor detector technology, going through the history of its advancements, which have led to the currently most promising sensors based on CMOS logic structure, the Monolithic Active Pixel Sensors (MAPS). As we have seen in the previous chapter, the VTX program wants to make the most of the technologies that have already proven reliable in precision measurements, like the TJ-Monopix development line, with also fast readout and high radiation tolerance. We will briefly present it, mentioning the peculiarities of its prototypes, to better understand how they could fulfill the Belle II requirements.

4.1 Semiconductor detectors

The detection of elementary particles, nuclei and radiations occurs through their interactions with matter. In particular, charged particles are often detected by atom ionization or excitation of the sensitive detector layers along the path they take passing through. This detection method can be used in gases, liquids and semiconductors. We want to focus on detectors that use semiconductors as sensitive material.

All solid can be divided into three categories based on their electrical conductivity: conductors, semiconductors and insulators. In a solid state lattice, the constituent atoms have a dense periodic arrangement and for this reason, the energy of single atoms are split due to the influence of other atoms in the vicinity. The energy levels of some level groups lie energetically so dense (order of meV) that one speaks of *energy bands*, separated from each other by a *band gap*, which represent the distance in energy between them (E_G , energy gap). The electrical conduction properties of materials are determined by the two highest energy bands, which are the **valence band (VB)** and the **conduction band (CB)**. The energy levels within the same band are so dense that the transitions to unoccupied levels, if they are not completely filled, are easily possible. Therefore the electrical conduction properties depend on the band gap between the two levels and on the band occupation, as we can see in Figure 4.1.

In insulators the valence band electrons are strongly bonded to neighbouring atoms, they are not free and they do not contribute to the conduction. In fact the VB is entirely occupied, the CB instead, is empty. As a consequence of the strong interatomic bond, there are large energy gap between the VB and the CB (typically E_G of about 9 GeV). Thus current flow is practically impossible.

In semiconductors weaker bonds between neighbouring atoms result in a

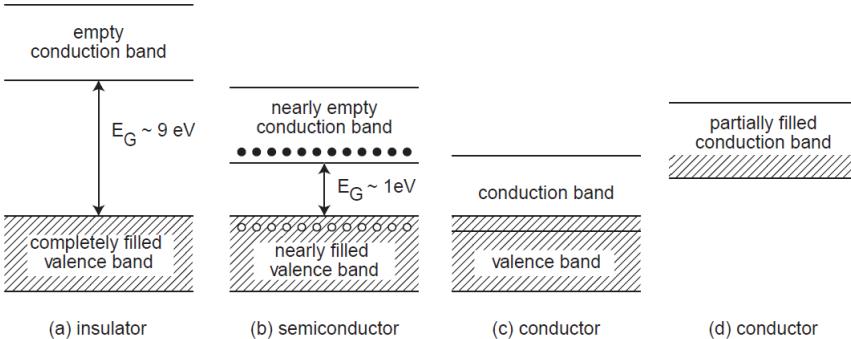


Figure 4.1: Schematic structure of the energy bands in insulators (a), semiconductors (b) and conductors (c,d).

smaller energy gap with respect to the insulator (for example 1.12 eV in silicon). In this way electrons from VB can easily overcome the gap, moving on the CB by thermal excitations or by external electric fields. When an electron makes this transition, it leaves a hole in the VB, which could be filled in turn by another electrons of the VB. If an electric field is applied, the free electrons in the CB and the holes in the VB start to move producing two different current flows, one negative and the other positive, respectively.

In conductors the conduction and valence band overlap or the conduction band is partially filled so transitions within the same band and between the two different bands are easy and current conductance requires minimal energy.

4.1.1 Movement of charge carriers and signal formation in semiconductors

Semiconductor materials are the only ones that allow the detection of charged particles by ionization of the sensitive matter (in conductors a current is always present, not only in case of particle or radiation crossing).

When a charged particle passes through the medium, it releases a certain amount of energy[, described by the *Bethe-Bloch formula*]. This energy loss, in turn, causes the ionization of the matter and so the formation of positive and negative charges, which in this context, are defined charge carriers. In semiconductors these carriers are the hole-electron pairs created by ionization, which start to move in opposite directions due to an electric field: the holes (positive) migrate towards the **anode** and the electrons (negative) towards the **cathode**, which sense the signal induced by this movement. In fact, their drift induces an accumulation of charges on the electrode surfaces, and it is possible to record this charge induction as a charge, current or voltage signal. It is worth noticing that the generation of the signal is dependent on the movement of the carriers relative to the electrodes, and not when they actually arrive, that is the moment in which the signal stops.

The charge carrier density of semiconductors can be modified by doping the material with specific chemical elements, and this process causes a modification of their conduction properties. Undoped semiconductors are called *intrinsic*

semiconductors.

Extrinsic semiconductors instead, are artificially doped with external impurities like:

- Pentavalent elements (P, As, Sb), called *donors*, added in a tetravalent material (Si, Ge) produce an excess of conduction electron with respect to the holes (n doping, Figure 4.2 (a)).
- Trivalent atoms (B, Al, Ga), called *acceptors*, create an excess of holes (p doping, Figure 4.2 (b)).

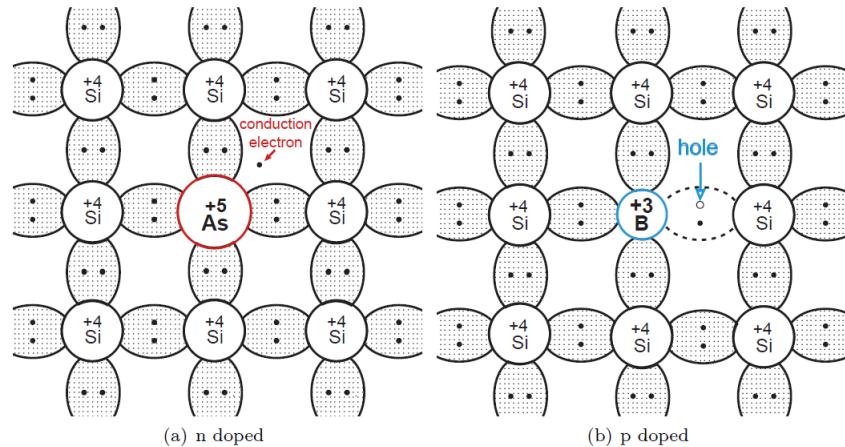


Figure 4.2: Schematic representation of atom bonding structure in n-doped and p-doped semiconductors.

4.1.2 The pn junctions as detector

The base material employed in semiconductor detectors is silicon. When a p-doped semiconductor get in contact with a n-doped material, a *pn junction* is formed. In particular, the p-doped part is the one where holes are the dominant charge carriers, called **majority carriers**; in the n-doped part of the crystal instead, the majority carriers are the electrons. The presence of these excesses of opposite charge in the two part of the junction, causes a migration of the majority carriers from each part to the opposite one. At the boundary the charges recombine, and this process creates a zone which is free of charge carriers, called **depletion zone**. After the recombination the atoms of this depletion region are ionized and so it is no longer neutral, but features a *space-charge*: a positive one in the n-layer, and a negative in the p-layer. Moreover these space charge densities are opposite in sign, so they generate an intrinsic electrical field that stop the original diffusion.

Moreover, the application of an external voltage V_{ext} between the two section of the junction, provokes a variation of the width of the depletion region, depending on the size and polarity of the applied voltage. It is possible to distinguish:

- *forward bias*, $V_{ext} > 0$: a positive external voltage applied to the p side with respect to the n side, causes a reduction of the depletion region ???
- *reverse bias*, $V_{ext} < 0$: if a negative voltage at the p side or positive at the n side relative to the respectively opposite side is applied, the depletion region gets wider.

The performance of the junction as a detector is determined above all by the boundary properties between p-and n- doped silicon. Boundaries of the same doping type, but with different concentration, also create a similar pn structure, like for example n^+n or p^+p . There are also Metal-semiconductor boundaries, used in metal contact with the outside and Metal-Oxide-Semiconductor interfaces (MOS) that we will see in more details in the following.

4.1.3 Complementary Metal-Oxide-Semiconductor (CMOS)

A MOS structure is a double interface made of three different media: an insulator (oxide) placed between the metal and the semiconductor, as shown in Figure 4.3.

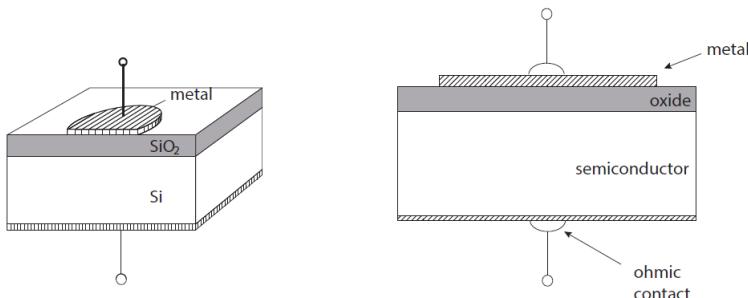


Figure 4.3: A perspective (on the left) and a cross-sectional illustration (on the right) of a MOS structure.

In recent transistor technology, the metal is almost entirely replaced by highly doped polysilicon, like n^{++} and p^{++} since these materials are more resistant to high temperature without reacting with the oxide (?). Despite this, the physics remains essentially the same. The MOS structure plays a very important role in chip electronics, including the readout of detectors, which employ a combination of NMOS and PMOS transistors, embedded in the substrate. This is at the base of the *Complementary* MOS (CMOS) electronics, which allows to develop more complex circuits. This technique consists in accomodate one of the two transistor type in a differently doped area, called *well*. For example, in a p-type substrate, a PMOS transistor are accomodate in n-wells, and vice versa with a n-type substrate.

4.2 Hybrid and monolithic pixel sensors

4.2.1 Hybrid

4.2.2 Pixel

4.3 CMOS Monolithic Active Pixel Sensors technology

4.4 History of Monopix developments

5. TJ-Monopix 2

In the previous chapter we have seen the fundamental steps that had lead to the development of the CMOS MAPS sensors technology and the history of their many different prototypes.

Here we will go through the main features of TJ-Monopix 2 designed to address efficiency degradation after irradiation, one of the main issues of its predecessor Tj-Monopix 1. The characterization of the chip has crucial consequences in the VTX upgrade program and therefore in the evolution of the next OBELIX chip. The chip W14R12 (figure 5.1 on the following page) which is one of the matrices tested during the Test Beam in Desy (July 2022) has been fully characterized in Pisa and in particular several aspects have been analyzed, among which:

- TOT calibration by internal charge injection;
- characterization with radioactive sources;
- systematic study of different registers' settings in order to operate the chip at low thresholds;
- investigation of an important issue with cross-talk, due to digital signal from the readout, discovered trying to operate the matrix at low threshold (below $250 e^-$).

This detailed analysis returned relevant results which have helped in Test Beam data reconstruction and in the simulations of the upgraded VTX detector with CMOS MAPS devices.

5.1 Matrix and flavors

Tj-Monopix 2 is the next generation small collection electrode DMAPS prototype in TowerJazz 180 nm. The need to create a sensor capable to maintain high efficiency even after irradiation, required improvements compared to Tj-Monopix 1 in two important fields: a lower operating threshold and different pixel layout to increase charge collection efficiency all over its area, especially in the corners.

To achieve these goals, a different in-pixel front-end circuit was implemented and a lot of efforts have been focused on optimizing pixel layout in order to reduce its size, which has been decreased to $33.04 \times 33.04 \mu m^2$ (pixel *pitch*). As a matter of fact we have seen (REFERENCE) that pixel dimensions are critical to accomplish faster charge collection across all active area, increasing

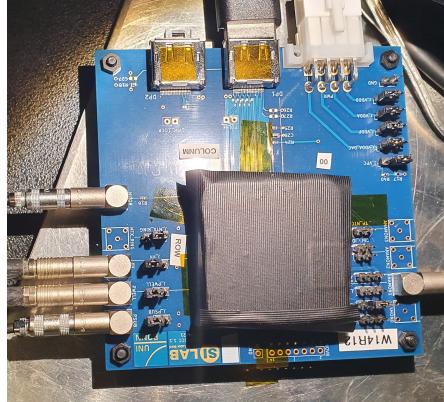


Figure 5.1: The W14R12 chip tested during the Test Beam in Desy.

the lateral electric field. For this reason it was necessary a special effort to design and create a smaller pixel but still adequate to embody the full digital readout. All of this required to work at the technology density limit and also to study for further modifications at the circuit design.

The BCID bus width has been increased to 7-bits because of higher gain and higher ToT slope with respect to Tj-Monopix 1.

It's worth mentioning here that both the large column height (≈ 17 mm) caused by large matrix area and the aggressive column-bus routing (which refers to the minimum line width and spacing) due to the smaller pixel size, generate a significant signal transmission delay due to the RC low pass filtering effect of the long metal wires. Consequently a special circuit has been studied that adds a variable delay to the hit pulse across the column that matches that of the BCID signal.

5.1.1 Flavors

The prototype is a $2 \times 2 \text{ cm}^2$ pixel matrix which consists of 512x512 pixels and all of them are designed with a reduced deep p-well geometry (RDPW) because, as it was demonstrated during the testing of TJ-Monopix1, this type of arrangement has superior charge collection properties compared to full deep p-well coverage (FDPW) (figure 5.2 on the next page). The total active area of the matrix is approximately 286 mm^2 and it is $300 \mu\text{m}$ thick.

As we can see in figure 5.3 on the facing page, the matrix is divided in four sectors, named **flavors** that implement different variation of the front-end circuit. In the first two flavors the collection electrode is DC-coupled directly with the readout electronics, the continuous baseline reset is implemented by a forward bias diode and they differ for the pre-amplifier circuit design. The second flavor indeed, named **Cascode FE**, includes an extra-cascode transistor that increases the pre-amplifier gain which in turn leads to a 50% reduction of the threshold dispersion compared to the first flavor, the **Normal FE**. The other two flavors consist of AC-coupled pixels (through a metal-oxide-metal MOM capacitor) and in particular, the **HV-Cascode FE** also incorporates the aforementioned pre-amplifier variation. AC-coupling allows to apply an high

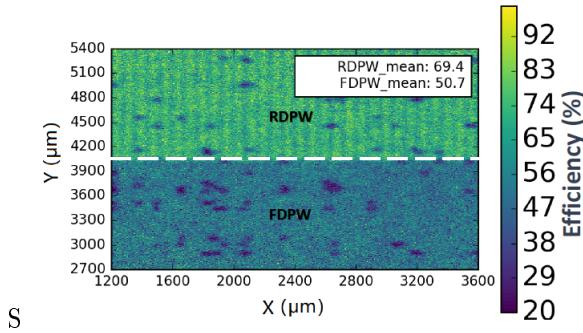


Figure 5.2: Detection efficiency map of a TJ-Monopix1 chip with $25 \mu\text{m}$ p-epitaxial layer that has been irradiated to $10^{15} n_{eq}/\text{cm}^2$ NIEL.

positive bias voltage (HV stands for High Voltage indeed) to the collection electrode which potentially enlarges the depleted region, but at the same time it also causes signal losses mainly due to the additional parasitic capacitance introduced at the sensitive input node.

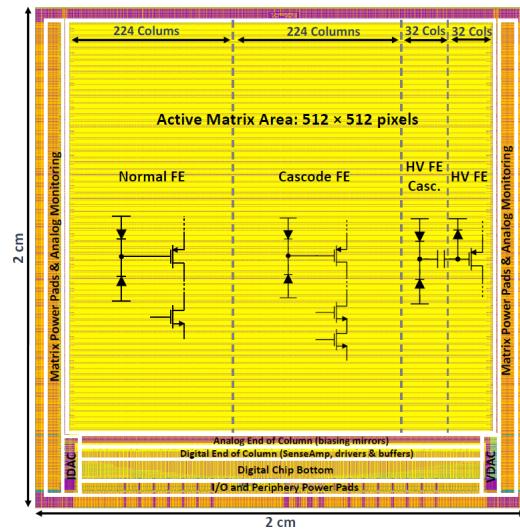


Figure 5.3: The layout of the TJ-Monopix2 prototype divided in four different flavors: **Normal**, **Cascode**, **HV-Cascode** and **HV FE**.

5.1.2 Pixel design

The 2×2 pixel core layout, shown in figure 5.4 on the next page, is fully optimized and is designed in order to share as much features as possible between the four pixels. The analog area incorporates the front-end circuit, the 3-bit threshold tuning DAC and the pixel configuration registers. The digital region is composed by the 7-bit LE and TE memory (14 SRAM cells per pixel), the 10 bit address

ROM (2 bit for the pixel position inside the core and 8 for the group address), the readout control logic and the hit delay circuit that is used to correct the BCID propagation delay. Two different token signals are used to set the priority of the pixels during the readout: the *fast* one that propagates across the double column establishing the priority between the cores and the *local* one, which arbitrates the reading order of the four pixels inside each core.

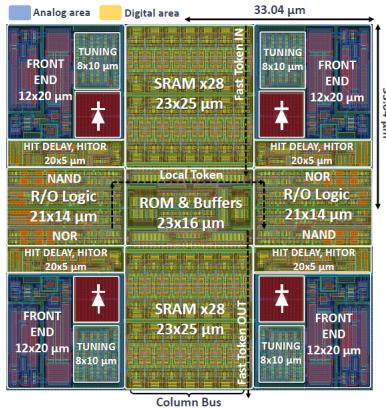


Figure 5.4: Layout of a TJ-Monopix 2 2x2 pixel core. In blue the analog area and in yellow the digital one.

5.1.2.1 Improved front-end circuit design

As we have seen above, there are two variations of the front-end circuit (figure on the facing page), such as *normal* and *cascode* type. The latter in particular includes an extra cascode transistor which increases the pre-amplifier gain and consequently reduces the threshold dispersion.

Moreover in TJ-Monopix 2 it was preferred to incorporate a simple diode to reset the input node instead of a PMOS transistor, which was the technology implemented in TJ-Monopix 1. A side effect of this choice is that the relationship between charge injected and the ToT of the detected signal is not linear anymore, because the diode is a not linear element and its discharge rate also depends on the collected charge. Indeed in the following analysis it was necessary to fit the ToT trend with a more complex function. But at the same time, the advantages are its simplicity (p^+ diffusion within the n-well collection electrode) and also the fact that it allows to increase radiation tolerance to TID effects, which was one of the key working area in the upgrade of the sensor.

In the last two AC-coupled flavors are implemented the same improvements, but here the different coupling provokes an important loss in the collected charge, as verified during the testing phase of TJ-Monopix 1 (50% losses), due at most to additional parasitic capacitances. Thus a lot of efforts have been made to improve this aspect, working on the coupling capacitor values. A signal loss of 41.5% has been reached in Tj-Monopix 2, which is a relevant enhancement with respect to its predecessor.

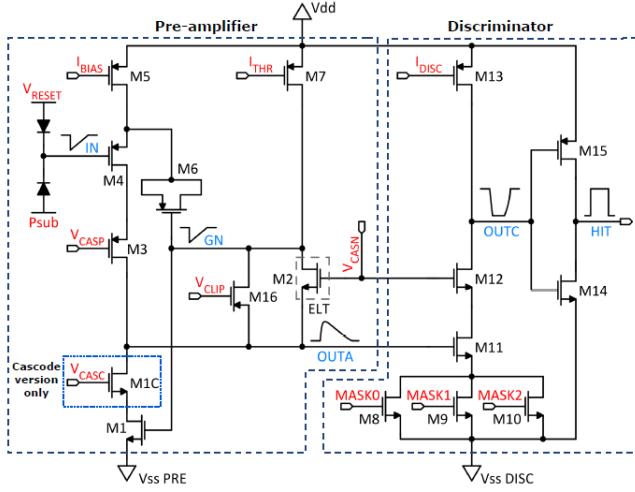


Figure 5.5: Schematic of the improved front-end circuit and its variation (extra-cascode transistor) in Tj-Monopix 2.

5.2 Threshold and noise

In order to achieve the absolute calibration of the whole matrix, the response of each pixel has been characterized by means of the internal charge injection.

The hit injection circuit included in TJ-Monopix 2 is similar to the one of TJ-Monopix 1. It allows to inject artificial hits through an injection capacitance C_{inj} connected at the collection electrode, which is equal to 230 aF for both the DC and AC coupling FEs. The injected charge is almost linear with the injection pulse amplitude (set by the two registers " V_L " and " V_H ", so that $\Delta V_{inj} = V_H - V_L$). Moreover the injection step is finer compare to the one of Tj Monopix 1 because of the higher voltage DAC resolution, which is indeed LSB (*Least Significant Bit*) = 7.03 mV. The injected charge resolution Q_{res} can be calculated by:

$$Q_{res} = Q_{inj} \cdot LSB = \frac{230 \text{ aF}}{q_{e^-}} \cdot \Delta V_{inj} = 1.4375 \frac{e^-}{mV} \cdot 7.03 \frac{mV}{DAC \text{ unit}} \approx 10.1 \frac{e^-}{DAC \text{ unit}} \quad (5.1)$$

Eventually this conversion factor has been used to convert the information of the injected charge from DAC unit to electrons unit, useful for further analysis. The four flavors have been separately analyzed to be able to study their main differences concerning their performance and features. The same method has been used for all of them, called *s-curve method* and explained below.

5.2.1 S-Curve method

In order to obtain the thresold and noise values for all pixels, each one of them has to be injected an arbitrary number of times (100 times in this work) for each value of the injection pulse between a minimum voltage value, chosen setting the

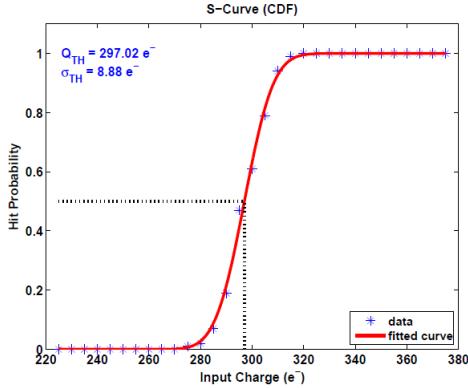


Figure 5.6: An example of the S-Curve fitted by the CDF to evaluate threshold and noise.

chip register "VL" and a maximum voltage set by the "VH" register, moving gradually with a step of 1 DAC unit (this is also adjustable). These two levels are provided by the voltage DAC.

So for each injection pulse height, the mean of 100 injection outputs are considered and it is represented as one marker in the plot. In this way, plotting the average number of detected hits in function of the injected charge, the typical curve better known as "*S-curve*" is reconstructed. It can be fit with the *Cumulative Distribution Function (CDF)*:

$$CDF(Q) = \frac{1}{2} \cdot \left(1 + erf\left(\frac{Q - \mu}{\sigma\sqrt{2}}\right) \right) \quad (5.2)$$

from which the value of the threshold is evaluated considering the value of the injected charge at half of the curve's maximum height. Specifically the half height corresponds to a charge value for which the pixel detects 50 hits of 100 injected and so when it has an occupancy of 0.5. This information is represented by the parameter μ obtained from the fit and the noise is evaluated from the fit parameter σ . The " $erf(x)$ " is the Gauss error function. In figure 5.6 is shown an example.

This method allows to study the noise and threshold of all pixels and also the threshold dispersion across an entire FE.

In the following are reported the results of this study for the flavors of all matrix, gradually injecting a charge from 0 to 140 DAC ($\approx 1414 \text{ e}^-$ adopting the conversion factor in equation 5.1)

5.2.1.1 Normal FE

The first flavor of the matrix is the **Normal FE**, which consist of 512 rows and 224 columns for a total of 114.688 pixels. The chip registers have been set with the same values used during the Test Beam at Desy (July 2022) which are different for the DC and AC-coupling case. They are called for simplicity "**GOE settings**" and they are reported in table 5.1 on the facing page, where the different biasing voltages used to power up the chip are also added.

Registers	Normal/Cascode FE ($P_{SUB}/P_{WELL} = -3$ V)	HV/HV-Cascode FE ($P_{SUB}/P_{WELL} = 0$ V, HV = +5 V)
I_{THR}	64	30
I_{BIAS}	50	60
V_{RESET}	143	100
I_{CASN}	0	8
V_{CASP}	93	40
V_{CASC}	228	228
I_{DB}	100	100
I_{TUNE}	53	53
V_{CLIP}	255	255
I_{COMP}	80	80
I_{DEL}	88	88
I_{RAM}	50	50

Table 5.1: Settings of the main registers used for all flavors (W14R12 chip) during the Test Beam in Desy.

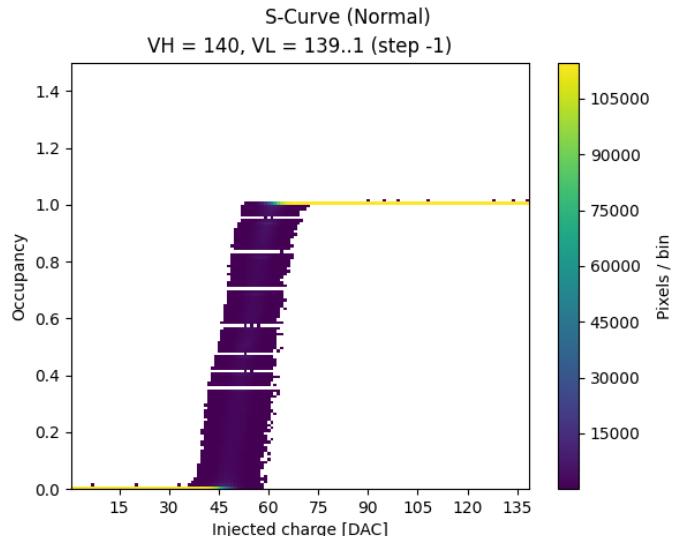


Figure 5.7: S-curves of all pixels of the Normal FE with a maximum injection pulse of 140 DAC.

In figure 5.7 are plotted all the s-curves of the all well-functioning Normal flavor pixels. The width of the figure is a first indication of the threshold dispersion of the whole flavor.

The threshold and noise distributions obtained injecting all pixels as explained above, have been fitted with a gaussian distribution and they are shown in figure on the following page with their maps, too.

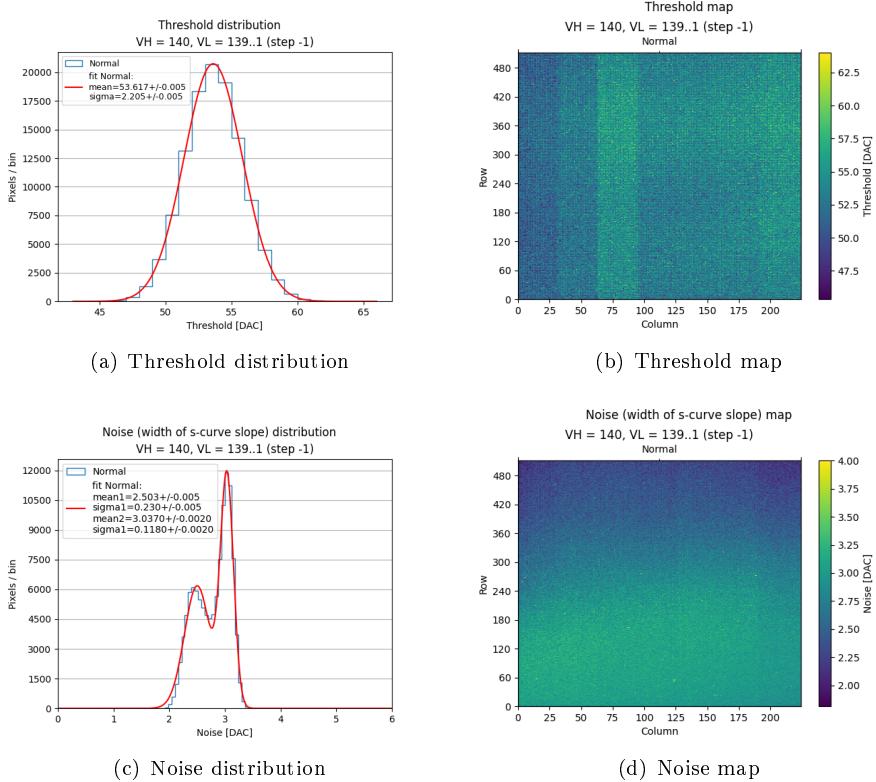


Figure 5.8: Normal FE.

5.2.1.2 Cascode FE

Cascode FE is the second flavor and like the previous one, it consists of 512 rows and 224 columns for a number of total pixels equal to 114.688. For this flavor the same procedure of Normale FE has been followed and also the same registers' values (table 5.1 on the previous page) have been used. In figure on the facing page the s-curves of all pixels are shown.

The maps and the fit of the threshold and noise distributions instead, are shown in figure 5.10 on the next page.

5.2.1.3 HV-Cascode FE

The third flavor is **HV-Cascode FE** and it is composed by 512 rows and 32 columns for a total number of pixel equal to 16384. Also for these last two flavors, the main chip registers are set with the same values tested and used during the Test Beam (@Desy) (but different from those used for the first two flavors). They are reported in table on the preceding page .

As we can see from the plot of the alle s-curves in figure on page 60, there were a lot of noisy pixels with this choice of register settings, but at this stage of measurements they were not masked. As a matter of fact along the y-axis of this plot is displayed the occupancy and when this value becomes higher than 1, it means that the pixel detects more hits than the injected ones, so it could

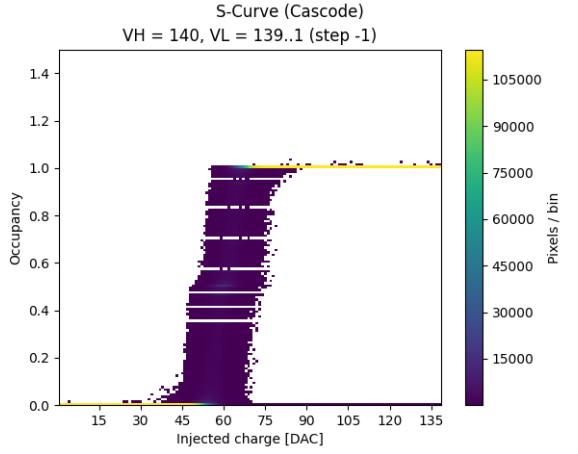


Figure 5.9: S-curves of all pixels in the **Cascode** flavor with a maximum injection pulse of 140 DAC.

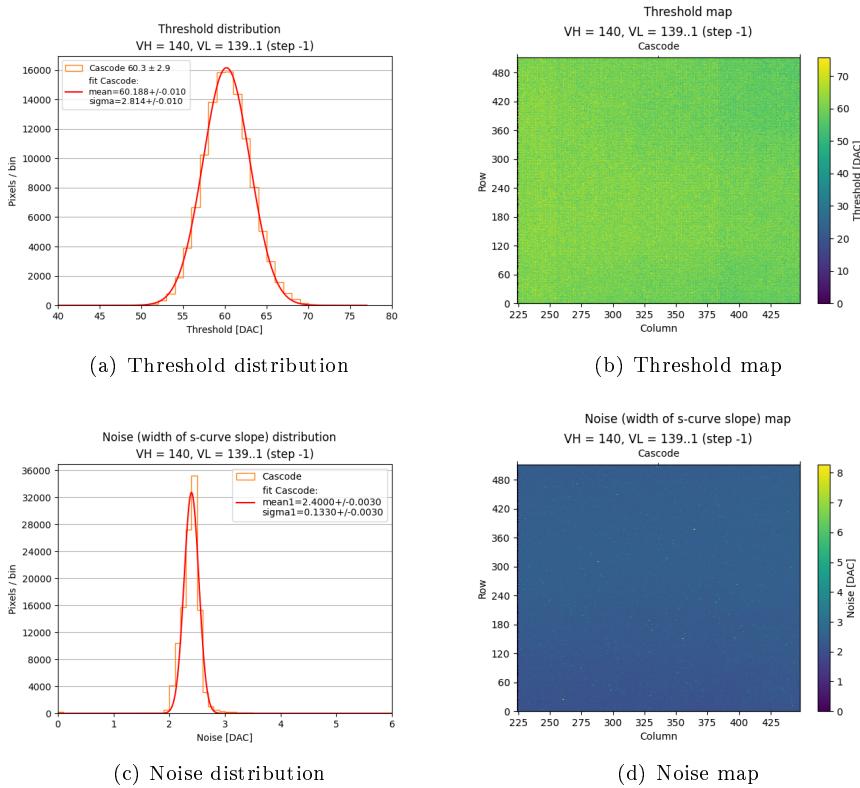


Figure 5.10: Cascode FE.

be identified as "noisy pixel" (because it results active regardless of the charge injection).

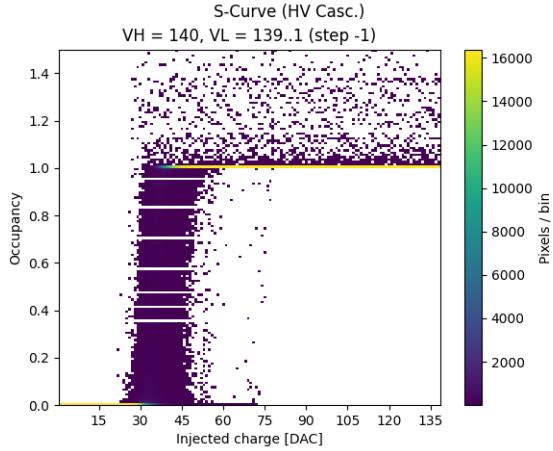


Figure 5.11: S-curves of all pixels in **HV Cascode** flavor with a maximum injection pulse of 140 DAC.

In figure 5.12 are shown the fit of the threshold and noise distributions.

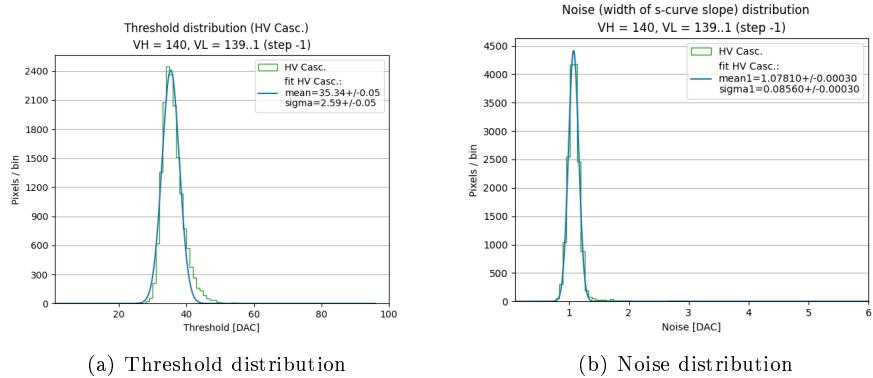


Figure 5.12: HV Cascode FE.

5.2.1.4 HV-Normal FE

The fourth and last flavor is the **HV-Normal FE** which has the same layout and so the same number of pixel of the previous FE. The main registers have been set with the values reported in table on page 57. In figure on the facing page, the s-curves of all pixel in this flavor. Also here we can see that there were some noisy pixels unmasked. Moreover the last 16 columns were not working (visible in the maps in figure on page 62) and as a matter of fact they had return a peak of threshold near the value 0, which is excluded from the threshold distributions plots.

So actually in this part of the matrix, the real number of pixel studied was the half of the total, such as 8192 pixels.

In figure 5.14 the fit of the threshold and noise distributions.

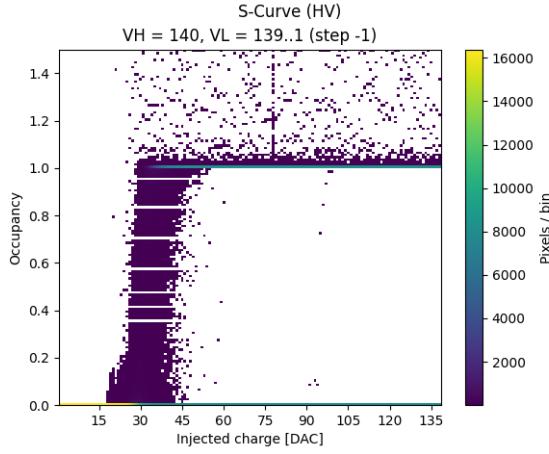


Figure 5.13: S-curves of all pixels in **HV Cascode FE** with an injection pulse of 140 DAC.

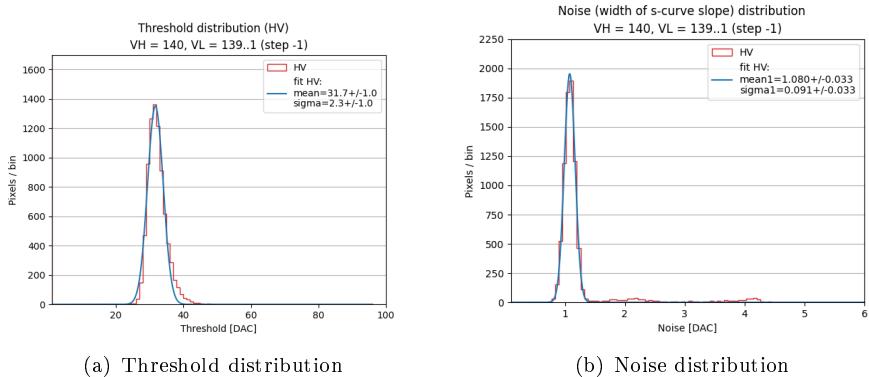


Figure 5.14: HV Normal FE.

At last, in figure on the next page the threshold and noise maps of the whole HV flavor.

As we will see in the following (section on page 78), the atypical s-curves in HVs flavors have been the first hint of the cross-talk problem, tied to a global lower threshold in these sectors with TB settings, compared with the first two.

5.2.1.5 Summary Table

In table on the next page a summary of results for threshold, noise and threshold dispersion of all FE.

5.2.2 Threshold dispersion and tuning

Despite its predecessor, Tj-Monopix 2 is equipped with a circuit which allows the *threshold tuning*. We have already mentioned that the analog part of the in-pixel front-end (figure 5.4 on page 54) includes the 3-bit threshold tuning DAC,

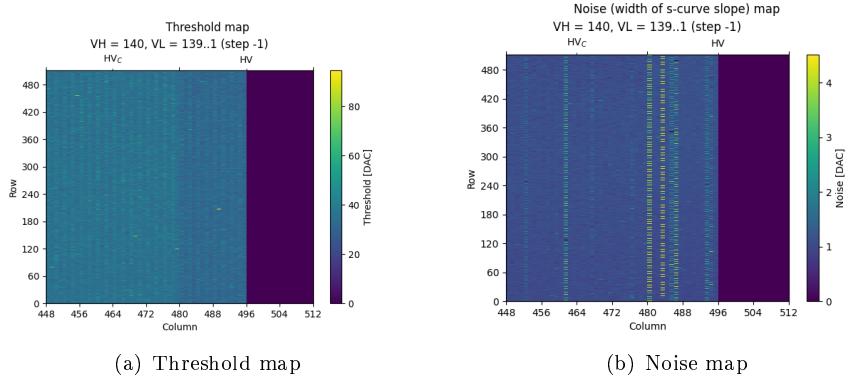


Figure 5.15: HV's FE.

Front-End	Threshold [e^-]	Threshold dispersion [e^-]	Noise [e^-]
Normal	53.62 ± 0.01	2.21 ± 0.01	2.503 ± 0.005
Cascode	60.19 ± 0.01	2.81 ± 0.01	3.037 ± 0.002
HV - Cascode	35.34 ± 0.05	2.59 ± 0.05	2.400 ± 0.003
HV	31.70 ± 0.10	2.30 ± 0.10	1.0781 ± 0.0003

Table 5.2: Summary table of threshold and noise values for all flavors of the W14R12 chip.

which not only improves the global threshold dispersion across the pixels, but also solves the issue with the unintentionally masked pixels (ghost), reducing the noise even more. In other words it can adjust every pixel threshold, even if only by few DAC, in order to have a global threshold on the matrix as uniform as possible, or in any case a dispersion as small as possible, especially after irradiation. This system has been design in order to counteract some negative effects that affect the threshold dispersion like systematics (for example related to biasing), process and temperature variations and radiation damage.

Specifically the TDAC circuit, shown in figure 5.16 on the next page, helps to make threshold trimming of each pixel. This component controls the discriminator active load current I_{DISC} which is partially responsible of the pixel threshold. It works as an analog multiplexer (consisting of simple PMOS transistor switches), which selects one of seven $I_{DISC,n}$ lines generated by the main 8-bit biasing DAC. So the possible value of the final I_{DISC} is given by the sum of two contributions:

$$I_{DISC} = I_{DISC,coarse} + (TCODE - 1) \cdot I_{DISC,fine}, \quad \text{where} \quad 1 \leq TDAC \leq 7 \quad (5.3)$$

$I_{DISC,coarse}$ is the current set by the primary value of threshold, resulting by the setting of the main registers which are responsible for it (listed in ...). $I_{DISC,fine}$ is the current selected by the fine tuning step (TDAC) and it depends on the 3-bit tuning code that is stored in the in-pixel tuning memory latch (the

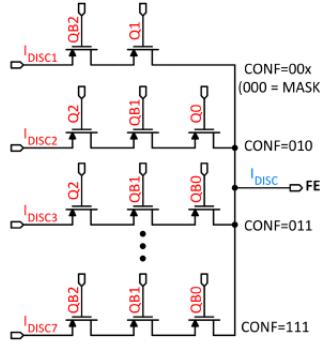


Figure 5.16: Schematic of 3-bit tuning DAC (TDAC)

in-pixel configuration memory). **TCODE** is the decimal representation of the TDAC code.

For example if the 3-bit DAC are set to "111", the decimal representation is 8 and the fine tuning provide a current $I_{DISC,7}$, which corresponds to the highest threshold. If the 3-bit are set to "010" the corresponding TCODE is 2 and the current $I_{DISC,1}$ is provided, which set the lowest threshold possible around the central value $I_{DISC,coarse}$. The particular combination "000" instead (TCODE = 0) masks the pixel by disabling the discriminator, without affecting the operation of the others.

5.2.2.1 First results from fine tuning

It has been trying to apply the fine tuning method to level out the threshold of some pixels as much as possible. We have considered about 12.000 pixels of the **Cascade FE** and in figure on the following page are shown the results before and after the threshold trimming for the s-curves and threshold distributions.

As we can see the dispersion has been reduced of the 42% after the tuning and as consequence also the estimation of the threshold is more precise. In figure on page 65 are displayed the maps of the threshold and of the TDAC values, such as the value of TCODE assigned at each pixel, in order to obtain a global threshold as uniform as possible.

In more details, there are three step for tuning:

- launch a first threshold scan through the internal charge injection, with a TDAC value equal for all pixels (usually default value is TDAC= 4).
- start a TDAC analysis which allows to choose a target threshold. During this phase, a tool tries to assign the optimized value of TDAC at each pixel in several steps, in order to get as close as possible to the set threshold. At the end it returns the final TDAC values for all selected pixels.
- At last another scan is launched setting the TDAC of each pixel passing the TDAC map values obtained from the previous.

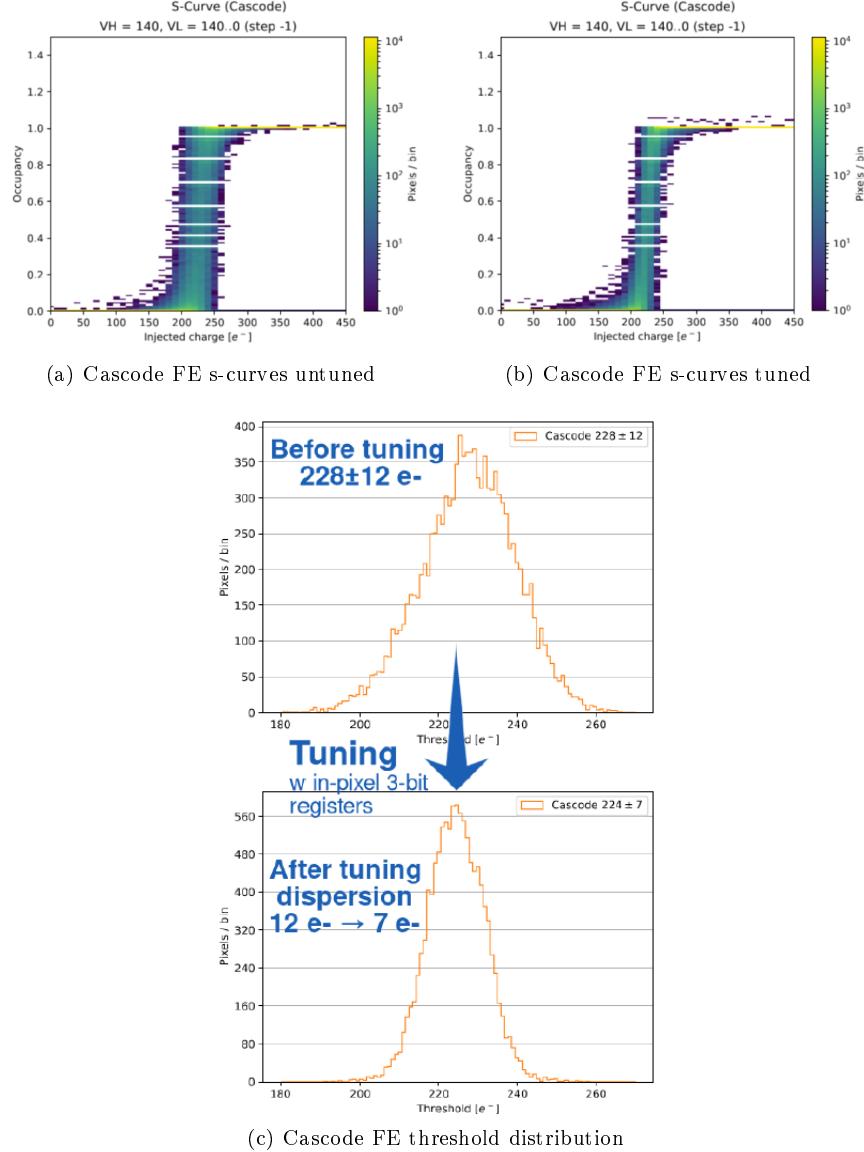
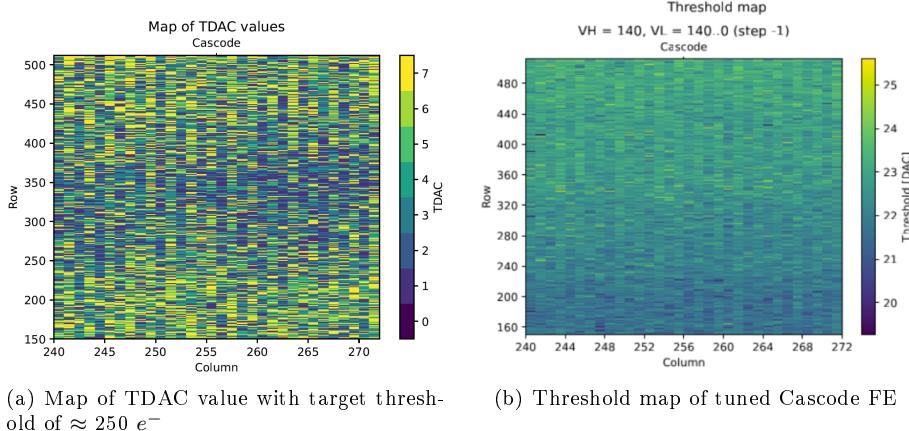


Figure 5.17: Cascode FE before tuning and after tuning comparison.

5.3 ToT calibration with internal injection

As it has been pointed out in the previous, the choice to use a simple diode instead of a PMOS transistor as reset input baseline element, increases the tolerance to TID radiation but at the same time it implicates a non-linear relationship between the injected charge and the ToT. For this reason, one of the goal of this analysis consisted to fit the trend of Q_{inj} vs. ToT, in order to obtain the absolute calibration of the whole matrix.



(a) Map of TDAC value with target threshold of $\approx 250 e^-$

(b) Threshold map of tuned Cascode FE

Figure 5.18: Maps of tuned Cascode FE.

5.3.1 Injection circuit issues

In carrying out the measurements mentioned above, we started to notice some issues with the injection circuit, which seemed to limit its working range. As a matter of fact the height of the injection pulse is expected to grow linearly increasing the value of charge to be injected. It actually happened up to a value of (about) ≈ 140 DAC, but for higher quantities the circuit seemed to increase not only the height of the signal, but also the threshold by a certain amount of ΔV (or equivalently of ΔQ , related by the conversion factor on page 55). Moreover, for injection heights grater than 200 DAC, only the threshold grows, without increasing the actual injected charge in any way.

However as we have seen in the previous section, the threshold depends on the settings of the chip registers and it can't be influenced by the injected charge, otherwise the whole response of the chip would be chaotic and it would not be reliable to take precise measurement of the impinging particles.

A method has been therefore devised to obtain reliable values of threshold and ToT up to a value of 170 DAC of effective charge injected. Moreover the characterization of the function used to describe the Q_{inj} - ToT relationship has allowed also to extrapolate ToT values in the forbidden region of charge (above $\approx 1717 e^-$), where the emission peaks of the radioactive sources available in the laboratory (usually) are.

5.3.2 Time Over Threshold (TOT) curves and fit [CHECK]

The function chosen for this purpose is:

$$y(x) = a \cdot x + b - \frac{c}{x - t} \quad (5.4)$$

with a , b , c and t free parameters and where the y represents the ToT corresponding to a precise value of collected charge, express by x .

Actually we know that the ToT distribution starts to grow near the threshold, so a random parameter among them could be computed in function of the

threshold value estimated from the previous measurements reported in section 5.2.1.

In more details, knowing that $y(x_{th})$ must be equal to 0 that is the ToT value at the threshold, it can be imposed for example:

$$0 = a \cdot x_{th} + b - \frac{c}{x_{th} - t} \quad \Rightarrow \quad c = x_{th}^2 \cdot a + x_{th} \cdot (b - a \cdot t) - t \cdot b \quad (5.5)$$

In this way the number of parameters to fit is reduced. In principle a similar equation could be equivalently solved for a , b and t .

Thus different fits have been made: one imposing a constraint on a free parameter (like shown above) and the other one leaving all parameters free. For all of them the value of χ^2 (MSE) have been computed and for all flavors it had its minimum when no parameters were fixed. So in the following the results from these last fits have been considered.

The same data collected in the previous measurements of thresholds have been used to fit the ToT curves of all pixels for each frontend. As a matter of fact we want to fully characterize the chip response with "GOE" settings (table on page 57), in order to use the results for analyzing TB data. In table 5.3 are reported the results of the fits for all parameters. In figure 5.19 on the next page the plots obtained for all Normal, Cascode and HVs FE.

	Normal	Cascode	HV Cascode	HV
$a \pm \Delta a$ [$\frac{\text{ToT}}{\text{DAC}}$]	0.12 ± 0.07	0.12 ± 0.01	0.257 ± 0.007	0.275 ± 0.008
$b \pm \Delta b$ [ToT]	4 ± 18	1.4 ± 3.1	3.2 ± 1.4	2.3 ± 1.6
$c \pm \Delta c$ [ToT-DAC]	200 ± 1100	140 ± 130	160 ± 70	140 ± 80
$t \pm \Delta t$ [DAC]	20 ± 90	40 ± 15	17 ± 6	13 ± 8

Table 5.3: Parameters obtained from the fit of ToT curve for each frontend.

5.4 Response to radioactive source and absolute calibration

The absolute calibration of the matrix consists in characterizing the response of each pixel to a known signal, like the emission peaks of radioactive sources. By the means of the Q_{inj} - ToT fit (section 5.3.2), it is possible to extrapolate the ToT signal value induced for whatever collected charge.

At this point then, three different X-rays radioactive sources were used to study the signal spectrum and the response time of the matrix, with emission lines from 6 to 60 KeV (corresponding approximately to 1600 and 16000 e- respectively). In fact the knowledge of the sources emission spectrum (in other words, the charge released in the matrix from particles emitted in decays) allows to compare the spectra obtained irradiating the chip, with the expected value of their peaks. It is worth mentioning that only the events in which all charge inducted is collected in a single pixel are part of the peaks reconstructed by the chip.

In addition these radioactive sources allowed to extend the ToT calibration at higher values with respect to the limit imposed by the saturation of the internal

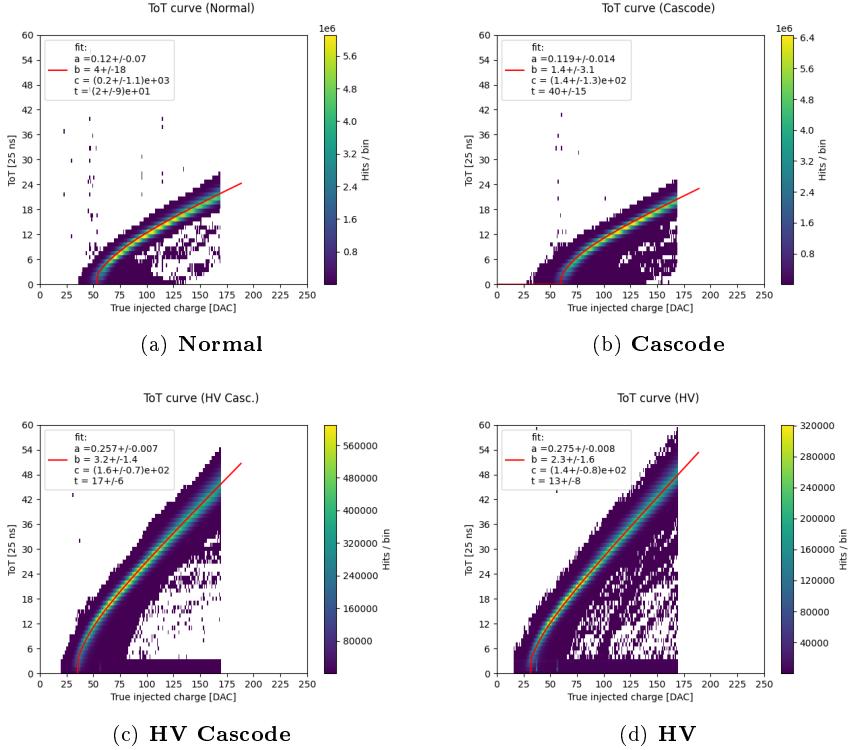


Figure 5.19: ToT curves fit for all frontend.

injection circuit (section 5.3.1). In table 5.4 on the following page are shown the emission energies of the employed sources, that it was possible to see with the chip under test.

Considering that the average energy necessary to produce an electron/hole pair in silicon is 3.65 eV, it is possible to convert the peak energies in a mean value of electrons released using the equation on the current page. So in the table are reported also the equivalent emission in electrons, which will be useful later.

$$N_{e^-} = \frac{E [eV]}{3.65[\frac{eV}{e/h\ pair}]} \quad (5.6)$$

Now we can go through the results obtained from three different sources: ^{55}Fe , ^{241}Am and ^{109}Cd .

5.4.1 ^{55}Fe

The ^{55}Fe source decays by **electron capture** to ^{55}Mn . One of the photons emitted in this transition has an energy of 5.9 KeV (K_α) and it produces in turn a photo-electron which deposits a ionization charge of about 1616 e^- in the sensor. All flavors were irradiate with a ^{55}Fe source available in the laboratory (with activity of 18 MBq). In figure 5.20 are shown the results. Each peak were fitted by a gaussian function, limited in the region of the peak itself. The

Source	Energy γ [KeV]	Equivalent charge [e^-]
^{55}Fe	5.9	1616
^{241}Am	13.9	3808
^{241}Am	17.7	4849
^{241}Am	20.7	5671
^{109}Cd	22	6027
^{241}Am	26.4	7233
^{241}Am	59.7	16356

Table 5.4: Emission lines of ^{55}Fe , ^{241}Am , ^{109}Cd sources visible by the sensor.

shoulder visible for smaller ToT is a consequence of charge sharing among pixels.

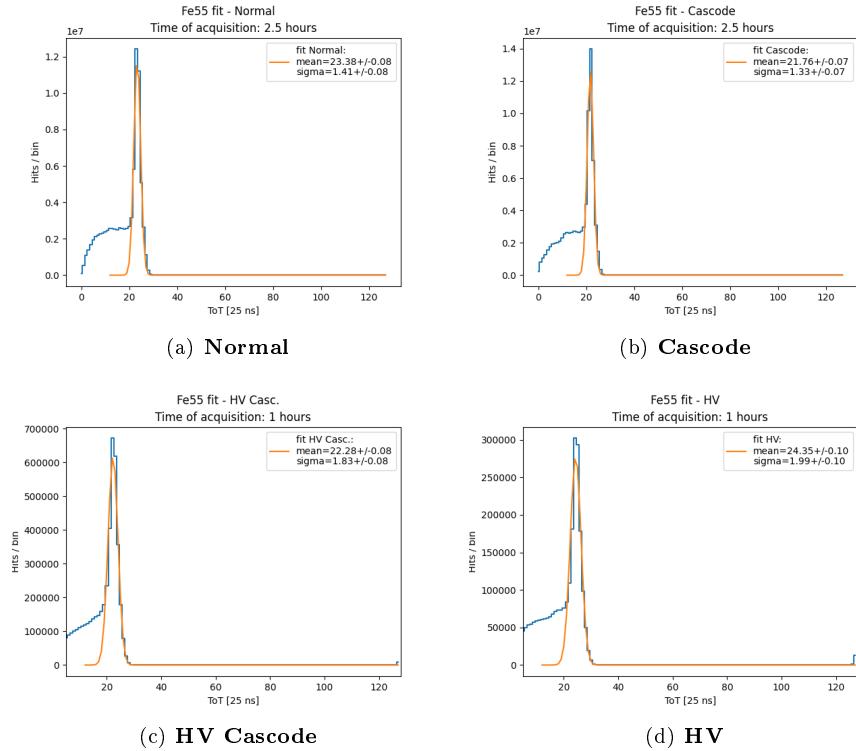


Figure 5.20: ^{55}Fe peaks for all frontends.

As it can be seen for the HV's FE a cut has been applied only to make clearly visible the emission line, indeed a lot of noisy pixels caused a sharp peak at 0 ToT. As a matter of fact in this flavors there were several columns of not-functioning pixels. In the box of each plot are also reported the results of the fit that will be crucial in the following.

5.4.2 ^{241}Am

The ^{241}Am source has a more complex spectrum (figure 5.21) and not all its peaks can be detected by the chip (because of the limited range of ToT available, depending on the number of bits dedicated to it). The spectrum shows other minor peaks besides the usual intense gamma peaks (59.5 and 26.3 keV) and several characteristic L X-rays from ^{237}Np (20.7, 17.7 and 13.9 keV).

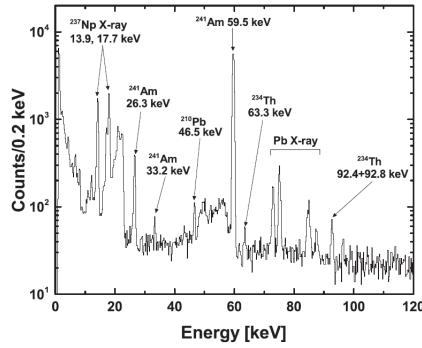


Figure 5. Gamma-ray spectrum of the ^{241}Am source.

Figure 5.21: ^{241}Am γ emission spectrum.

Results are reported in figure 5.22 on the following page.

In the case of the first two flavors, it could be possible to fit four peaks of the emission lines. In case of the HV's flavors instead, only three peaks for the HV-Cascode FE and two for the HV. As already discussed (section on page 54) the AC-coupling causes about 41.5% of signal loss, so they are much less evident and more difficult to fit as isolated peak.

5.4.3 ^{109}Cd

The third source employed was the ^{109}Cd . This isotope decays in ^{109}Ag by electronic capture, producing a photon of 22 KeV in the transition. In figure on page 71 the results obtained irradiating all FEs.

5.4.4 Injection capacitance calibration

Here it is necessary to point out that for iron source more statistics were collected so in this case a complete analysis of each pixel has been done. For the other sources instead, there weren't enough statistics on every pixel so the injection capacitance has been estimated only as a mean value for the whole front-end, just to compare with the results obtained from the iron analysis.

In case of ^{55}Fe source, we managed to fit the emission peak for each working pixel of the whole matrix. The value of the charge corresponding to the ToT peak of the emission line was extrapolated considering the parameters' values obtained by fitting the Q_{inj} - ToT relationship (section on page 65).

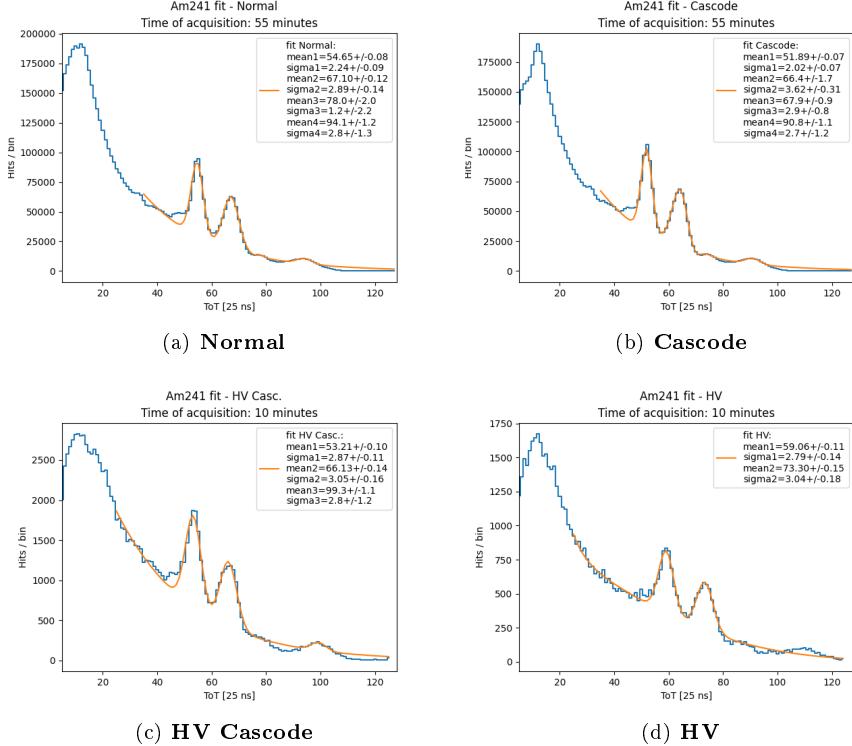


Figure 5.22: ^{241}Am peaks for all frontends.

Specifically the fit function on page 65 was inverted obtaining:

$$x(y) = \left(\frac{t}{2} - \frac{b}{2a} + \frac{y}{2a} \right) \pm \sqrt{\left(\frac{t}{2} + \frac{b}{2a} - \frac{y}{2a} \right)^2 + \frac{c}{a}} \quad (5.7)$$

where x represents the charge corresponding to the ToT labeled by y .

As shown in table on page 68, the charge released in the sensor (considering that collected from only one pixel) corresponds roughly to 1616 e^- . Therefore it was possible to estimate the conversion factor for each pixel as follows:

$$C_f \left[\frac{e^-}{DAC} \right] = \frac{1616\text{ e}^-}{ToT \frac{DAC}{ToTunit}} \quad (5.8)$$

By these steps, a value of the injection capacitance was estimated for each well-functioning pixel. In figure on page 72 is reported the distributions of the injection capacitance estimated, fitted by a gaussian function.

Regarding the other sources, it was impossible to fit the distributions for each pixel due to low statistics. For this reason only a mean value for all flavor could be extrapolate. In table on the current page the results obtained with the same method used with the iron source, but considering all pixels of each flavor.

Bringing equation on page 55 back to mind, the conversion factors of the flavors are not those expected. In particular for the HV's flavors, they are almost the double and it could mean that the injection capacitance is greater

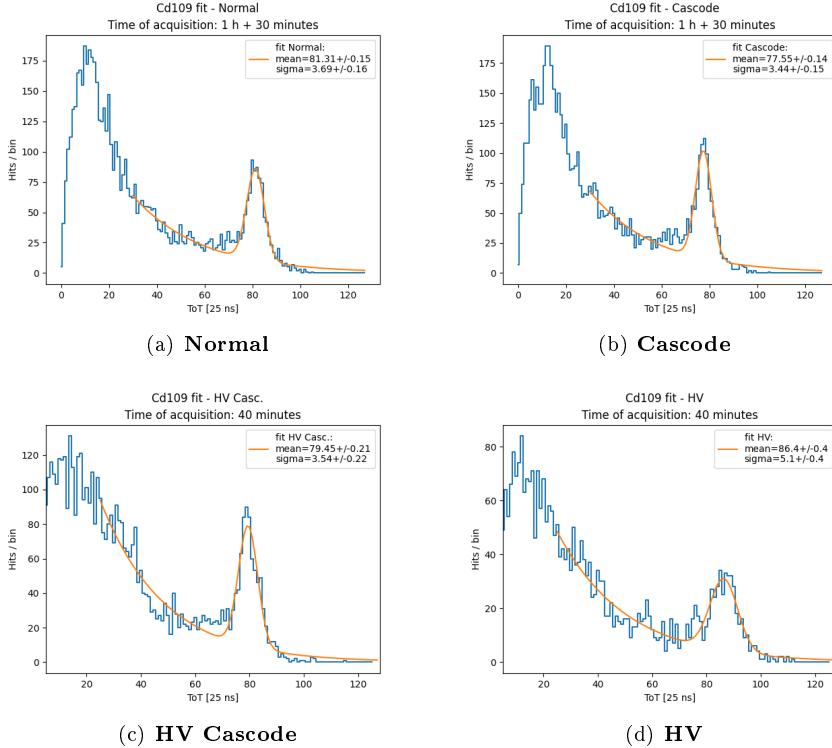


Figure 5.23: ^{109}Cd pekas for all frontends.

Source peak	C_{Normal}	$C_{Cascode}$	$C_{HVCascode}$	C_{HV}
^{55}Fe (5.9 KeV)	9.37	9.00	19.33	18.56
^{241}Am (13.9 KeV)	8.94	8.91	19.23	18.22
^{241}Am (17.7 KeV)	9.16	8.84	19.59	18.63
^{241}Am (20.7 KeV)	9.15	10.11	-	-
^{109}Cd (22 KeV)	9.32	9.39	20.16	19.6
^{241}Am (26.4 KeV)	9.60	9.61	19.25	-
Mean value	9.26	9.31	19.51	18.75

Table 5.5: Estimation of injection capacitance of all flavors for different source emission peaks.

than expected. But as mentioned in section on page 55, the prototype under test was designed in order to have the same injection capacitance for all flavor, equal to 230 aF (despite this doesn't seem to be the case). The cause of the different results for the conversion factors might be due also to the expected loss of the collected charge for HV FEs (section on page 54).

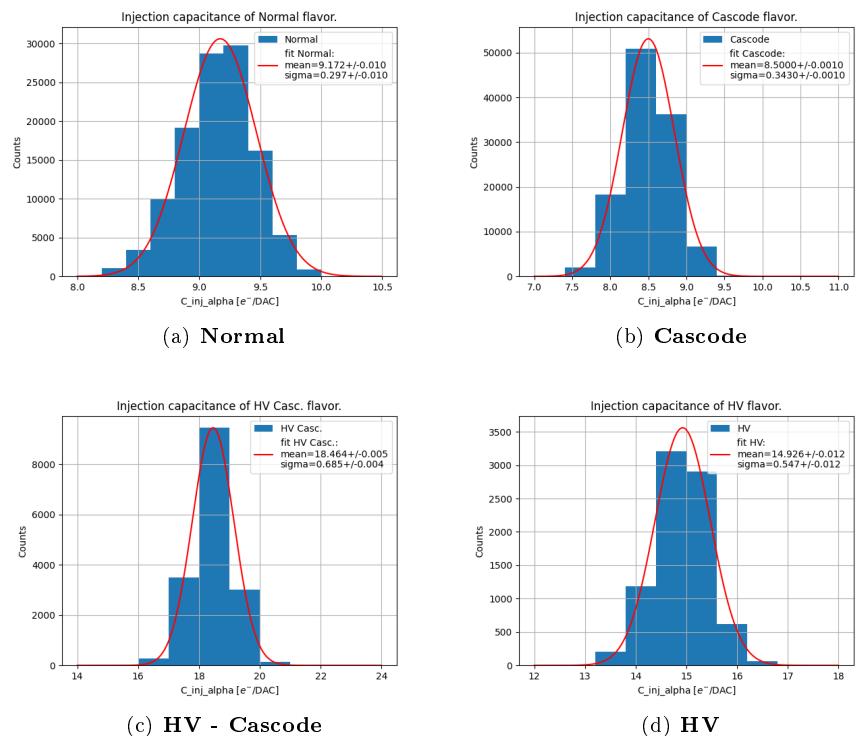


Figure 5.24: Injection capacitance distributions of all FE.

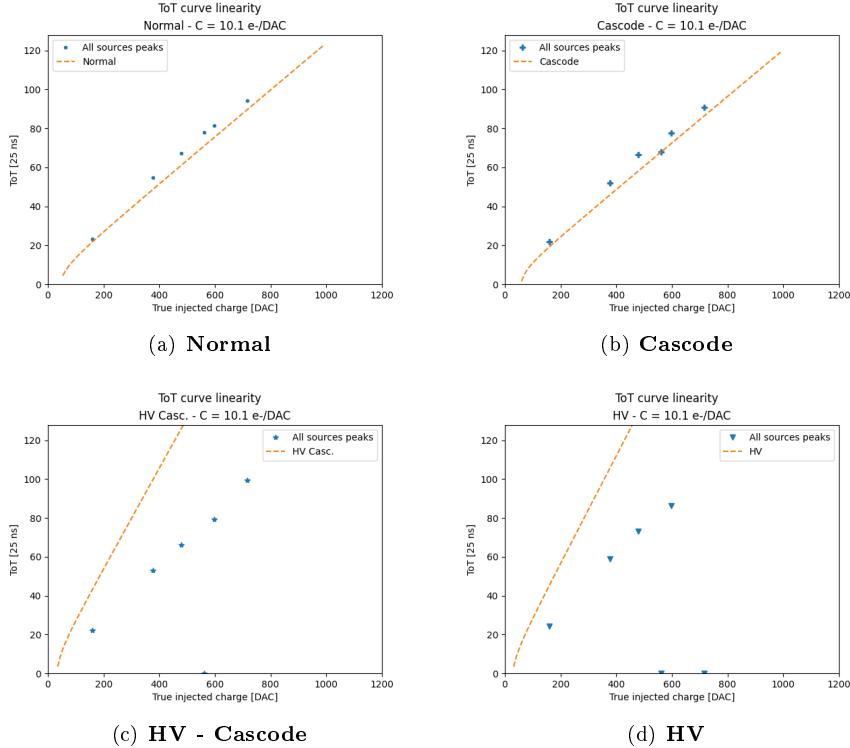


Figure 5.25: ToT linearity of all flavors assuming the nominal(expected) conversion factor equal to $10.1 \frac{e^-}{DAC}$.

5.4.5 Check on linearity of tot fit

Eventually, the values of all emission peaks (obtained by the fit) from the several sources have been plotted for each frontend, in order to verify the agreement between their trend and the ToT-Q relationship studied by the internal injection.

Initially the average quantity of electrons, expected to be released for each peak, has been calculated with the nominal conversion factor equal to $10.1 \frac{e^-}{DAC}$. As it could be seen from results in figure on the current page there isn't good agreement between data and ToT relationship studied in section on page 65.

After the calibration instead, assuming the average value of injection capacitance (*Mean value*) reported in table on page 71, the charge corresponding to the emission peaks have been recalculated and results are shown in figure on the following page.

As we can see, a better agreement is therefore obtained and this shows that through calibration we are now able to interpret data with greater precision, which was the main purpose of this analysis.

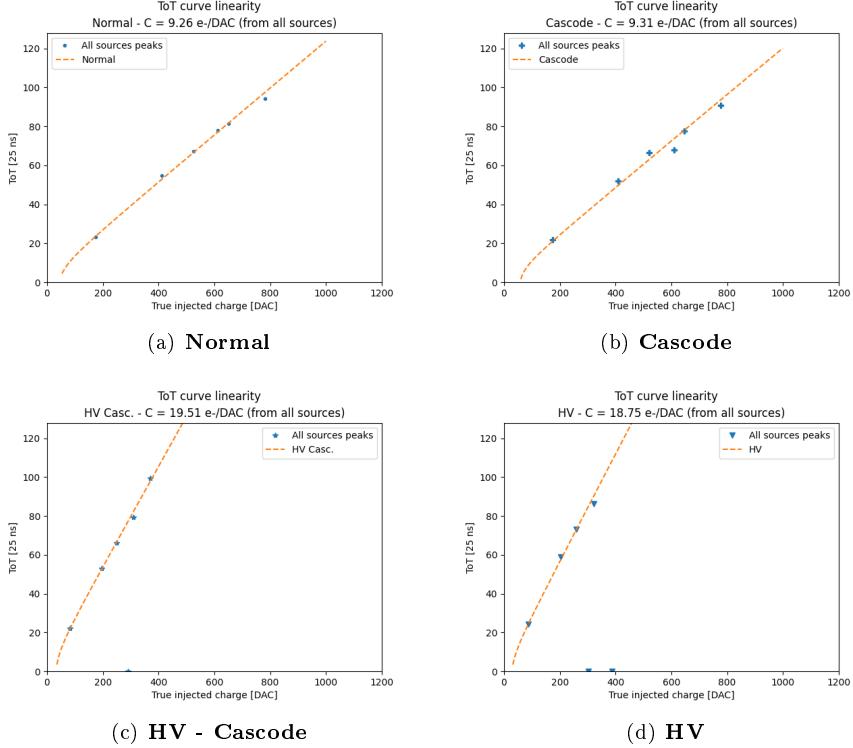


Figure 5.26: ToT linearity of all flavors assuming the conversion factor obtained from calibration for each FE.

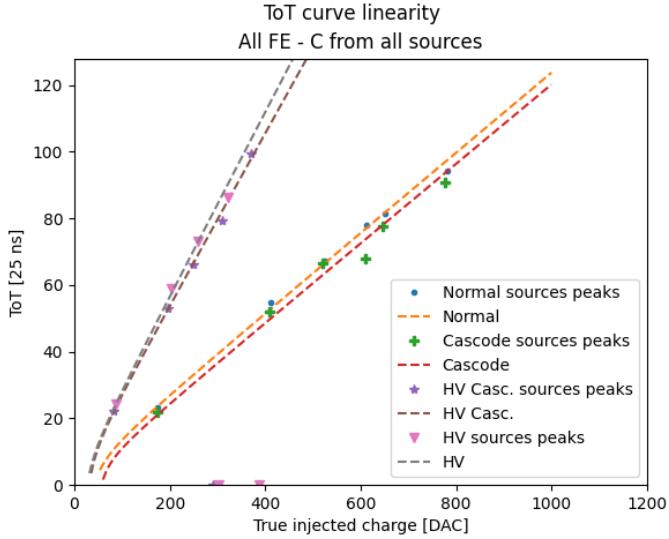


Figure 5.27: Summary of trends.

5.5 Operation with low threshold

One of the most important target in sensor design is to keep high efficiency even after irradiation damages. All experimental environments indeed, are exposed to high doses of radiations, so it's crucial to ensure the functionality of the detectors, even after being irradiated.

For this reason, many tests were done in order to understand the chip behaviour at low threshold where a good value of efficiency could be preserved. Moreover working at low threshold allows to detect low charge events due to charge sharing or charge trapping (effect which increases after irradiation), especially in case of thin epitaxial material.

5.5.1 Register optimization

As we have seen in section (REFERENCE), there are a lot of registers which control the discriminator threshold and also the readout sequence. So preliminary it was necessary to explore their possible settings in order to operate the chip at low threshold.

Now we will go through the main registers used for this purpose, in order to explain their functionality. There are several dozens of registers but we focused on some of the most important and crucial to set the threshold and its dispersion.

- I_{CASN} : this current is responsible of the output baseline signal. In particular it sets the baseline of the FE output that goes to the input discriminator. In a few words, higher this value, higher the baseline, lower the threshold and also a little bit the gain and vice versa by decreasing it.
- I_{THR} : it controls the pre-amplifier feedback strength and speed, so it is responsible for the output reset rate. Increasing this current increases the gain and the time that the analog output takes to get back to the baseline and as consequence, it increases a lot the maximum value of the ToT. In fact it is recommend to set I_{THR} to low value (e.g. 8 nA[ref]) in order to avoid high ToT slope.
- I_{DB} : this current corresponds to $I_{DISC,coarse}$ explained in section 5.2.2 on page 62. It represents the primary current that sets the discriminator threshold, to which another current is added by the tuning.
- I_{TUNE} : it corresponds to $I_{DISC,fine}$ instead (always section on page 62). Remembering the equation from tuning¹ (on page 62), this is the current to multiply by the TDAC value, which is added to I_{DB} , during the tuning process.
- I_{BIAS} : this current acts on the pre-amplifier input transistor and influences the threshold dispersion and the gain. In particular increasing this value, the dispersion decreases and the gain becomes greater. Nevertheless it can't be increased a lot because it affects the power consumption, too.

¹

$$I_{DISC} = I_{DISC,coarse} + (TCODE - 1) \cdot I_{DISC,fine}, \quad \text{where} \quad 1 \leq TDAC \leq 7 \quad (5.9)$$

- V_{RESET} : this register influences the threshold dispersion. Lowering its value, the dispersion decreases and vice versa.

5.5.2 Comparison between data and simulation

In the interest of understanding how the registers' setting of the chip influences the threshold, several measurements have been taken with different configuration of their values. The results are compared with simulations done by Hung Pham. (...). [??]

I_{CASN}

In figure 5.28, we can see the simulated behaviour of the threshold and the gain, increasing the value of I_{CASN} .



Figure 5.28: Trends of Gain and Threshold increasing I_{CASN} .

To verify the trend of threshold as this current varies, three different acquisitions have been taken by fixing $I_{THR} = 20, 40, 64$ and increasing I_{CASN} from 0 to 30 DAC, with a step of 5 DAC. We have done this enabling 200 pixels in the Cascode FE (rows: 472 - 512, cols: 225 - 230).[??]

Each threshold distribution has been fitted with a gaussian function in order to estimate the average threshold value and its dispersion.

In figure 5.29 on the next page are reported all trends obtained.

I_{THR}

Reusing the same data of the previous measurements, the trend of the threshold have been studied by changing the value of I_{THR} and fixing that of I_{CASN} . In this case only I_{CASN} from 0 to 15 DAC is considered, because for higher values we don't have enough measures of the threshold (specifically only two for $I_{THR}=40, 64$). The results are shown in figure 5.30 on the facing page.

As expected, increasing I_{THR} results to lower gain and faster return to baseline, so higher threshold. We can compare them with the simulation shown in figure 5.31 on page 78.

Time over Threshold (ToT)

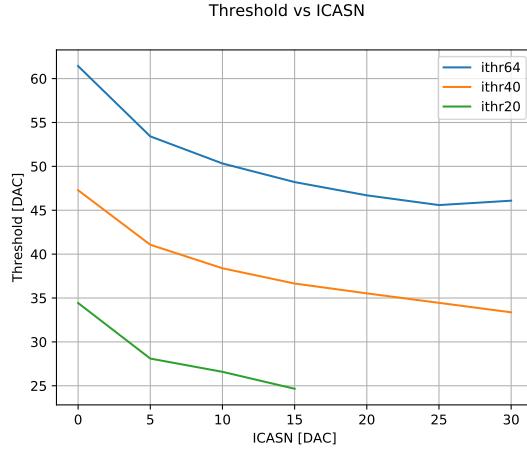


Figure 5.29: Threshold vs. I_{CASN} for $I_{THR}=20, 40, 64$.

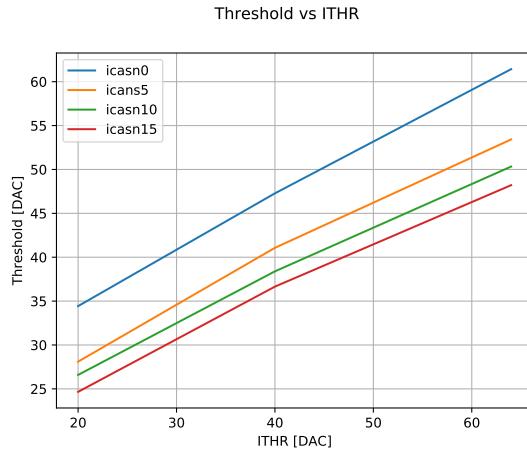


Figure 5.30: Threshold vs. I_{THR} for $I_{CASN}=0, 5, 10, 15$.

The last analysis done to make a comparison with the simulations, is about the trend of the ToT changing the value of I_{CASN} for a fixed value of I_{THR} and vice versa. In particular we consider the data obtained with I_{CASN} fixed to 0 DAC and I_{THR} to 64 DAC, which are the values studied and used for these registers during the Test Beam in Desy.

Results are reported in figure on page 79 and on page 80. Also here we can see a good agreement between data and simulations.

5.5.2.1 some nice picture of the optimized thr and tuning

!!!!

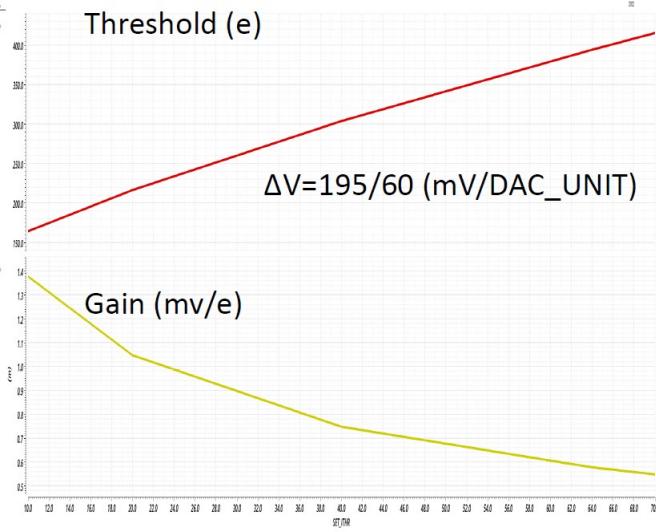


Figure 5.31: Trends of Gain and Threshold increasing I_{CASN} .

5.6 Cross talk issue and mitigation

As it was already pointed out, during the measurements of the average threshold of all FEs, there were something atypical in the s-curves of the HV flavors (subsection 5.2.1.4 on page 60), indeed some pixels seem to have occupancy greater than 1. This behavior threatens the good functionality of the overall matrix response, because these *noisy pixels* flood the readout, giving unreliable results.

Also the systematic study of different register configurations has revealed the presence of the hot pixels which prevent to use certain settings and as consequence to reach lower global thresholds.

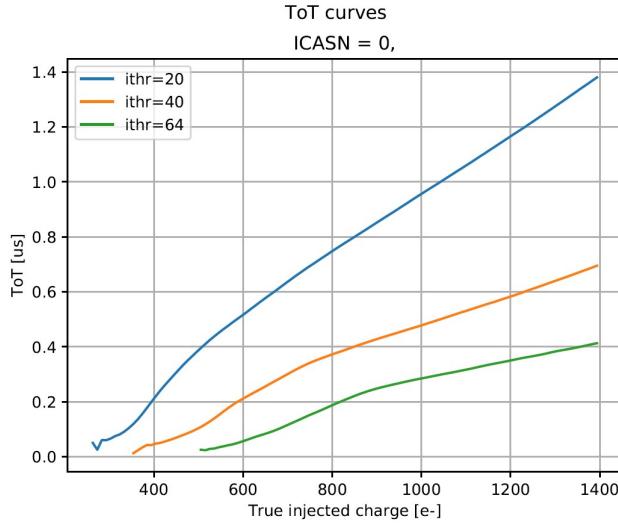
For this reason an investigation has been conducted in order to understand the reasons why they start to fire and how to cure them as far as possible. In the meantime an important issue with cross-talk of the readout signals was discovered, and so in this section we examine this effect and some attempts to mitigate it using different settings/bias.

5.6.1 Hot pixel issue

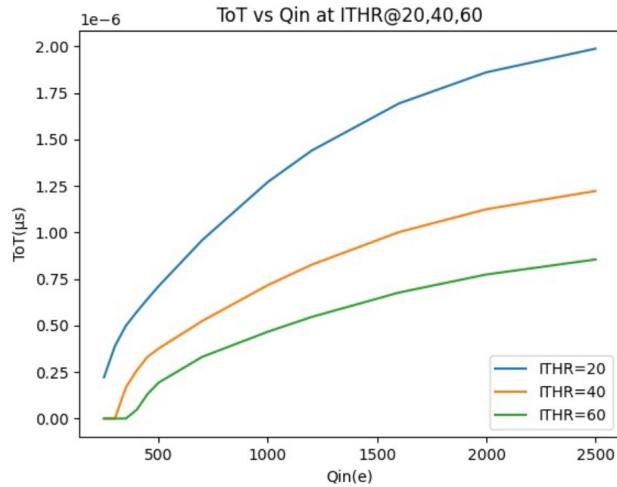
First of all we noticed that in the s-curves oh the HV flavors (e.g. that of HV-Cascode reported in figure on page 81), the atypical behavior could be triggered by a digital signal sent to the matrix during the readout activity at low threshold. This consideration is based on two main reasons:

- when the matrix has high threshold, like for Normal and Cascode FE, all pixels seem to behave as expected, without *hot pixels*.

Lowering the threshold and running some source acquisitions without any source, no strange behaviour was observed. Acquiring data with a radioactive source instead, even Normal and Cascode FE seem to reveal the



(a) ToT vs I_{THR} ($I_{CASN}=0$ DAC) - Data (Cascode)

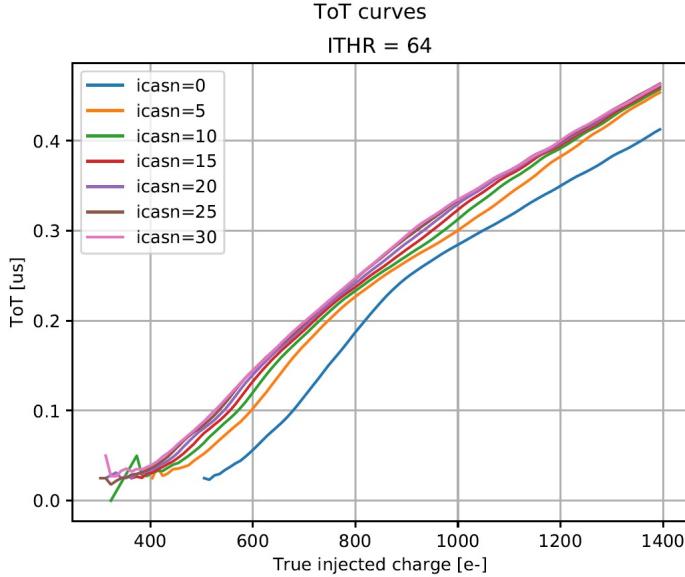


(b) ToT vs I_{THR} ($I_{CASN}=0$ DAC) - Simulation

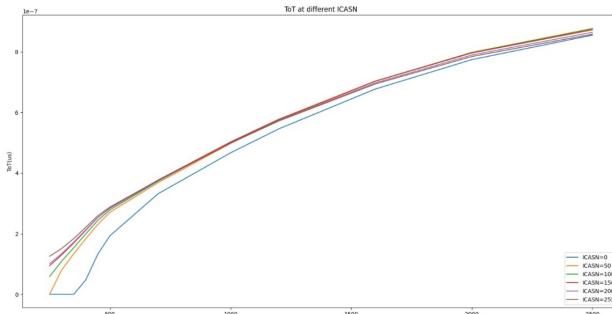
Figure 5.32: ToT vs I_{THR}

same problem. This led to thinking that during the readout of good pixels an induced signal is created which couples with some other pixels, in particular with those at lower threshold with respect to the average value. If the height of this signal exceeds the threshold of the single pixel, it causes some spurious hits, making the pixel "hot".

- Moreover, always considering the HV Cascode s-curves, it could be noticed that in the region before the threshold ($Q_{inj} < \text{threshold}$, pointed by the blue arrow) there isn't an anomalous activity which means that the induced signal is not due to the BCID that is always sent to the matrix during the injection or an acquisition with the source, regardless of being



(a) ToT vs I_{CASN} ($I_{THR}=64$ DAC) - Data (**Cascode**)



(b) ToT vs I_{CASN} ($I_{THR}=64$ DAC) - Simulation

Figure 5.33: ToT vs I_{CASN}

above or below the threshold. The atypical behaviour indeed, is in the region above the threshold ($Q_{inj} > \text{threshold}$, pointed by the red arrow) where the occupancy of some pixels becomes greater than 1. This means that these *hot pixels* detect more hits of those injected.

From these first observations, we have reached the conclusion that the cross talk could be tied to the readout activity. So we have started investigating the timestamp of those hits not synchronize with the timestamp of the injection.

5.6.2 Hot pixel strategy (study)

At first, it has been lowered the threshold in order to "create" hot pixel also in the first two flavors of the matrix. In fact with TB settings the threshold was too high and the hypothetical induced signal did not cause spurious hits. For

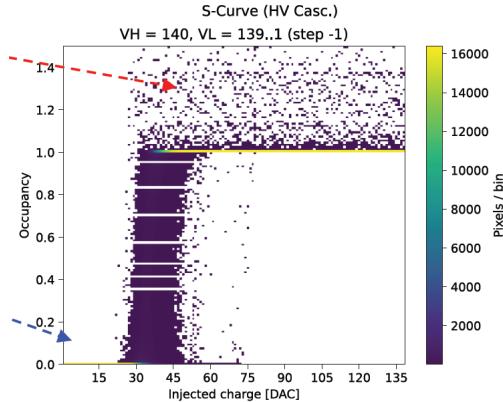


Figure 5.34: HV-Cascode s-curves.

this purpose different settings were tried, changing some fundamental registers responsible for the threshold like those listed and explained on page 75.

Then several tests have been run under controlled conditions:

- one healthy(good) pixel was injected;
- one *hot pixel* (two or three in different tests) was enabled but not injected;
- the all matrix except these pixels was disabled.

In this way (remembering the readout sequence [REFERENCE]) the readout cycle has a known duration and two different timing info has been used to study the induced signal with greater precision:

- ΔTS (TimeStamp) between two consecutive hits: the TimeStamp is assigned from the FPGA when the TOKEN rises on the TE of the first hit to read, but only if the previous readout frame is completed. So, if the hit coming from a *hot pixel* is after the hit from the injected one, the minimum ΔTS has to be equal to the readout time of 1 pixel and so to the duration of the FREEZE_STOP signal .
This info has allowed to verify if the hot pixel fires after the good injected one or not.
- LE(hit) - TE(previous hit): this quantity measures the elapsed time between a hit and the previous one. This is a finer info than the ΔTS because it allows to correlate the hit with the induced digital signal, originate from the readout cycle.

Moreover, since a 7-bit BCID is sent to the matrix during its activity, it was important to keep short (<128 clock cycle) the duration of the full readout sequence and to not enable too many pixels in order to not extend too much the readout frame. Otherwise the information on the leading edge of the pixel could not be correlated with the token of the previous hit. In other words, if the readout frame exceed 128 clock cycles, since the token could be raised if the matrix is **not** freeze, even if an hit is arrived before it could be read only in

the next frame when it could rise again the token, but in this case it will have different TimeStamp. So in this case the TS is useless for our purpose.

5.6.3 Cross-talk (Results)

Referring to the readout sequence, in order to understand which signal could induce cross talk, each register value has been moved one by one. In table on the current page just an example from the several settings tried is reported.

Register	Value
FREEZE_START_CONF	10
READ_START_CONF	13
READ_STOP_CONF	15
LOAD_CONF	30
FREEZE_STOP_CONF	31
STOP_CONF	31

Table 5.6: Register values of the Readout cycle.

Doing so indeed, the LE-TE info has to shift by the same value with which the signal that cause the cross talk has been changed. This step of the procedure is tied with the necessity to keep the readout sequence within the maximum 128 BCID range.

For example, if FREEZE_START_CONF is responsible for the cross talk signal, we expect that shifting its value by a certain amount, the hot pixels start to fire after that this signal arises due to the hit on the injected pixel. So the value of LE-TE has to be FREEZE_START_CONF + some potential delay. Same argument for the other registers.

By this procedure, repeated for each readout register, we have come to the conclusion that the cross-talk could be related to the raising and falling edge of the FREEZE signal.

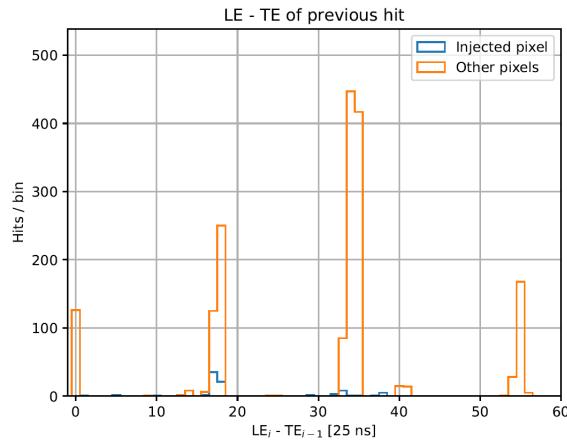
In figure on the facing page an example of some results obtained. It is the histogram of the time last between the leading edge of an hit and the trailing edge of the previous hit, when one pixel is injected and two are read. It's possible to see several peaks (referring to the readout setting reported in table on the current page):

- one at 0, that represent the situation in which both hits come from hot pixel firing simultaneously after the injection. This means that they are activated by the same signal and so it is the most important confirmation that is cross-talk and not random firing pixel signal;
- one at ≈ 18 equal to FREEZE STAR raising + 8 \rightarrow first induced signal;
- one at ≈ 35 equal to FREEZE STOP falling + 4 \rightarrow second induced signal;
- one at ≈ 55 equal to FREEZE STOP falling + 4 when two different pixels are read. In more details, after the first 30 time unit until the first

LOAD CONF, a distinct pixel reading starts and it takes another 20 time unit (LOAD - FREEZE START) + 1 unit time to conclude the frame with the FREEZE STOP and so $51 + 4$ unit time wrote above. Therefore when two pixels are read, the FREEZE STOP falls after 51 clock cycles, and it is compatible with the last peak in the plot.

Injected pixel: (217, 140) Other pixels: [(218, 155) (222, 188)] Assuming timestamp clock = 40.00 MHz Green = injected pixels							
Row	Col	LE	TE	ALE	ATE	$\Delta TS[25ns]$	$TS[25ns]$
140	217	12	29	124	123	5627.0000	635972.0000
140	217	8	25	124	124	5628.0000	641600.0000
155	218	60	60	52	35	35.0000	641635.0000
188	222	59	60	127	0	0.0000	641635.0000
155	218	115	115	56	55	55.0000	641690.0000
188	222	114	115	127	0	0.0000	641690.0000
155	218	42	43	56	56	55.0000	641745.0000
188	222	42	42	0	127	0.0000	641745.0000
140	217	4	21	90	107	5482.0000	647227.0000

(a) An example of the time quantity used in the analysis.



(b) An example of the LE(hit)-TE(previous hit) histogram.

Figure 5.35: Some results of the cross-talk studies.

As already stated, we run several tests varying the number of pixels to read, the value of the readout registers, different combination of hot and good pixels and also different spatial location of them in the matrix to exclude the possibility that the problem was related to particular columns. All results are in agreement with the interpretation explained above.

?????...

Furthermore it has been tried to estimate the height of the induced signal from the threshold of the hot pixel. For this reason we have tried different setting of the currents cited above to make a pixel *hot* in order to understand when the induced signal went above the threshold. We have found that the signal could (may) correspond to $100/150 e^-$.



Figure 5.36: Cross-talk of the FREEZE signal on oscilloscope's analog output, for different value of FREEZE_START_CONF register.

??????

In figure on this page an analog acquisition of the readout signals taken by an oscilloscope.

In these tests one pixel was injected from 0 to 140 DAC (in the acquisition it can be seen in the increasing signal height). There are two different group of spikes: the first which is smaller and represent the cross-talk from the raising of the FREEZE signal and the second, larger and corresponding to the cross talk from the falling edge of the same signal. Moreover it's possible to see that in the two different pictures, the cross talk signals move according to the different settings of the FREEZE START/STOP edge.

5.6.4 Mitigation

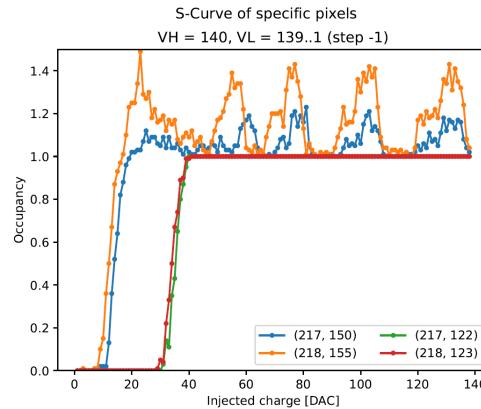
As seen in the previous, the problem of the hot pixel is tied to the induction signal produced during the readout, which cause cross-talk. It becomes even more serious when there is a grater dispersion threshold.

Potentially every pixel could become *hot* if its threshold is lower than the height of the cross-talk signal, since the FREEZE is sent across the entire matrix.

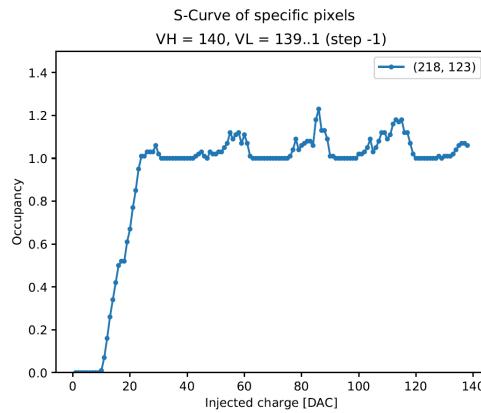
As an example, in figure on the next page it si possible to compare the behaviour of pixel (218, 123) for different register settings adopted in order to reduce the threshold.

Threshold trimming

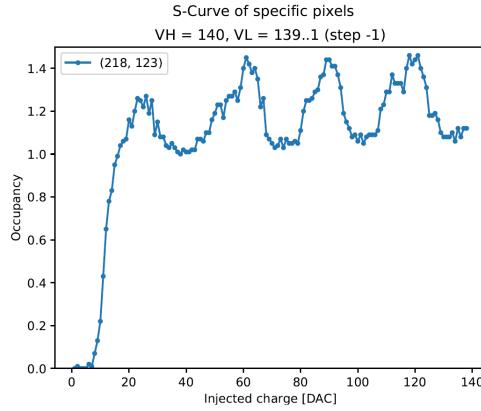
Therefore a possible treatment could be related to the threshold tuning, explained in section (5.2.2 on page 62) , which could allow to make the pixel



(a) $I_{DB}=100$, $I_{TUNE}=53$ - Good behavior (the red one)

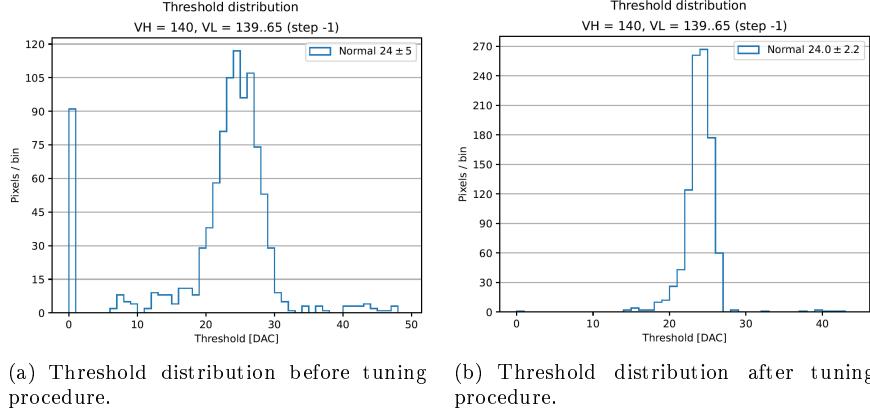


(b) $I_{DB}=60$, $I_{TUNE}=150$ - Pixel starts to misbehave



(c) $I_{DB}=55$, $I_{TUNE}=150$ - Pixel becomes *hot*

Figure 5.37: S-curve of the pixel (218, 123) for different register settings.



(a) Threshold distribution before tuning procedure.
(b) Threshold distribution after tuning procedure.

Figure 5.38: Threshold tuning to reduce hot pixels.

threshold more uniform (less threshold dispersion) and simultaneously to target a value greater than the induced signals.

In figure on the current page an example of the results obtained.

It's evident the reduction of the tail in the threshold distribution, in fact the dispersion is reduced by 56%. Also the hot pixels decrease from 18% to 1.2% of the total number of pixels studied. We can also notice that the peak at 0 threshold disappears.

Bias Voltage

Moreover it has been tried to increase the voltage bias of the all matrix, too. We remember that all previous test has been run with P_{WELL}/P_{SUB} set to -3 V. This value was increased to -6 V and indeed there were some improvements. In fact increasing the bias, we expected a decrease of the diode capacitance thus higher gain and lower threshold dispersion. In addition the coupling with the cross talk signal is reduced too and so the induced signal height.

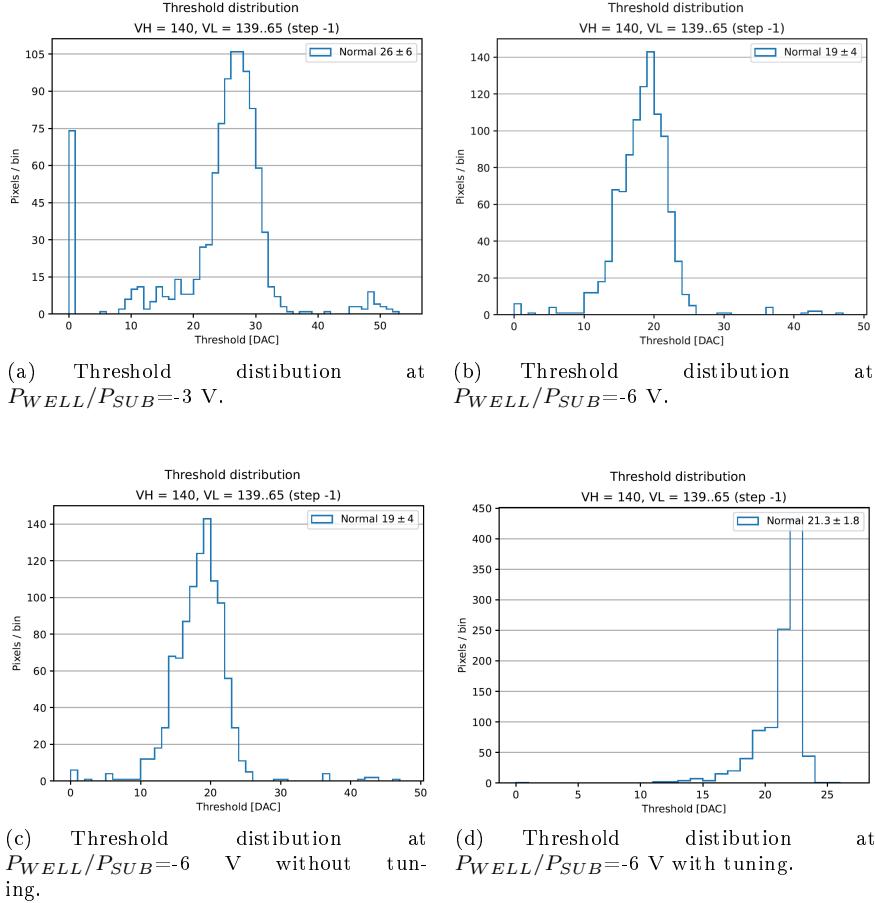
In figure on the facing page a comparison between the threshold distribution respectively at -3 V and -6 V, with same registers setting and without tuning.

At higher bias voltage not only is the threshold lower (higher gain), but also its dispersion, as expected. And despite that there are fewer hot pixels: 1.3% at -6V against 17% at -3 V. Also here there is clear a reduction of the threshold distribution tail.

5.6.4.1 Final results?

Eventually the final results obtained with both threshold tuning and a bias voltage on $P_{WELL}/P_{SUB}=-6$ V.

As we can see in figure on the next page, the threshold dispersion decreases together with the number of the hot pixels. In fact without the tuning, there are 1.3% of them, instead with tuning procedure there are none at all.



5.7 Test Beam results

This full characterization of the chip allowed to interpret data collected during the Test Beam done in Desy in June 2022. Several measurements have been taken, among which: pixel response, noise level, cluster signal, full depletion depth, voltage scan and angular scan. In particular the aim was to study electrical characteristics and hit detection efficiency of unirradiated modules.

5.7.1 Experimental apparatus and DUTs

The measurements have been performed in a 4 (or 5) GeV electron beam at DESY II testbeam facility at DESY, Hamburg. The experimental apparatus consisted of a beam telescope, a Trigger Logic Unit to provide trigger and control signals employed during test beams, a scintillator trigger and a rotation-translation stage on which install the device under test. (figure on the following page).

Three different modules have been tested with different sensor geometry, among which the chip W14R12 that we have studied in depth through laboratory measurements. All results described in previous sections have been crucial to interpret data obtained during these tests.

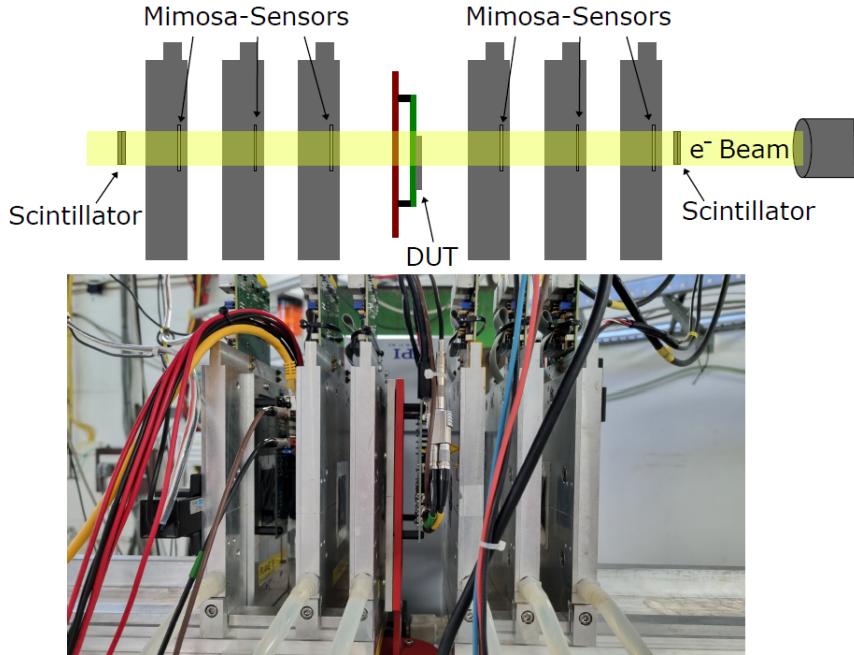


Figure 5.39: Test Beam experimental apparatus.

In the following we will briefly mention the results obtained.

5.7.2 Hit detection efficiency measurements

Initially a depletion depth study have been done and with a voltage bias of -3 V (Normal and Cascode FE), for the chip W14R12 resulted $33\mu\text{m} \pm 0.04$ (stat.) ± 2.53 (sys.).

Voltage scan measurements have been performed for both the two type of pixel couplings, such as DC and AC coupled. For the latter, only some first results have been obtained.

Normal FE

In figure on the next page are reported results obtained for **Normal FE** of all modules. As expected, efficiency and clusters grow as the bias voltage is increased.

Cascode FE

Same trends for **Cascode FE** are shown in figure on the facing page.

In particular in figure on the next page are reported the hit map efficiency for DC-coupled pixels for the W14R12 chip, which includes Normal and Cascode FE.

HV Casc. FE

HV FE

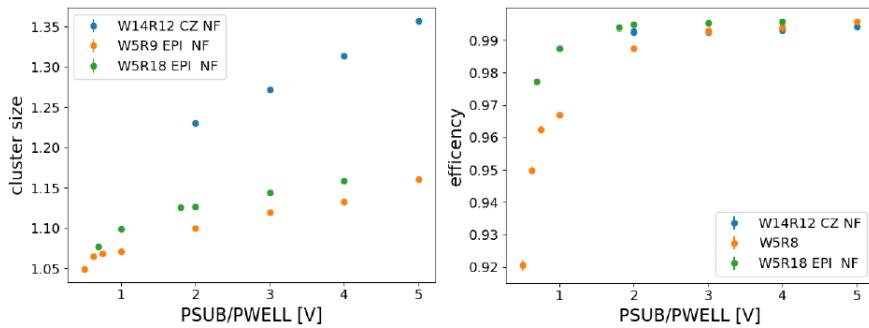


Figure 5.40: Cluster size and efficiency results for Normal FE.

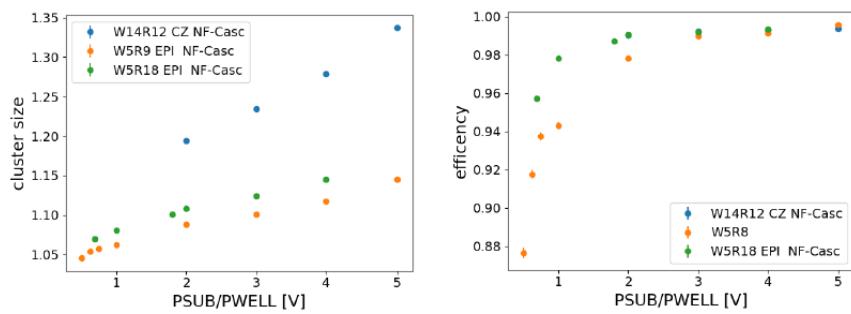


Figure 5.41: Cluster size and efficiency results for Cascode FE.

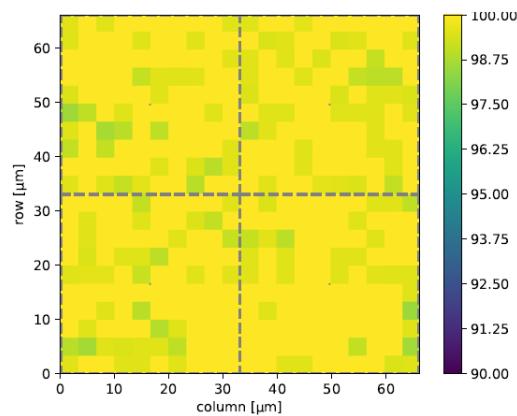


Figure 5.42: Efficiency for DC-coupled pixel of W14R12 chip: (99.79 ± 0.10) %

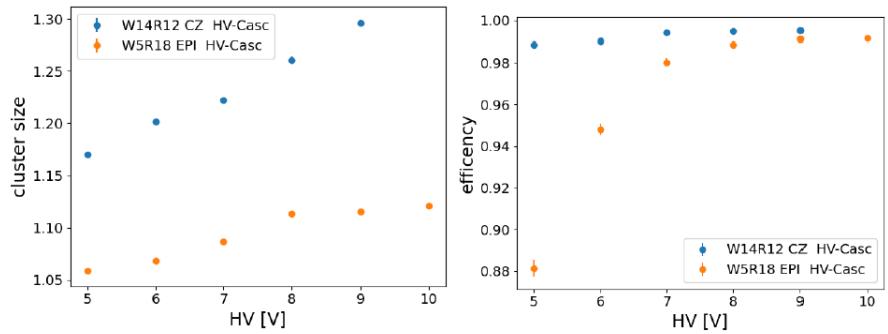


Figure 5.43: Cluster size and efficiency results for HV Cascode FE.

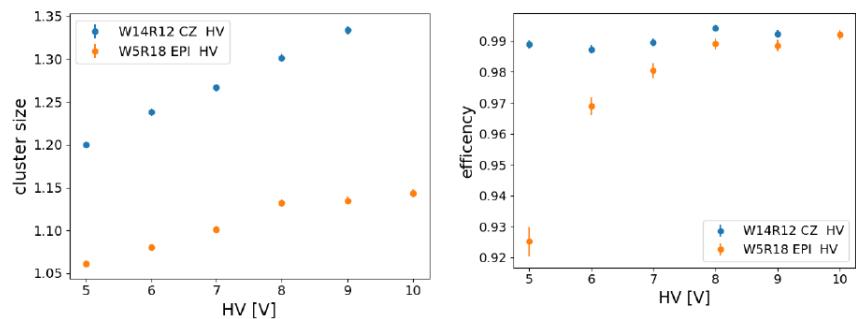


Figure 5.44: Cluster size and efficiency results for HV FE.

6. Conclusions

Bibliography

- [1] E. Kou et al. The Belle II Physics Book. *Progress of Theoretical and Experimental Physics*, 2019(12), December 2019.
- [2] M. Baszczyk et al. SuperB Technical Design Report. 6 2013.