

UNIVERSITÉ DE GENÈVE

Département de physique nucléaire et
corpusculaire

FACULTÉ DES SCIENCES

Professeur Giuseppe Iacobucci

Picosecond Time Stamping in Fully-Monolithic Highly-Granular Silicon Pixel Detectors

THÈSE

Présentée à la Faculté des sciences de l'Université de Genève
Pour obtenir le grade de Docteur ès sciences, mention physique

par

Matteo MILANESIO

Santa Vittoria d'Alba (Italie)

Thèse N° 5918

GENÈVE

Atelier d'impression ReproMail
2025



**UNIVERSITÉ
DE GENÈVE**

FACULTÉ DES SCIENCES

DOCTORAT ÈS SCIENCES, MENTION PHYSIQUE

Thèse de Monsieur Matteo MILANESIO

intitulée :

**«Picosecond Time Stamping in Fully-Monolithic
Highly-Granular Silicon Pixel Detectors »**

La Faculté des sciences, sur le préavis de

Monsieur G. IACOBUCCI, professeur ordinaire et directeur de thèse
Département de physique nucléaire et corpusculaire

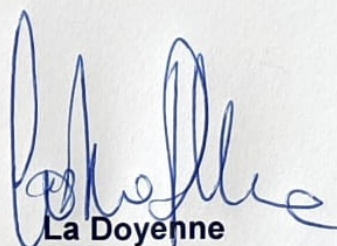
Monsieur F. SANCHEZ NIETO, professeur ordinaire
Département de physique nucléaire et corpusculaire

Monsieur W. RIEGLER, docteur
Département de physique, Conseil Européen de Recherche Nucléaire (CERN),
Genève, Suisse

autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 30 juin 2025

Thèse - 5918 -



La Doyenne

UNIVERSITÉ DE GENÈVE

Abstract

FACULTÉ DES SCIENCES
Département de physique nucléaire et corpusculaire

Docteur ès sciences, mention physique

Picosecond Time Stamping in Fully-Monolithic Highly-Granular Silicon Pixel Detectors

by Matteo MILANESIO

This thesis presents the development and characterization of monolithic silicon detectors with precise time resolution developed in the context of the MONOLITH project. Three generations of prototypes were fabricated using SiGe BiCMOS technology, using hexagonal pixels with a pitch of $100\text{ }\mu\text{m}$ and innovative sensor designs to achieve outstanding efficiency and timing performance while addressing power consumption and radiation tolerance challenges.

The first proof-of-concept PicoAD prototype, featuring a double-junction gain layer, demonstrated a detection efficiency of 99.9 % and a time resolution of 17.3 ps across the full pixel area at a power density of 2.7 W/cm^2 . Performance remained excellent down to 0.4 W/cm^2 , proving the viability of the continuous gain layer for efficient, low-power monolithic detectors.

A second prototype, produced without a gain layer, confirmed the potential of SiGe HBT front-end electronics. Tested with a 120 GeV/c pion beam, it achieved a 20 ps time resolution at 0.9 W/cm^2 , with minimal dependence on hit position. After proton irradiation up to $1 \times 10^{16}\text{ n}_{\text{eq}}/\text{cm}^2$, a detection efficiency above 99.6 % and time resolutions below 50 ps were maintained by adjusting bias voltages. Measurements with a femtosecond laser helped disentangle the various contributions to the time resolutions, supported by a dedicated Garfield++ simulation framework that successfully reproduced experimental results.

A third prototype, combining the second-generation ASIC with a PicoAD-like gain layer, achieved detection efficiencies up to 99.99 % and time resolutions of 12.1 ps at 2.6 W/cm^2 . Remarkably, at power densities below 1 W/cm^2 , time resolutions of approximately 18 ps were obtained, confirming the suitability of these devices for large-area applications like 4D trackers.

SiGe BiCMOS-based monolithic pixel detectors demonstrated scalability, radiation tolerance, and performance, establishing this technology as a promising solution for future collider experiments and high-precision timing applications, from particle physics to LiDAR. Integration with a low-gain avalanche may further enhance time resolution towards the picosecond time stamping.

Keywords: Timing, Detectors, Pixel, Monolithic, SiGe, BiCMOS, HEP, Front-end Electronics, MAPS, Gain layer, PicoAD, MONOLITH, Picosecond.

UNIVERSITÉ DE GENÈVE

Résumé

FACULTÉ DES SCIENCES

Département de physique nucléaire et corpusculaire

Docteur ès sciences, mention physique

Picosecond Time Stamping in Fully-Monolithic Highly-Granular Silicon Pixel Detectors

by Matteo MILANESIO

Cette thèse présente le développement et la caractérisation de détecteurs monolithiques au silicium dans le cadre du projet MONOLITH. Trois générations de prototypes ont été fabriquées en utilisant la technologie SiGe BiCMOS, en implementant des pixels hexagonaux avec un pas de $100\ \mu\text{m}$ pour atteindre une efficacité ainsi qu'une performance temporelle remarquables, tout en relevant les défis de la consommation d'énergie et de la tolérance aux niveaux de radiations des expériences de physique des particules à haute énergie.

Le premier prototype PicoAD, doté d'une couche de gain à double jonction, a démontré une efficacité de détection de 99.9 % et une résolution temporelle de 17.3 ps sur l'ensemble de la surface du pixel à une densité de puissance de $2.7\ \text{W}/\text{cm}^2$. Les performances sont restées excellentes jusqu'à $0.4\ \text{W}/\text{cm}^2$, ce qui prouve la viabilité de la couche de gain continue pour les détecteurs monolithiques efficaces et à faible consommation.

Un deuxième prototype, produit sans couche de gain interne, a confirmé le potentiel de l'électronique de frontend basés sur la technologie SiGe HBT. Testé avec un faisceau de pions de $120\ \text{GeV}/c$, ce prototype a atteint une résolution temporelle de 20 ps à $0.9\ \text{W}/\text{cm}^2$, avec une dépendance minimale de la position de l'impact. Après une irradiation contrôlée avec des protons cumulant jusqu'à un fluence de $1 \times 10^{16}\ n_{\text{eq}}/\text{cm}^2$, une efficacité de détection supérieure à 99.6 % et des résolutions temporelles inférieures à 50 ps ont été maintenues. Les mesures effectuées avec un laser femtoseconde ont permis de démêler les différentes contributions aux résolutions temporelles, soutenues par un cadre de simulation Garfield++ dédié qui a reproduit avec succès les résultats expérimentaux.

Un troisième prototype, combinant l'ASIC de deuxième génération avec une couche de gain de type PicoAD, a atteint des efficacités de détection allant jusqu'à 99.99 % et des résolutions temporelles de 12.1 ps à $2.6\ \text{W}/\text{cm}^2$. Fait remarquable, à des densités de puissance inférieures à $1\ \text{W}/\text{cm}^2$, des résolutions temporelles d'environ 18 ps ont été obtenues, ce qui confirme l'adéquation de ces dispositifs pour des applications à grande surface telles que les trackers 4D.

Les détecteurs de pixels monolithiques basés sur le SiGe BiCMOS ont démontré leur évolutivité, leur tolérance aux rayonnements et leurs performances, faisant de cette technologie une solution prometteuse pour les futures expériences de collisionneurs et les applications de chronométrage de haute précision, de la physique des particules au LiDAR. L'intégration avec une avalanche à faible gain peut encore améliorer la résolution temporelle vers l'horodatage à la picoseconde.

Mots-clés: Timing, Detectors, Pixel, Monolithic, SiGe, BiCMOS, HEP, Front-end Electronics, MAPS, Gain layer, PicoAD, MONOLITH, Picosecond.

Publications

Publications

Some ideas and figures have appeared previously in the following publications:

- L. Paolozzi et al. “Picosecond Avalanche Detector — Working Principle and Gain Measurement with a Proof-of-concept Prototype”. In: *Journal of Instrumentation* 17.10 (Oct. 2022), P10032. DOI: [10.1088/1748-0221/17/10/P10032](https://doi.org/10.1088/1748-0221/17/10/P10032)
- G. Iacobucci et al. “Testbeam Results of the Picosecond Avalanche Detector Proof-of-concept Prototype”. In: *Journal of Instrumentation* 17.10 (Oct. 2022), P10040. DOI: [10.1088/1748-0221/17/10/P10040](https://doi.org/10.1088/1748-0221/17/10/P10040)
- M. Milanesio et al. “Gain Measurements of the First Proof-of-concept PicoAD Prototype with a ^{55}Fe X-ray Radioactive Source”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1046 (Jan. 2023), p. 167807. ISSN: 0168-9002. DOI: [10.1016/j.nima.2022.167807](https://doi.org/10.1016/j.nima.2022.167807)
- S. Zambito et al. “20 ps Time Resolution with a Fully Efficient Monolithic Silicon Pixel Detector without Internal Gain Layer”. In: *Journal of Instrumentation* 18.03 (Mar. 2023), P03047. DOI: [10.1088/1748-0221/18/03/P03047](https://doi.org/10.1088/1748-0221/18/03/P03047)
- M. Milanesio et al. “Radiation Tolerance of SiGe BiCMOS Monolithic Silicon Pixel Detectors without Internal Gain Layer”. In: *Journal of Instrumentation* 19.01 (Jan. 2024), P01014. DOI: [10.1088/1748-0221/19/01/P01014](https://doi.org/10.1088/1748-0221/19/01/P01014)
- M. Milanesio et al. “Time Resolution of a SiGe BiCMOS Monolithic Silicon Pixel Detector without Internal Gain Layer with a Femtosecond Laser”. In: *Journal of Instrumentation* 19.04 (Apr. 2024), P04029. DOI: [10.1088/1748-0221/19/04/P04029](https://doi.org/10.1088/1748-0221/19/04/P04029)
- T. Moretti et al. “Testbeam Results of Irradiated SiGe BiCMOS Monolithic Silicon Pixel Detector without Internal Gain Layer”. In: *Journal of Instrumentation* 19.07 (July 2024), P07036. DOI: [10.1088/1748-0221/19/07/P07036](https://doi.org/10.1088/1748-0221/19/07/P07036)
- L. Paolozzi et al. “Testbeam Characterization of a SiGe BiCMOS Monolithic Silicon Pixel Detector with Internal Gain Layer”. In: *Journal of Instrumentation* 20.04 (Apr. 2025), P04001. DOI: [10.1088/1748-0221/20/04/P04001](https://doi.org/10.1088/1748-0221/20/04/P04001)

*This, in fact, is the power of the imagination, which
combining the memory of gold with that of the mountain,
can compose the idea of a golden mountain.*

Umberto Eco, The Name of the Rose

*The greatest ideas are the simplest.
Now there was something to be done
they worked with passion.*

William Golding, Lord of the Flies

*To zio Domenico,
I finally crafted something with my hands.
As you taught me.*

Acknowledgements

Completing this work would not have been possible without the support and encouragement of many people to whom I am deeply grateful.

First and foremost, I would like to thank my entire working group. I'm deeply grateful to my supervisor, Peppe, for all the things I've learned from him, apart from the obvious physics-related topics, I've also gained many insights into how a good academic environment works. Thank you for creating such a collaborative and stimulating research environment. I've been inspired to create the same atmosphere in my future endeavors. I am especially grateful to Lorenzo and Magdalena, who guided me through the most perilous steps of my journey. I learnt a lot from you both. A special thank you goes to my colleagues, who were always ready to discuss and, most importantly, to laugh with me.

I want to thank my family, including my parents, brothers, uncles, aunts, and cousins. We are a big family. Bad times are beyond us, and I hope we can find tranquility and peace in our future. Thank you, *nonna*, for being so curious about what I do; you have always been the most interested one, and you are, and will always be, the target of my outreach. Marco and Enrico: I hope you can find your path and understand what I have come to realize is more relevant in life.

I would also like to thank my friends from Italy. I'm thrilled that we've stayed connected over the years. Every time we meet, no matter how long it's been, it feels as natural and familiar as if only a day had passed. In a way, it's as though I never truly left Santa Vittoria, despite how much life has changed since. Roberta, you are a part of this group. I'm so grateful for the time we shared and for the support you provided during a significant part of my studies.

Finally, I would like to express my sincere appreciation to all the young physicists I have had the pleasure of meeting over the years, during conferences, schools, or simply around the CERN and university ambients. Many of you are still in the process of achieving a significant goal: some are pursuing their PhD, others are seeking grants to stay in academia, and still others are looking for jobs outside academia. I strongly believe in you and support you all. Meeting you enriched both my personal life and my research experience.

Contents

Abstract	i
Resumé	iii
Publications	v
Acknowledgments	vii
I Background	1
1 Introduction	3
1.1 Particle Physics Background	4
1.1.1 The Standard Model of Elementary Particles	4
1.1.2 History and Achievements of the Standard Model	6
1.1.3 Physics Beyond the Standard Model	8
1.1.4 Present and Future Particle Colliders	10
1.2 SiGe BiCMOS Monolithic Pixel Detectors	11
1.2.1 Motivations of the MONOLITH Project	11
1.2.2 Purpose of the MONOLITH Project	12
1.2.3 PicoAD Sensor Concept	12
1.2.4 PicoAD Wafer Preparation	14
1.2.5 PicoAD Proof-of-Concept Prototype	14
1.2.6 MONOLITH Prototype 2 without a Gain Layer	16
1.2.7 MONOLITH Prototype 2 with a Gain Layer	18
1.2.8 The Printed Circuit Board	19
1.2.9 Other Projects with SiGe BiCMOS Technology	20
2 Silicon-Based Devices	21
2.1 Physics of Semiconductors	22
2.1.1 Energy Bands	22
2.1.2 Charge Carrier Concentration	24
2.1.3 Intrinsic Semiconductor and Doping	25
2.1.4 Mobility, Drift Velocity, and Resistivity	27
2.1.5 Avalanche Due to Impact Ionization	29
2.2 The p - n Junction	30
2.2.1 Electrostatic of the p - n Junction	30
2.2.2 Electric Field and Built-in Potential of the p - n Junction	31
2.2.3 Depletion Depth of the p - n Junction	33
2.2.4 Depletion Capacitance of the p - n Junction	33
2.2.5 External Bias to the p - n Junction	34
2.2.6 The p - n Junctions for the MONOLITH Project	35
2.3 Transistors	37

2.3.1	Bipolar Junction Transistors	37
2.3.2	Heterojunction Bipolar Transistors	39
2.3.3	Metal-Oxide-Semiconductor Field Effect Transistors	40
2.3.4	High Voltage CMOS Devices	41
3	Silicon-Based Particle Detectors	43
3.1	Brief History of Silicon-Based Particle Detectors	44
3.2	Interaction of Particles with Matter	45
3.2.1	Interaction of Charged Particles with Matter	45
3.2.2	Interaction of Photons with Matter	48
3.3	The Shockley-Ramo Theorem	50
3.4	Readout Electronics	52
3.4.1	Signal Amplification and Shaping	53
3.4.2	Noise Sources	54
3.5	Time Resolution	55
3.5.1	Landau Fluctuations Contribution to the Time Resolution	57
3.5.2	Time Walk Contribution to the Time Resolution	58
3.5.3	Noise Contribution to the Time Resolution	59
3.5.4	Other Contributions to the Time Resolution	60
3.6	Spatial Resolution	60
3.6.1	Rectangular Pixel	61
3.6.2	Hexagonal Pixel	62
3.6.3	Comparison between Squared and Hexagonal Pixels	63
3.7	Power Consumption	65
II	Characterization of the Prototypes	67
4	PicoAD Proof-of-Concept Prototype	69
4.1	TCAD Simulations	70
4.2	Laboratory Measurements	72
4.2.1	Probe Station Measurements	72
4.2.2	Measurements with a ^{55}Fe Radioactive Source	75
4.3	Test Beam	79
4.3.1	Experimental Setup	80
4.3.2	Data Samples	86
4.3.3	Signal Amplitudes	86
4.3.4	Efficiency	88
4.3.5	Time Resolution	92
5	MONOLITH Prototype 2 without a Gain Layer	101
5.1	Laboratory Measurements	102
5.1.1	Experimental Setup	102
5.1.2	ASICs Characteristics	102
5.1.3	Characterization Before the Femtosecond Laser Measurements	105
5.1.4	Characterization Before the Garfield++ Simulations	106
5.2	Test Beam	107
5.2.1	Experimental Setup	108
5.2.2	Data Samples	109
5.2.3	Efficiency	109
5.2.4	Time Resolution	113

5.3	Femtosecond Laser	117
5.3.1	Experimental Setup	118
5.3.2	Time Resolution of the Reference ASIC	122
5.3.3	Time Resolution of the DUT	123
5.3.4	Comparison Between Laser and Test-beam Time Resolutions	126
5.4	Radiation Tolerance	129
5.4.1	Proton Irradiation of the SiGe HBT ASICs	129
5.4.2	Characterization with a ^{90}Sr Radioactive Source	132
5.4.3	Estimation of the Time Jitter of the Signal	134
5.4.4	Test Beam Experimental Setup	138
5.4.5	Test Beam Data Samples	139
5.4.6	Efficiency	140
5.4.7	Time Resolution	142
5.5	Garfield++ Simulations	146
5.5.1	The Simulation Framework	147
5.5.2	Test Beam Time Resolution Matching	147
5.5.3	Charge-Collection Noise Estimation	153
5.5.4	Non-Gaussianity of the Distributions of Time of Arrival	155
6	MONOLITH Prototype 2 with a Gain Layer	159
6.1	Laboratory Measurements	159
6.2	Test Beam	161
6.2.1	Experimental Setup	162
6.2.2	Data Samples	162
6.2.3	Signal Amplitudes	163
6.2.4	Efficiency	163
6.2.5	Time Resolution	167
III	Conclusions	173
7	Conclusions	175
	Bibliography	179

Part I

Background

Chapter 1

Introduction

The goal of this chapter is twofold. First, it wants to provide a comprehensive overview of the theoretical framework and technological motivations behind the research presented in this thesis. Second, it introduces the innovative monolithic pixel silicon detectors developed at the University of Geneva, utilizing SiGe BiCMOS technology. The research and development conducted, particularly within the context of the MONOLITH ERC advanced project, aimed to enhance the performance of these detectors by improving their timing resolution through the use of cutting-edge electronics and sensor design.

Section 1.1 offers an outline of the significant milestones in the development of particle physics, culminating in the Standard Model. Despite its remarkable success in describing the known fundamental particles and interactions (excluding gravity), the Standard Model remains incomplete. We highlight some of the most pressing open questions, such as the nature of dark matter, the origin of neutrino masses, and the matter-antimatter asymmetry, which strongly motivate physics beyond the Standard Model. The section also introduces the landscape of current and future particle colliders, such as the present *Large Hadron Collider* (LHC) at CERN, its upgrade, the *High-Luminosity Large Hadron Collider* (HL-LHC), and the proposed *Future Circular Collider* (FCC). These colliders will have stringent requirements for tracking detectors in terms of time and spatial resolution, efficiency, radiation hardness, and low power consumption.

Section 1.2 introduces the monolithic pixel silicon detectors developed at the University of Geneva, utilizing SiGe BiCMOS technology, particularly in the context of the MONOLITH project. This section outlines the motivations and purpose of the MONOLITH project, a collaborative effort at the *Département de Physique Nucléaire et Corpusculaire* at the University of Geneva, aimed at designing and testing ultra-fast monolithic silicon sensors for particle detection. The project leverages analog front-end electronics built with SiGe transistors and integrated in a BiCMOS process. The section then introduces the concept of the PicoAD sensor, an innovative sensor architecture optimized for precise timing measurements. The MONOLITH research project is funded by the H2020 ERC Advanced grant 884447. Finally, this section shifts the focus to a broader context. Other research initiatives have been conducted in parallel within the same research group, adopting the same SiGe BiCMOS technology as the MONOLITH project, demonstrating its versatility and increasing interest, and highlighting their potential.

1.1 Particle Physics Background

Particle physics aims to comprehend the fundamental constituents of matter and the forces that govern their interactions. Finding them has driven the development of increasingly sophisticated experimental apparatus and theoretical frameworks in the past century. Among these, the Standard Model (SM) theory is one of the most outstanding achievements in recent physics. It accurately describes the known elementary particles and their interactions, apart from gravity. However, some issues suggest that it is not the final theory.

The journey of particle physics began in the twentieth century with the discovery of the quantum nature of particles, the nucleus, and the electron. As experiments got better, so did our understanding of the subatomic universe. Important milestones, such as the discovery of the neutron, the positron, and the development of quantum electrodynamics, enabled us to construct a unified framework. This theoretical and experimental work, combined, led to the creation of the Standard Model, which encompasses all known elementary particles and the three fundamental forces: electromagnetic, weak, and strong interactions.

The Standard Model was not constructed in one go, but rather emerged through decades of theoretical predictions followed by experimental validations, including the discoveries of the W and Z bosons, the top quark, and the Higgs boson. High-energy collider experiments, such as the Large Hadron Collider (LHC), continue to test the model's predictions with greater accuracy today.

The concepts of the Standard Model presented in this section are primarily based on the book [9] and the 2024 annual review of the Particle Data Group [10]. For the present and future colliders, some ideas have been taken from [11].

1.1.1 The Standard Model of Elementary Particles

The Standard Model (SM) is the quantum field theory that categorizes all known elementary particles and describes their interaction. Elementary particles are subdivided into two categories based on the quantum property of the particle's spin.

The first category is made of fermions, particles with spin $1/2$ that follow the statistics of Fermi-Dirac [12]. Fermions obey the Pauli exclusion principle. Hence, two identical particles cannot occupy the same quantum state. Fermions make up matter.

The second category is made of bosons, particles with integer spin that follow the statistics of Bose-Einstein [13]. Bosons don't obey the Pauli exclusion principle, and two identical bosons can occupy the same quantum state, leading to superconductivity and Bose-Einstein condensation. Bosons mediate the three fundamental interactions: strong, weak, and electromagnetic interactions.

Fermions are further divided into quarks and leptons, each coming in three generations. Quarks carry color charge and interact via the strong force, while leptons do not. All the fermions interact via the weak force, and all the charged fermions interact via the electromagnetic force, hence excluding the neutrinos. Fermions are composed of matter particles and their respective antimatter particle, with the same mass and spin but opposite charge.

Bosons are further divided into vector bosons, i.e., with spin 1, and scalar bosons, i.e., with spin 0. The vector bosons are the gluons (mediators of the strong force), the W and Z bosons (mediators of the weak force), and the photon (mediator of the electromagnetic force). The only known scalar boson is the Higgs boson, whose interaction with the W and Z bosons and the fermions is responsible for their mass.

Figure 1.1 shows a summary of the elementary particle content of the Standard Model. The subdivisions between fermions and bosons and quarks and leptons are color-coded. Fermions are grouped by a faint line with the bosons that mediate their interactions. This means, for example, that quarks are connected to the gluon, all the charged particles with the photon, and all the fermions with the W and Z bosons. Each particle's mass, charge, and spin are also indicated.

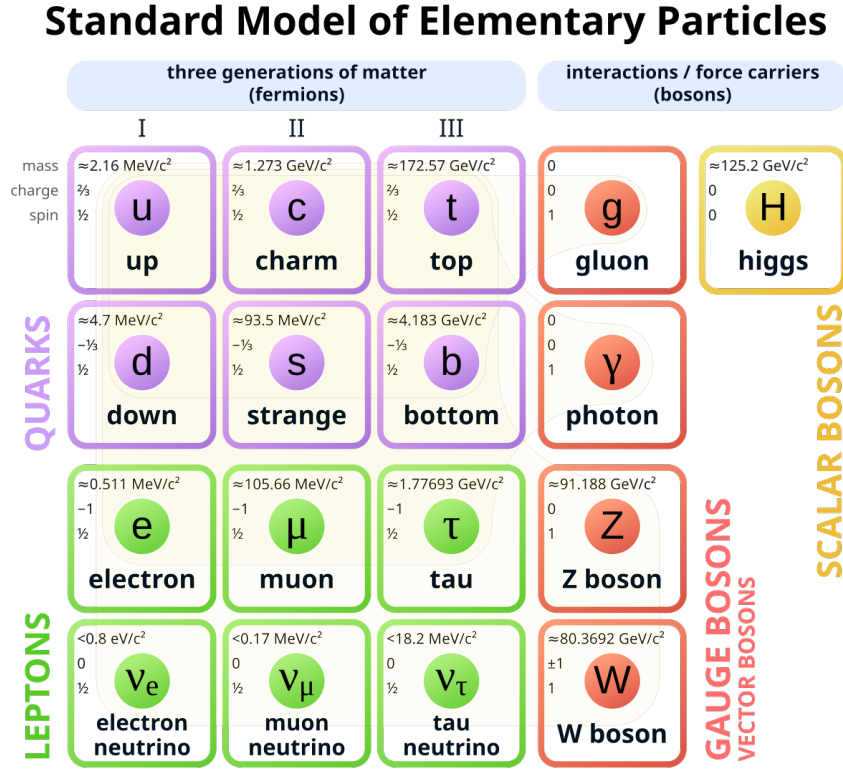


FIGURE 1.1: Elementary particles content of the Standard Model theory. The color code underlies the category of appartenance. A faint line groups together the fermions with the bosons that mediate their interactions.

All the fundamental interactions derive from one basic principle, the requirement of local gauge invariance, i.e., the invariance of the physical system under certain local transformations. For electromagnetism, the physics does not change if the phase of a charged particle's wavefunction is altered. For the weak interaction, the physics does not change with the rotation of the components of weak isospin doublets (like the electron and its neutrino). Finally, for the strong interaction, the physics remains unchanged if quarks are rotated in the color space.

Each invariance is related to a gauge group, a gauge field, and the corresponding gauge boson (or bosons). The gauge group structure of the overall Standard Model is the unification of the three gauge groups $SU(3)_C \times SU(2)_L \times U(1)_Y$, corresponding to the strong, weak, and electromagnetic interactions, respectively. The gauge group of the electromagnetic force is $U(1)$ with one gauge boson, which is the photon. The number of gauge bosons of a generic $SU(N)$ group is $N^2 - 1$, with N being the dimension of the matrices representing the group elements. Hence, the gauge group of the weak interaction is $SU(2)$ with three gauge bosons, the Z and W bosons (the W boson can be positively and negatively charged). The gauge group of the strong interaction is $SU(3)_C$ with eight gauge bosons, which are the eight gluons.

A central feature of the Standard Model (SM) is the Higgs mechanism, which is responsible for the masses of the W and Z bosons, as well as the fermions. The Higgs field also breaks the electroweak symmetry, giving rise to the observed differences between the weak and electromagnetic forces.

Gravity is excluded from the Standard Model because it fundamentally differs from the other three fundamental forces of nature. While these forces are integrated within the framework of quantum field theory, gravity is described by Einstein's general relativity. This classical theory does not conform to the principles of quantum mechanics. Attempts to quantize gravity, which have led to the theoretical prediction of a particle known as the graviton, have faced significant challenges and lack experimental support. Moreover, gravity's effects are negligible at the subatomic scale, where the Standard Model operates: the gravitational force between a proton and an electron is approximately 10^{-36} times weaker than the electromagnetic force between them, rendering gravity's influence negligible in particle interactions at this scale.

Modern particle physics experiments continue to test the limits of the Standard Model and seek evidence for phenomena beyond it. The precision achieved in both theoretical predictions and experimental measurements has reached a level where even slight deviations from the Standard Model can be significant, potentially indicating the presence of new physics.

1.1.2 History and Achievements of the Standard Model

Since its formulation in the 1960s and 1970s, the Standard Model has provided an exceptionally accurate description of particle interactions, unifying the electromagnetic, weak, and strong forces within a single quantum-field framework. Its principal triumphs include the theoretical unification of forces, the prediction and discovery of fundamental particles, and precision tests that continue to validate its structure. These successes have been recognized by multiple Nobel Prizes, both for theoretical insights and experimental discoveries.

Figure 1.1 shows a timeline of the elementary particle content of the Standard Model, emphasizing the time passed between the theoretical prediction and experimental discovery (or, vice versa, between experimental discovery and theoretical explanation). In the following sections, we will detail the significant milestones of the Standard Model.

The ancestors of the Standard Model are considered to be the quantum electrodynamics (QED) theory, formulated by Tomonaga [14], Schwinger [15], and Feynman [16] in the late 1940s, and the Fermi theoretical formulation of beta decay in 1934 [17]. QED was the first fully relativistic formulation of a quantum field theory explaining the electromagnetic interaction between the electron and the photon. Tomonaga, Schwinger, and Feynman were awarded the Nobel Prize in Physics in 1965. Fermi's theoretical formulation of beta decay, in which he introduced a four-fermion contact interaction and the concept of the neutrino to conserve energy and momentum in beta emission, laid the groundwork for the physics of weak interaction.

The successive discovery of new elementary particles provided experimental confirmation of the Fermi theory. In particular, the muon was discovered in 1937 by Anderson and Neddermeyer [18]. The electron neutrino was discovered in 1956 by Cowan and Reines [19], for which Reines was awarded half the Nobel Prize in 1995. The muon neutrino was discovered in 1962 by Lederman, Schwartz, and Steinberger [20]. They were awarded the 1988 Nobel Prize for the neutrino beam method and for demonstrating the doublet structure of leptons by discovering the muon neutrino.

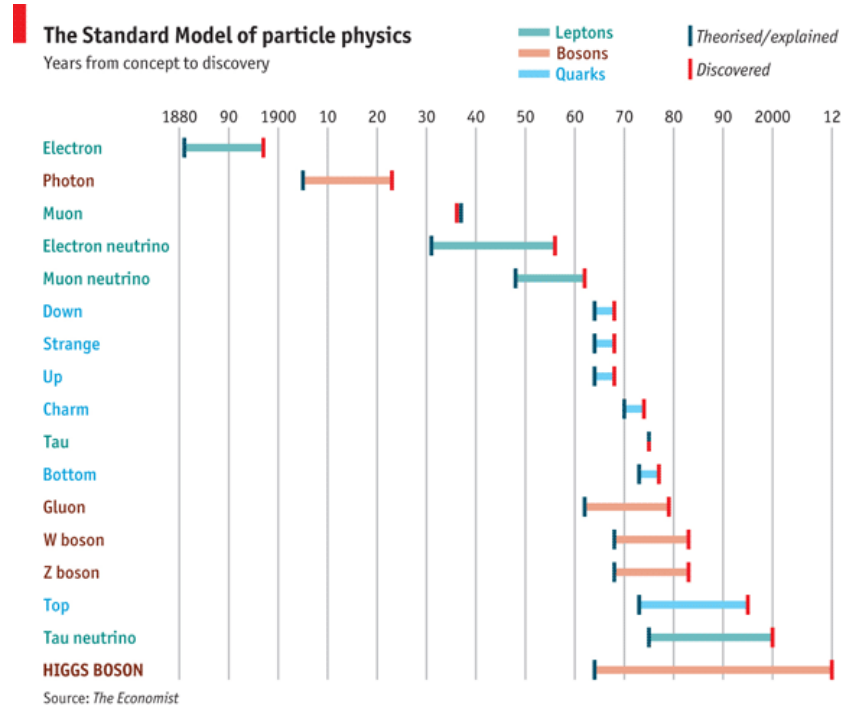


FIGURE 1.2: Timeline of the elementary particle content of the Standard Model theory. The color code underlies the category of appartenance: leptons, quarks (both fermions), or bosons.

The new particles also clarified that a unified theory of electromagnetic and weak interactions was needed. The foundations of the Standard Model as we know it today can be traced back to Glashow, who in 1961 formulated the original $SU(2)_L \times U(1)_Y$ gauge theory for electroweak interactions, introducing the neutral weak current and laying the groundwork for unified forces [21]. Weinberg, in 1967, incorporated spontaneous symmetry breaking into this gauge theory, predicting the masses of the W and Z bosons [22]. Salam independently developed the same framework in 1968 [23]. Glashow, Weinberg, and Salam were awarded the Nobel Prize in 1979 for their contributions to the theory of unified weak and electromagnetic interactions. The Gargamelle bubble chamber experiments at CERN first observed neutrino-induced weak neutral current events in 1973 [24, 25], confirming a key prediction of the electroweak theory, which had never been observed before.

Veltman and 't Hooft in the early 1970s showed how a non-Abelian gauge theory for the weak and electromagnetic forces could be consistently renormalized, placing the electroweak sector of the Standard Model on a firm mathematical footing [26]. Their work made it possible to perform precise calculations of physical observables, including the masses and couplings of the W and Z bosons, which were subsequently confirmed by experiments at CERN's LEP collider and Fermilab's Tevatron. They were jointly awarded the 1999 Nobel Prize in Physics.

The GIM mechanism, proposed by Glashow, Iliopoulos, and Maiani in 1970 [27], elegantly suppresses flavor-changing neutral currents by requiring the necessity of a fourth quark. In the so-called *November Revolution*, Brookhaven's group, guided by Ting [28], and SLAC's group, guided by Richter [29], simultaneously observed the J/ψ meson on 11 November 1974, confirming the existence of a fourth quark, called the charm quark, and consequently the GIM mechanism. Ting and Richter were awarded the Nobel Prize in 1976 for this breakthrough.

This insight explained the historic CP violation observed in neutral kaon decays by Cronin and Fitch in 1964 [30] (for which they received the Nobel Prize in 1980). In 1973, Kobayashi and Maskawa [31] extended this framework by showing that CP violation could be accommodated only if there were at least three generations of quarks, leading to the 3×3 unitary CKM matrix that governs quark mixing and CP-violating phases. Kobayashi and Maskawa were awarded the Nobel Prize for these theoretical breakthroughs in 2008. The quark mixing proposed by Kobayashi and Maskawa predicted the existence of bottom and top quarks, whose eventual discoveries completed the SM's three-generation structure. The bottom quark was discovered by observing the Y meson at Fermilab by Lederman in 1977 [32]. The top quark was discovered in 1995 by the CDF [33] and D0 [34] collaborations at the Tevatron.

Gross, Wilczek [35], along with Politzer [36], received the Nobel Prize in 2004 for discovering in 1973 that the non-Abelian gauge theory $SU(3)_C$ becomes weakly coupled at high energies, establishing a consistent theory of the strong force. This phenomenon, called *asymptotic freedom*, explains why isolated quarks have never been observed.

UA1 and UA2 collaborations at CERN's SPS collider independently reported the discovery of the W boson in early 1983 [37, 38], followed by the discovery of the Z boson later that same year [39, 40]. Rubbia and Van der Meer were awarded the Nobel Prize in 1984 for their decisive roles in this discovery.

The SM's three-generation structure was completed on the quark side, but not yet on the lepton side. The tau lepton was discovered in 1975 by Perl and collaborators at SLAC in the process of electron-positron annihilation [41]. Perl was awarded one-half of the 1995 Nobel Prize. Finally, in 2000, the DONUT experiment at Fermilab reported the discovery of the tau neutrino [42], completing the Standard Model's three-generation structure for both leptons and quarks.

In 1964, Higgs demonstrated how a scalar field can endow gauge bosons with mass while preserving renormalizability [43] through the so-called *Higgs mechanism*. Spontaneous symmetry breaking is related to the creation of a massive scalar boson, known as the Higgs boson. ATLAS [44] and CMS [45] experiments at the LHC announced the observation of a Higgs-like boson in 2012, confirming the mechanism of electroweak symmetry breaking. He was awarded the Nobel Prize in 2013, along with Englert, following the 2012 discovery of a Higgs-like boson at the Large Hadron Collider (LHC), which confirmed the Higgs mechanism. This confirmation brought to a close many years of experimental observations carried out with a global effort. The Standard Model was now complete, consistent, and highly predictive.

1.1.3 Physics Beyond the Standard Model

Despite its extreme success, the SM is known to be incomplete. It does not include a quantum theory of gravity, nor does it explain the nature of dark matter, dark energy, the matter-antimatter asymmetry, the hierarchy problem, or the origin of neutrino masses. Nevertheless, it remains the best-tested theory of particle interactions, providing a solid foundation for current and future explorations.

The incompatibility between general relativity and quantum mechanics, coupled with gravity's negligible impact on particle-scale phenomena, has led to its exclusion from the Standard Model of particle physics. This omission highlights the model's limitations and underscores the need for a more comprehensive theory that unifies all fundamental forces, including gravity.

The SM does not explain the existence of dark matter and energy, constituting over 95 % of the universe's energy content. Dark matter has multiple lines of evidence, including galactic rotation curves [46], gravitational lensing [47], and cosmic microwave background anisotropies [48]. The earlier COBE discovery of primordial anisotropies [49] led to the 2006 Nobel Prize in Physics to Smoot and Mather. Roughly 85 % of the matter in the universe is non-luminous and weakly interacting, properties incompatible with the SM's neutrinos or baryons. Leading candidates, ranging from weakly interacting massive particles (WIMPs) to axions or primordial black holes, lie outside the Standard Model (SM) spectrum, motivating a broad range of experimental programs that include underground detectors, colliders, and astronomical searches. Dark energy was evidenced by observations of distant Type Ia supernovae in the late 1990s [50], which revealed that the universe's expansion is accelerating, implying the presence of a dominant energy component. The simplest candidate is a cosmological constant as quantum vacuum energy, but contributions calculated within the Standard Model framework exceed the observed value by around 120 orders of magnitude.

Another puzzle is the matter-antimatter asymmetry in the universe. Although the SM includes CP violation in the quark sector via the CKM matrix, quantitative analyses show it falls far short of producing the observed baryon asymmetry of the universe. Sakharov [51] pointed out that three conditions are necessary to explain the matter-antimatter symmetry: baryon number violation, C and CP violation, and departure from equilibrium. These conditions can, in principle, occur within the SM, but do so too weakly, pointing toward new sources of CP violation or novel baryogenesis mechanisms.

The naturalness problem, also known as the hierarchy problem, questions why the mass of the Higgs boson is so much lighter than the Planck scale [52]. Quantum corrections to the Higgs boson's mass in the SM scale quadratically with the cutoff energy, leading to an enormous fine-tuning if that cutoff approaches the Planck scale. Solutions such as supersymmetry, composite Higgs models, or extra dimensions introduce new particles or dynamics near the TeV scale to stabilize the Higgs mass. Still, experimental confirmation at the LHC has yet to be found.

Neutrino oscillation experiments, such as Super-Kamiokande [53] and SNO [42], have definitively shown that neutrinos have nonzero masses. The 2015 Nobel Prize was awarded jointly to Kajita (Super-Kamiokande) and McDonald (SNO) for this reason. The smallness of the neutrino masses suggests a high new-physics scale or suppressed couplings, motivating searches for neutrinoless double beta decay and precision oscillation studies.

These open questions drive the search for new physics through direct searches at colliders and indirect methods such as precision measurements and astrophysical observations.

1.1.4 Present and Future Particle Colliders

Protons, neutrons (made by quarks), and electrons are the basic stable constituents of ordinary matter. For more exotic particles, the primary production sources are cosmic rays and particle accelerators. Cosmic rays are a significant source of elementary particles, particularly those with enormous energies. Still, the rate at which they strike any detector of reasonable size is very low, and they are completely uncontrollable. For these reasons, particle accelerators are the preferred choice.

With particle accelerators, everything starts with electrons or protons, which are accelerated to high energy, and smashed into a target or in head-on collisions. Head-on collisions have a significant advantage in center-of-mass energy, enabling the production of heavier particles. The heavier the particle to be created, the higher the energy of the collision must be. That's why, historically, lightweight particles were often discovered first (see Figure 1.2).

Many contemporary experiments involve colliding beams from intersecting storage rings. Unfortunately, when a charged particle accelerates, it radiates, thereby losing energy. In the case of circular motion, this phenomenon is known as synchrotron radiation, which severely limits the efficiency of storage rings for accelerated particles. However, the bigger the radius of the circular collider, the less synchrotron radiation. Thus, historically, bigger and bigger circular colliders were built.

The largest particle collider in the world is the Large Hadron Collider (LHC) at CERN. It is currently operational and has enabled significant discoveries, including the 2012 discovery of the Higgs boson. Its two general-purpose detectors, ATLAS and CMS, have played a key role in these achievements. These detectors are equipped with advanced tracking systems that allow for the precise reconstruction of particle trajectories. The near future of the LHC is the High-Luminosity LHC (HL-LHC), for which the same tunnel and machinery as the LHC will be used, but the experiment will be renewed and improved to sustain the higher rate of produced particles.

Looking ahead, future projects such as the Future Circular Collider (FCC) and the International Linear Collider (ILC) aim to provide higher collision energies and luminosities. These machines require even more sophisticated detectors, with improved spatial and temporal resolution, lower power consumption, and greater radiation hardness. Figure 1.3 shows an aerial view of the land around Geneva, with the comparison in size between the existing SPS and LHC, and the proposed FCC.

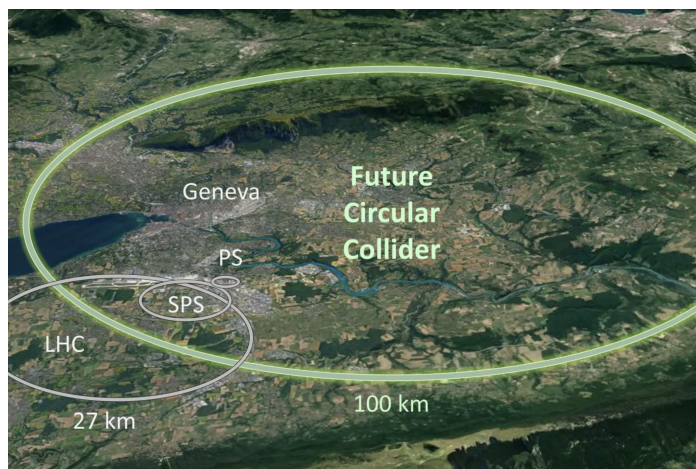


FIGURE 1.3: Aerial view of the land around Geneva, with the comparison in size between the existing SPS and LHC, and the proposed FCC.

1.2 SiGe BiCMOS Monolithic Pixel Detectors

The research group at the *Département de Physique Nucléaire et Corpusculaire* at the University of Geneva has been working with fast silicon pixel detectors for several years, exploiting the very high speed and low noise of SiGe BiCMOS electronics for multiple projects.

After a demonstrator with discrete components [54], time resolutions of 50 ps were achieved with the first monolithic silicon pixel detector prototype [55, 56] using the SG13G2 [57] 130 nm process available at IHP Microelectronics (*Leibniz Institut für High-Performance Microelectronics*) with 100 μm pixel pitch. These initial results were achieved at the cost of detection efficiency, as a relatively high discrimination threshold was imposed to maintain a good signal-to-noise ratio. A second monolithic silicon pixel ASIC prototype was produced in the same 130 nm IHP process, aiming to surpass the performance of the previous prototype. A time resolution of 36 ps and a detection efficiency of 99.9 % were measured [58].

In the context of this research, the MONOLITH H2020 ERC Advanced project [59] started in 2020. It introduces a novel sensor structure called the Picosecond Avalanche Detector (PicoAD) [60], devised to overcome the intrinsic limits of present p - n -junction sensors and provide simultaneous picosecond timing and high spatial resolution in a monolithic implementation. The fully depleted multi-junction contains a continuous deep gain layer fully decoupled from the pixel structures, thus permitting a small-pitch pixelated design.

Three monolithic silicon pixel detector prototypes were produced within the MONOLITH H2020 ERC Advanced Project [59]. The prototypes were all made by IHP Microelectronics, with which the group doing monolithic pixel detectors at the University of Geneva has a long history of collaboration. The three prototypes come from three different fabrication batches, and various factors, such as the ASIC design, doping profiles, and substrate, may change from one production to another. The prototypes listed in chronological order are:

- the *PicoAD Proof-of-Concept Prototype* (described in Section 1.2.5 and characterized in Chapter 4);
- the *MONOLITH Prototype 2 without a Gain Layer* (described in Section 1.2.6 and characterized in Chapter 5);
- the *MONOLITH Prototype 2 with a Gain Layer* (described in Section 1.2.7 and characterized in Chapter 6).

1.2.1 Motivations of the MONOLITH Project

With the advent of the High-Luminosity Large Hadron Collider (HL-LHC) program at CERN, precise time measurement becomes essential to cope with the large number of collisions per bunch crossing. The so-called *4D tracking* [61], i.e., the usage of the time measurement as information for the particle tracking, will be achieved at the HL-LHC using separate detectors [62, 63], which are specialized in either position or time measurement. This expensive and complex solution to the problem of significant pile-up events at the HL-LHC will not be sustainable in future hadron-collider programs. Therefore, a silicon detector with precise time and space resolutions is needed. Other fundamental requirements include low power consumption for simple cooling systems and a limited material budget to minimize multiple scattering and enhance tracking capabilities.

In the quest for the next generation of pixel sensors, the particle physics community working on silicon detectors is developing Monolithic Active Pixel Sensors (MAPS) [64], which contain the sensor in the same CMOS substrate used for the electronics. MAPS are particularly appealing since they offer all the advantages of industrial-standard processing, thereby avoiding the production complexity and high costs associated with bump-bonded hybrid pixel sensors commonly used in high-energy physics. So far, monolithic active pixel sensors with nanosecond time resolution have been adopted in the DEPFET [65] implementation in Belle-II [66]. At the same time, MAPS are used in STAR [67], Mu3e [68], and the upgraded ALICE tracker [69], which is installed to take data during LHC Run 3, and they were considered for the ATLAS ITk upgrade [70].

Producing thin sensors capable of providing excellent position and time resolution simultaneously requires going beyond the diode structure of present silicon pixel sensors. LGAD sensors with internal gain [71] reached the approximately 30 ps intrinsic limit imposed on diodes by the charge-collection noise. A promising research approach is represented by the resistive AC-coupled LGAD [72, 73], in which the n^+ implant is resistive and the LGAD is read out using AC coupling. The TIMESPOT project [74] adopts an alternative approach that exploits the 3D sensor structure to produce hybrid detectors, which are intrinsically radiation-tolerant. In the version with trenches, using SiGe discrete electronics amplification, these sensors achieved time resolutions of 11.5 ps [75]. The efficiency varies between 80 % for tracks perpendicular to the sensor and full efficiency for track angles larger than 10° .

1.2.2 Purpose of the MONOLITH Project

The MONOLITH project aims to design a novel silicon pixel sensor capable of detecting particles with picosecond time resolution and precise spatial resolution. Building on the knowledge gained from MAPS and LGAD, the MONOLITH project aims to incorporate aspects from both technologies, notably the simplified assembly process and reduced production costs offered by monolithic implementation in standard CMOS processes, as well as gain layer amplification to overcome the intrinsic limitations of silicon detectors.

The MONOLITH project relies on two key ingredients:

- the fast and low-noise amplification made in analog front-end electronics with *SiGe Heterojunction Bipolar Transistors* (SiGe HBTs, see 2.3.2);
- the novel sensor concept of the *Picosecond Avalanche Detector* (PicoAD) [60] devised to overcome the intrinsic limits of present p - n -junction sensors.

The MONOLITH project may provide the required breakthrough and offer a sustainable solution for fundamental research that requires precise tracking and picosecond time resolution, which includes future high-energy physics experiments at colliders, rare decay experiments featuring active targets or precise tracking, space-borne astroparticle physics experiments, and nuclear physics measurements of lifetimes of excited states in short-lived nuclei.

1.2.3 PicoAD Sensor Concept

The PicoAD introduces design flexibility to avalanche-based sensors by moving the gain layer away from the pixel matrix and separating the region where primary charge is produced and amplified from the area dedicated to signal induction on the

readout electrodes. Figure 1.4 shows the cross-section of the PicoAD. The detector is characterized by a fully depleted multi-junction $(n-p)_{\text{pixel}}-(n-p)_{\text{gain}}$ structure, which provides a new degree of freedom for engineering the electric field. Indeed, when the bias voltage is applied, the sensor depletion region expands from the two junctions that operate in inverse polarization: the one close to the top surface (pixel junction) serves to isolate the pixels, and the one close to the bottom edge of the depletion region (gain junction) provides an electric field large enough to generate impact ionization and, therefore, avalanche gain. The other two junctions isolate the central, direct junction until complete depletion is achieved.

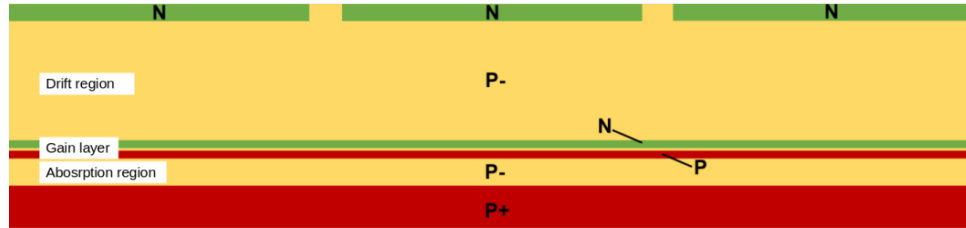


FIGURE 1.4: Simplified cross section of the PicoAD detector. The sensor presents n -type pixels on a high-resistivity epitaxial layer with a boron background. A second junction, deep inside the sensor volume, produces a continuous avalanche gain layer. The epitaxial layer and the deep junction are operated in full depletion.

When crossed by ionizing radiation, the primary holes generated in the region between the pixel and gain junctions will drift towards the gain junction, resulting in little or no amplification in silicon. In contrast, the primary electrons generated below the gain junction will drift upwards to be amplified, producing secondary charges. This behavior allows identifying three central regions:

- The primary *absorption region*, constituted by the bottom p -doped region, near the backside contact. This region does not produce impact ionization, but the PicoAD will amplify the e^- signal generated here.
- The uniform and continuous deep $n-p$ junction is the *gain layer*, operated in proportional mode.
- The *drift region* corresponds to the thicker top p -doped region between the pixels and the gain layer. Primary charges and secondary electrons moving in this region induce most of the signal current in the pixel above, according to the Shockley-Ramo theorem (see Section 3.3).

This sensor structure offers multiple advantages. First and foremost, the gain layer is a uniform implant far from the pixels. Therefore, it is less subject to substantial electric field variations and can grant uniform signal amplification on all detector surfaces. This characteristic enables reducing the pixel size with minimal impact on avalanche gain performance, as long as the inter-pixel distance is smaller than or comparable to the thickness of the drift region. Another advantage is that the absorption region can be arbitrarily thin, reducing charge-collection noise (see Section 3.5.1) without significantly increasing pixel capacitance. This fundamental feature enables the exploitation of advances in fast, low-noise electronics, which boosts the signal-to-noise ratio to achieve picosecond time resolution. Finally, the design of the pixel matrix is independent of the gain layer's presence, which simplifies the production of prototypes by allowing for the integration of the gain layer into an existing sensor design.

1.2.4 PicoAD Wafer Preparation

PicoAD prototypes were produced in a monolithic implementation by IHP Microelectronics. The manufacturing of the prototypes was divided into wafer preparation and CMOS processing using their SG13G2 130 nm SiGe BiCMOS process. The procedure adopted for the wafer preparation is depicted in Figure 1.5.

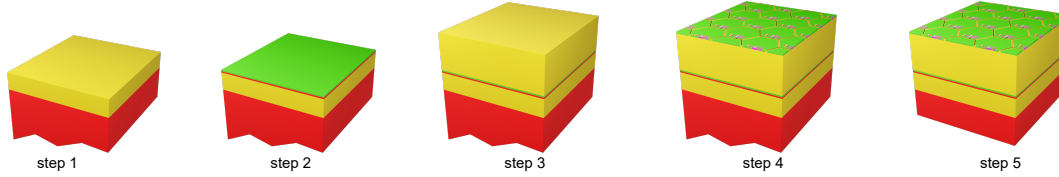


FIGURE 1.5: Manufacturing process of the PicoAD. Step 1: A high-resistivity epitaxial (absorption) p -layer is grown on a low-resistivity, boron-doped p -substrate. Step 2: A uniform n - p gain layer is implanted. Step 3: A second epitaxial p -layer corresponds to the sensor drift region. Step 4: The top surface is processed to produce the pixel matrix. In the case of a monolithic PicoAD implementation, this involves full CMOS processing. Step 5: The substrate can be thinned from the backside and metalized.

The PicoAD was produced on boron-doped p -substrates with a resistivity of $0.1 \Omega \text{ cm}$. To guarantee the uniformity of the electric field in the gain layer, the drift region of the PicoAD should be thicker than the inter-pixel distance of the matrix.

One of the risks associated with manufacturing the multi-junction structure is the presence of a direct junction at the transition between the drift and gain regions. When the epitaxial layer is fully depleted during sensor operation, this region is emptied of all free charges. However, this is not the case during the sensor bias ramp-up: at this stage, if the gain layer implantation reaches the sensor edge, a conductive path to the substrate may form where the silicon is cut, resulting in excess current in the device. To avoid this effect, it was decided to use an implantation mask to guarantee that the gain layer ended at $80 \mu\text{m}$ from the chip edge.

The wafers were produced using a dichlorosilane-based process available at IHP, which is optimized for sub-micron epitaxial processing at 850°C . This process is unsuitable for producing large volumes of wafers as it is time-consuming and has a non-negligible risk of introducing defects on the surface of the thicker epitaxial drift layer.

1.2.5 PicoAD Proof-of-Concept Prototype

The first prototype characterized in this thesis is the ASIC named the PicoAD proof-of-concept prototype. It was produced by the IHP Microelectronics foundry using the IHP SG13G2 130 nm BiCMOS technology. It was submitted to the foundry in 2020 and was received in 2021. It's called a proof-of-concept prototype because it was the first time the Picosecond Avalanche Detector (PicoAD) sensor concept was used to create an ASIC prototype.

Figure 1.6 shows the monolithic silicon pixel ASIC layout of the PicoAD proof-of-concept prototype. The four analog pixels are visible in the magnified view.

The ASIC has a total size of $2.3 \text{ mm} \times 2.5 \text{ mm}$. The pixel matrix consists of 12 rows by 12 columns of hexagonal pixels, totaling 144 pixels. Outside the pixel matrix lies the periphery of the ASIC, which features a slow-control interface for regulating the bias current supplied to the front-end electronics. The periphery is then insulated from the substrate through 18 guard rings. The pixel matrix is divided into four

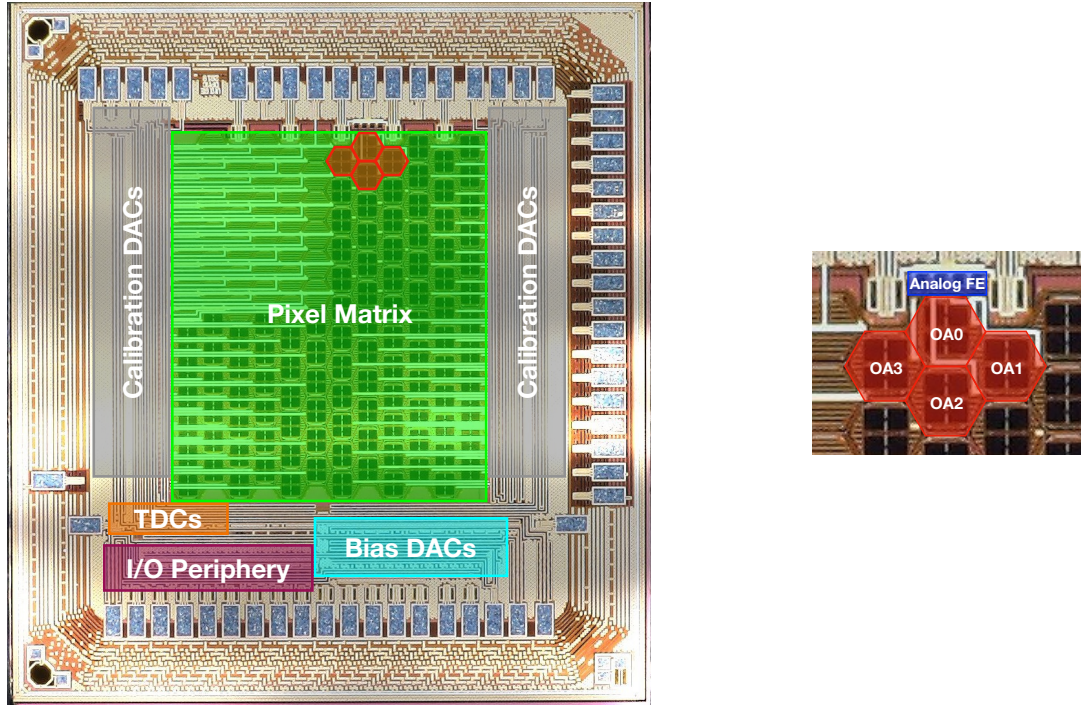


FIGURE 1.6: (Left) Microscope picture of the monolithic silicon pixel ASIC of the PicoAD proof-of-concept prototype. The main building blocks are underlined in color. (Right) Magnified view of a detail of the ASIC, showing the four pixels with analog readout.

submatrices with different amplifier and discriminator designs, each containing 6×6 pixels. Each pixel is a regular hexagon with a side length of $65 \mu\text{m}$. Hence, the pitch is comparable to that of squared pixels, with a $100 \mu\text{m}$ pixel pitch.

Four of the 144 pixels are connected to charge-sensitive amplifiers and analog drivers, which measure precisely the charge produced by the sensor. These four pixels with analog readout are characterized in the following sections. The remaining pixels are connected to a digital readout with various discrimination stages. They were also tested, but since the thesis focuses on optimizing the sensor and analog front-end electronics, the results with the digital readout will not be discussed. The four analog output pixels are hereafter referred to as OA0, OA1, OA2, and OA3. The pixel matrix of the PicoAD proof-of-concept prototype is identical to that of a prototype submitted in 2019 and discussed in [58].

Figure 1.7 shows the schematic of the analog front-end electronics connected to four pixels of the matrix of the PicoAD proof-of-concept prototype.

The thickness of the drift region should be larger than the inter-pixel distance, which is $10 \mu\text{m}$ in the present prototype. This condition was not compatible with a complete depletion of the sensor using the epitaxial process, which had a minimum boron concentration of $3 \times 10^{14} \text{ cm}^{-3}$ available at the time of manufacturing at IHP. For this reason, the prototype wafer produced had a $5 \mu\text{m}$ thick epitaxial layer under the gain layer (the absorption region) and a $10 \mu\text{m}$ thick epitaxial layer above the gain layer (the drift region), both with a resistivity of $50 \Omega \text{ cm}$. The thickness of the drift region of this PicoAD proof-of-concept prototype is thus sub-optimal. It may reduce avalanche gain and introduce distortions of the weighting field, which is responsible for signal induction in the inter-pixel region.

Four implantation doses were used to form the gain layer, with dose one being

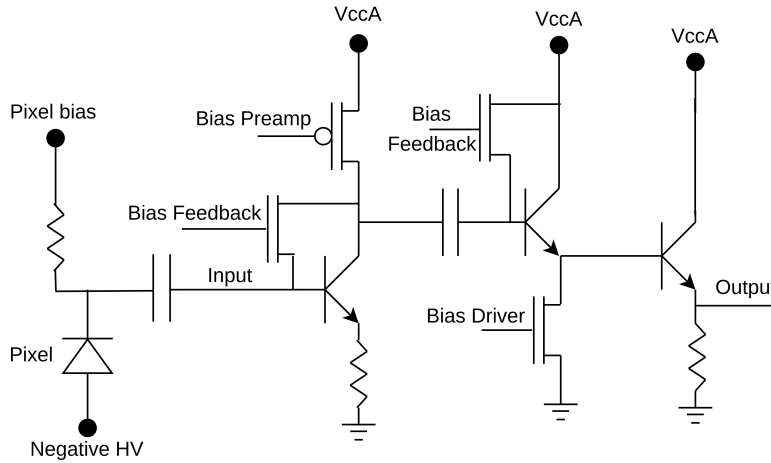


FIGURE 1.7: Schematic view of the analog front-end electronics and driver configuration for the four analog pixels of the PicoAD proof-of-concept prototype.

the lowest and dose four being the highest, targeting a sensor working point at a bias voltage between 100 V and 200 V. In all cases, the dose of the *p*-type gain layer implant was set at 60 % of the dose of the *n*-type gain layer implant to ensure a reduction in the electric field in the absorption region when complete depletion is achieved.

1.2.6 MONOLITH Prototype 2 without a Gain Layer

The second MONOLITH prototype without a gain layer was also produced by the IHP Microelectronics foundry using the IHP SG13G2 130 nm BiCMOS technology. It was submitted to the foundry in 2021 and was received in 2022. The bulk of the sensor doesn't feature the PicoAD sensor concept. Hence, the wording *without Gain Layer* was added to the name.

Figure 1.8 shows the monolithic silicon pixel ASIC layout of the second MONOLITH prototype without a gain layer. The four analog pixels are highlighted in red.

The ASIC is slightly smaller than the PicoAD proof-of-concept one, with a total size of 2.0 mm × 2.5 mm. The main features of the ASIC are identical to those of the PicoAD Proof of Concept. In summary: a pixel matrix of 144 hexagonal pixels with a side length of 65 μm, a periphery outside the matrix with the slow-control interface system, and four pixels, called OA0, OA1, OA2, and OA3, connected to analog front-end electronics.

The front-end electronics of the analog pixels differ slightly from those in the previous PicoAD proof-of-concept prototype, with some design improvements. Figure 1.9 shows the schematic of the analog front-end electronics connected to four pixels of the matrix of the second MONOLITH prototype without a gain layer.

The improvements include separating the power supply of the preamplifier from that of the driver stage to remove the feedback path that caused crosstalk between the channels and instability in the PicoAD proof-of-concept prototype. In addition, the driver stage was equipped with a semi-differential analog output, which reduces common-mode noise outside the ASIC and provides more stable amplifier operation at high gain in the front-end electronics. The new driver also reduced the signal rise time, significantly improving the detector timing jitter.

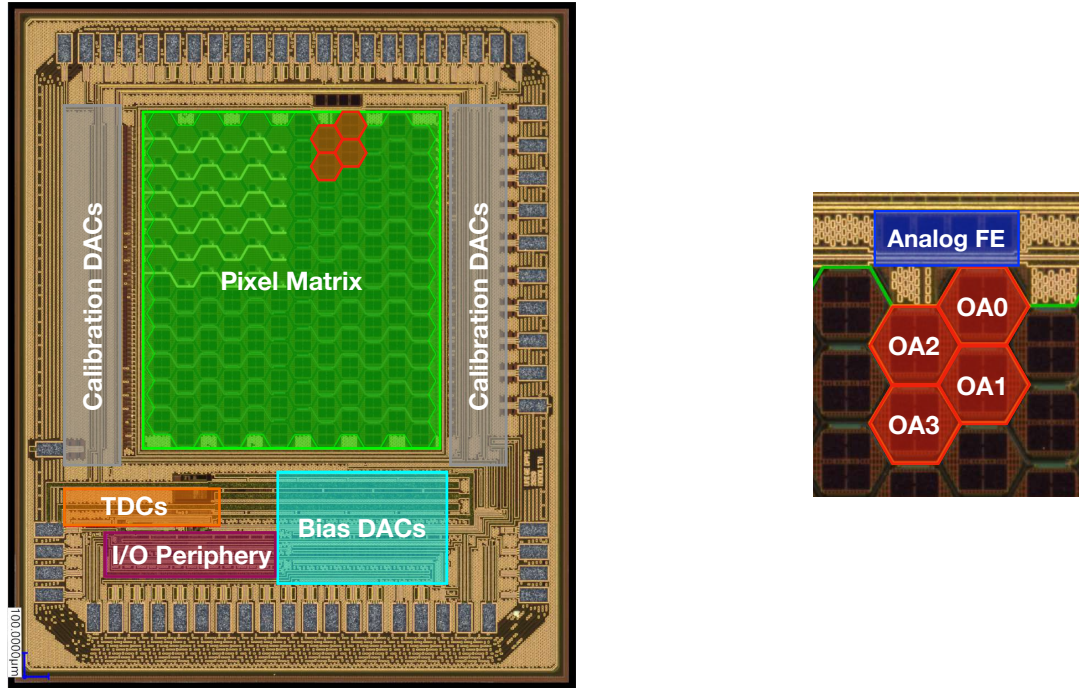


FIGURE 1.8: (Left) Microscope picture of the monolithic silicon pixel ASIC of the second MONOLITH prototype without a gain layer. The main building blocks are underlined in colors. (Right) Magnified view of a detail of the ASIC, showing the four pixels with analog readout.

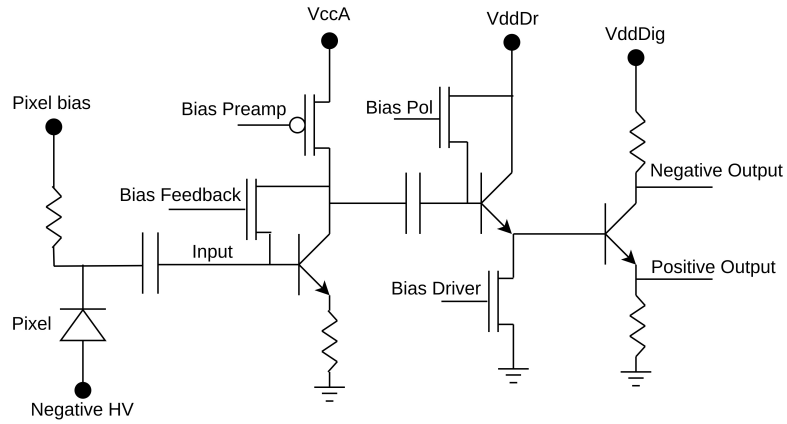


FIGURE 1.9: Schematic view of the analog front-end electronics and driver configuration for the four analog pixels of the second MONOLITH prototype without a gain layer.

On the sensor side, the second MONOLITH prototype without a gain layer was manufactured on a standard boron-doped substrate with a very low resistivity of $0.1 \, \Omega \, \text{cm}$ (the same substrate as the previous prototype). Then, a $50 \, \mu\text{m}$ -thick epilayer with $350 \, \Omega \, \text{cm}$ resistivity was grown on top of the substrate. The CMOS processing, with the layout shown in Figure 1.8, was manufactured on top of it.

1.2.7 MONOLITH Prototype 2 with a Gain Layer

The second MONOLITH prototype with a gain layer was also produced by the IHP Microelectronics foundry using the IHP SG13G2 130 nm BiCMOS technology. It was submitted to the foundry in 2023 and was received in 2024. This prototype features the same ASIC design as the second MONOLITH prototype without a gain layer. Moreover, the PicoAD sensor concept was used to produce a gain layer for the sensor. Hence, the wording *with a Gain Layer* was added to the name.

The same CMOS masks used for the second MONOLITH prototype without a gain layer were also used for this prototype, leading to an ASIC layout visually identical to the one already shown in Figure 1.8.

The difference between this prototype and the previous one is exclusively in the active region of the sensor's bulk. The ASIC was manufactured on a standard boron-doped substrate with a very low resistivity of $0.1 \Omega \text{ cm}$ (the same substrate as the previous prototype). A first epitaxial layer (which can be thick either 3 or 5 μm) was grown on the substrate, then the gain layer was implanted, and then a second epitaxial layer (which can be thick either 15 or 25 μm -thick) was grown on top. Both epitaxial layers have a resistivity of $350 \Omega \text{ cm}$, compared to the $50 \Omega \text{ cm}$ resistivity of the epitaxial layers in the PicoAD proof-of-concept prototype. The combination of 2×2 epitaxial-layer thicknesses yields four possible combinations. Moreover, the gain layer was implanted with different doses at each thickness (in a variable number, depending on the thickness, and up to 6 per thickness), resulting in 15 different flavors. These various flavors of PicoAD were created to explore the time resolution and signal-to-noise ratio for other combinations, rather than the one used in the PicoAD proof-of-concept prototype.

Figure 1.10 shows a simplified drawing of the active region of the second MONOLITH prototype with a gain layer: the two epitaxial layers and the gain layer are visible. The pixel implants, realized with a CMOS process, are also visible on top of the sensor.

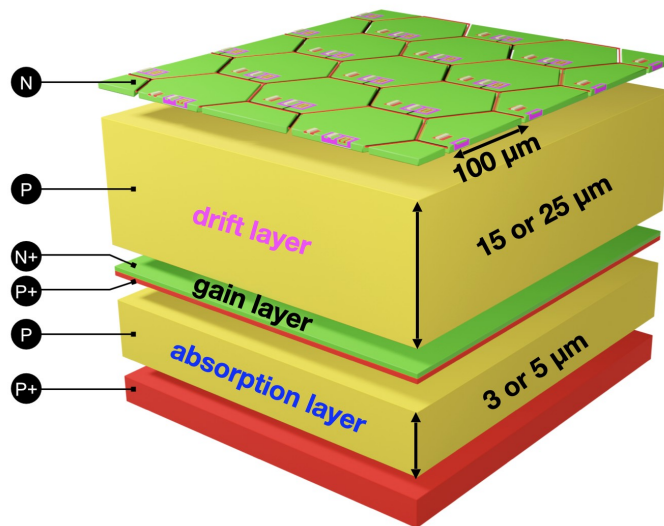


FIGURE 1.10: Simplified drawing of the active region of the second MONOLITH prototype with a gain layer. Starting from the bottom, it is visible: the first epitaxial layer (absorption layer), which can be 3 or 5 μm thick; the gain layer, which can have up to 6 different doping doses, the second epitaxial layer (drift layer), which can be 15 or 25 μm thick, and pixel implants made with a CMOS process.

1.2.8 The Printed Circuit Board

It is crucial to say that each prototype is interfaced with the experimental setup through some *Printed Circuit Board* (PCB), designed by the same research group at the University of Geneva. Identical PCBs were used to interface the second MONOLITH prototypes with and without the gain layer, since they have the same ASIC design. A different PCB is used to host the PicoAD proof-of-concept prototype.

Figure 1.11 shows a PCB hosting an ASIC of the second MONOLITH prototype without a gain layer.

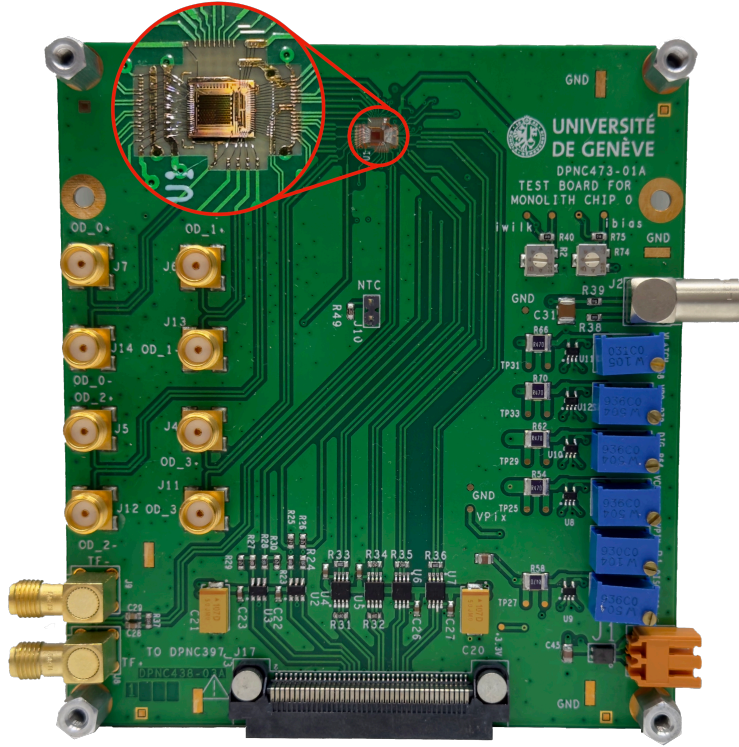


FIGURE 1.11: Printed Circuit Board (PCB) hosting an ASIC of the second MONOLITH prototype without a gain layer. The insert shows a zoom of the ASIC, where the structure of the matrix and the wire bonding scheme that interfaces with the ASIC are visible.

The PCB interfaces to the ASIC via a series of aluminum wire bonds, which deliver all bias voltages and carry both the analog and digital signals (see insert of Figure 1.11). Figure 1.11 highlights the following key components:

- **Gold SMA connectors (left):** feed the fast oscilloscope with analog signals.
- **Orange low-voltage connector (bottom right):** supplies 3 V from the low-voltage power supply module.
- **Blue trim potentiometers (right):** adjust the various ASIC bias voltages.
- **Metallic High-Voltage (HV) connector (right):** connects to the sensor bias from the HV power supply module.
- **Black GPIO connector (bottom):** provides digital control and communication.

1.2.9 Other Projects with SiGe BiCMOS Technology

The research group working with monolithic silicon pixel detectors at the Département de Physique Nucléaire et Corpusculaire at the University of Geneva is not only involved in the MONOLITH project but also in other projects that exploit the many advantages of the same SiGe BiCMOS technology. Two notable examples are the *100 μ PET Project* and the *FASER Preshower Upgrade*, which were carried out in parallel with the MONOLITH project.

The 100 μ PET project is a Swiss collaboration led by the University of Geneva, EPFL, and the University of Lucerne, aiming to develop a small-animal positron emission tomography (PET) scanner with unprecedented spatial resolution, down to a few hundred micrometers [76]. The scanner employs stacks of monolithic silicon pixel sensors interleaved with thin tungsten converter foils. When a positron annihilates, the resulting 511 keV photons convert in the tungsten, producing electrons that are tracked by 270 μ m-thick silicon sensors for precise depth-of-interaction measurements. Each of the four scanner towers contains 60 such lead-and-silicon layers, yielding a simulated spatial resolution of approximately 0.2 mm full width at half maximum (FWHM) and a sensitivity of approximately 3 to 4 % for mouse-scale objects. Monte Carlo studies using Geant4 have addressed the design, demonstrating excellent performance in reconstructing point sources in preclinical models. Beyond instrumentation, the project integrates advanced image-reconstruction algorithms to fully exploit the high granularity of the detector and correct for photon scatter and attenuation. Scheduled for completion in 2025 under an SNSF Sinergia grant, 100 μ PET's first biomedical application will probe mouse atherosclerotic plaque development, monitoring metabolic changes under different diets and treatments.

The Forward Search Experiment (FASER) at the LHC is installing a new tungsten–silicon preshower detector upstream of its calorimeter to enhance sensitivity to axion-like particles that decay into closely spaced TeV-energy photons [77]. Planes instrumented with monolithic silicon pixel detectors, each with a granularity of 100 μ m, are interleaved between tungsten plates to convert photons and initiate electromagnetic showers. The monolithic pixel ASICs, designed in the same 130 nm SiGe BiCMOS process by IHP Microelectronics that is used for the MONOLITH project, feature hexagonal pixels (65 μ m side) arranged in 13 supercolumns, each with its readout logic and flash ADCs, providing a dynamic range from 0.5 fC to 64 fC. The group at the University of Geneva has led the development of ASICs, prototyping three variants in 2021 and conducting beam tests to validate timing (less than 200 ps) and charge measurement performance under realistic backgrounds. Allpix² simulations predict more than 95% two-photon separation efficiency for 1 TeV photons at 200 μ m spacing, a key requirement to distinguish axion signals from single-photon backgrounds. Funded in part by an SNSF grant to Geneva, module assembly, surface commissioning, and installation in FASER's tunnel were completed by the end of 2024, and the first physics data are expected in Run 3 of the LHC. This upgrade significantly broadens FASER's discovery potential for long-lived, weakly interacting particles, leveraging cutting-edge monolithic-pixel technology pioneered at the University of Geneva.

Chapter 2

Silicon-Based Devices

This chapter introduces the foundational principles of semiconductors and the working mechanisms of the devices at the heart of modern experimental particle physics, particularly silicon detectors. Understanding the physics of semiconductors is crucial for appreciating how these materials are utilized to detect and process signals generated by charged particles.

We begin with an overview of the physics of semiconductors. Section 2.1 discusses the origin and structure of energy bands in crystalline solids, which give rise to the unique electrical properties of semiconductors. We then examine the concentration of charge carriers, followed by how introducing impurities, also called dopants, can alter the energy bands. The concepts of mobility, drift velocity, and resistivity are introduced to describe how carriers move under the influence of electric fields. Finally, we conclude the section by discussing avalanche multiplication through impact ionization. This phenomenon is exploited in sensors with a gain layer to enhance the primary charge, as seen in the PicoAD sensor concept.

In Section 2.2, we move to the fundamental building block of most semiconductor devices: the p - n junction. We explore its electrostatics, the built-in potential, and the depletion region. The electrical characteristics of the p - n junction under reverse bias are described, which make the p - n junction a suitable technology for particle detection. The section describes the p - n junctions in the MONOLITH project detectors.

Section 2.3 introduces semiconductor transistors and their relevance for integrated electronics in particle detectors. We present both bipolar junction transistors (BJTs) and their advanced implementations, heterojunction bipolar transistors (HBTs), which offer enhanced performance at high frequencies and low temperatures. The chapter then describes metal-oxide-semiconductor field-effect transistors (MOSFETs), the cornerstone of modern CMOS technology. We conclude with high-voltage CMOS (HV-CMOS) devices. This key technology enables the monolithic integration of sensors and readout electronics, preserving the fast collection of the charge for precise detection and radiation hardness.

This chapter thus sets the stage for understanding the design and functionality of silicon-based detectors by grounding the reader in the fundamental physics and device technologies that enable precise particle tracking and energy measurements.

The concepts presented in this chapter, along with some of the figures shown, are primarily based on three books, listed in chronological order: [78], [79], and [80].

2.1 Physics of Semiconductors

Solid-state materials can be divided into three classes according to their behavior when conducting electricity, as defined by measuring the conductivity σ (or the resistivity $\rho = \sigma^{-1}$) of the material: metals (or conductors), semiconductors, and insulators.

Since the early nineteenth century, silicon (Si) and germanium (Ge) have been the most extensively studied semiconductors composed of single atoms. Starting in the early 1950s, silicon gradually supplanted germanium, mainly due to economic considerations (silicon makes up almost 28 % of the Earth's crust) and also because of its better properties at room temperature. Additionally, high-quality silicon dioxide can be grown thermally. That's why silicon is mainly considered in the rest of the chapter.

2.1.1 Energy Bands

When considering a single atom, the electrons in the atom's shells have discrete energies, which can be calculated, for example, with Bohr's model. The energy levels depend on the interaction between the electrons and the atom's nucleus. In the case of silicon, two energy levels are possible for the outer shell electrons.

When bringing together N atoms of silicon to form a crystal, the electrons of each single atom also start to feel the interaction with all the other $N - 1$ atoms. Instead of two energy levels, N separate but closely spaced levels are formed. When N is large, the result is an essentially continuous energy band. Figure 2.1 shows N silicon atoms being brought together to form a lattice crystal. The lattice spacing is proportional to the distance between two atoms, and $5.43 \text{ \AA}$ (or $5.43 \times 10^{-10} \text{ m}$) is the real lattice spacing in a crystal of silicon.

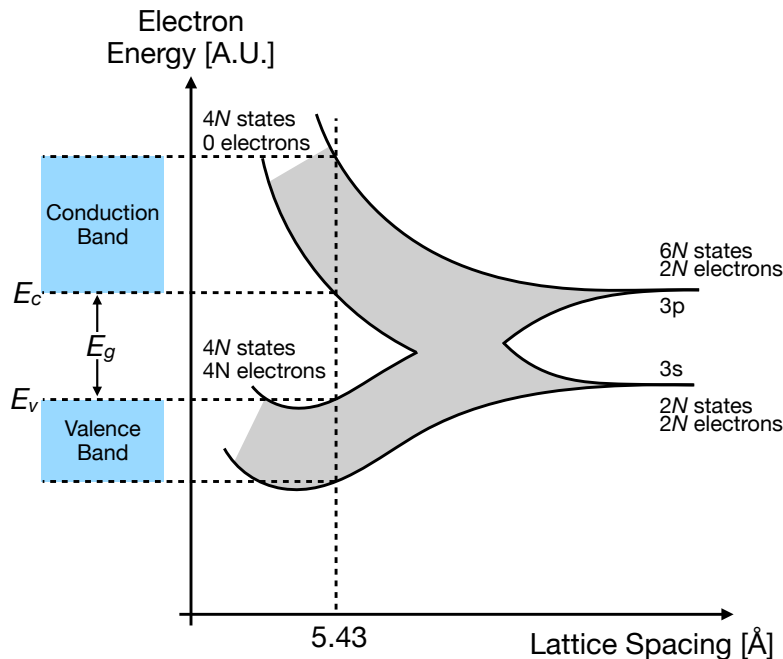


FIGURE 2.1: Energy levels in silicon as a function of the lattice spacing. When N isolated atoms are brought together to form a lattice crystal, the electron energies form a conduction and a valence band. The valence band is filled at a temperature of 0 K, and the conduction band is empty. Adapted from [78].

At a temperature of absolute zero, electrons occupy the lowest energy states so that all states in the lower band (the valence band) will be full, and all states in the upper band (the conduction band) will be empty. The bottom energy level of the conduction band is called E_C , the top energy level of the valence band is called E_V , and the bandgap energy between the bottom of the conduction band and the top of the valence band is called $E_g = E_C - E_V$. When the temperature rises, electrons in the valence band can acquire enough energy to break the bond in the silicon crystal and jump into the conduction band, leaving a hole in the valence band. For intrinsic semiconductors at thermal equilibrium, the concentration of electrons in the conduction band is always equal to the concentration of holes in the valence band, as a hole is only created when an electron leaves the valence band.

The energy band structure is related to the subdivision of materials into three classes based on their conductivity: metals, semiconductors, and insulators. We can summarize the main characteristics as follows:

- metals (conductors) have a very low resistivity value, and can either have a conduction band partially filled with electrons or an overlap with the valence and conduction bands. In both cases, electrons can move in the crystal with only a small electric field.
- Semiconductors have the valence and conduction bands separated by an energy gap E_g , which is small. Thermal excitations can easily cause an electron to transition to the conduction band, allowing it to conduct current.
- Insulators have the valence and conduction bands separated by an energy gap E_g greater than that of semiconductors. Even when the temperature is increased, it is difficult for an electron to go to the conduction band; hence, the crystal does not conduct current.

Figure 2.2 shows the typical band structures for the three groups.

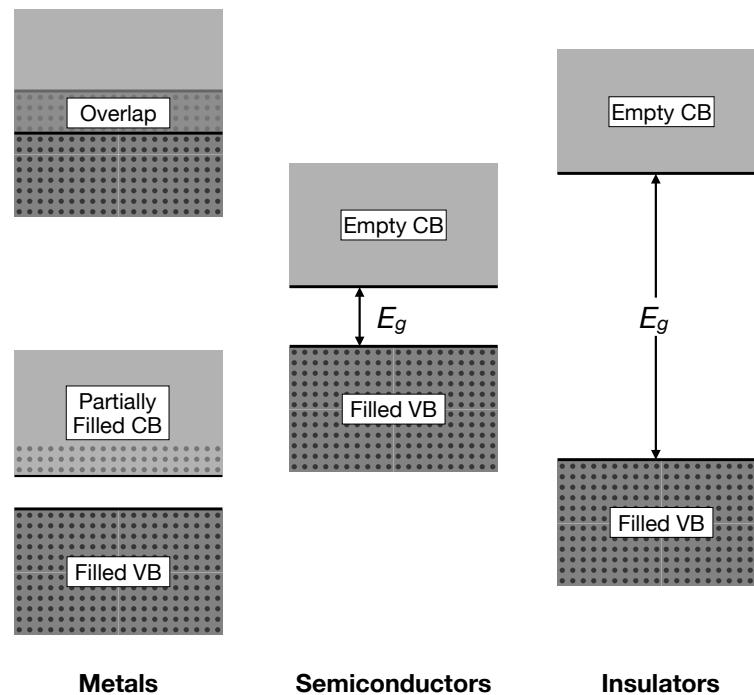


FIGURE 2.2: Typical conduction and valence band structures for metals (left), semiconductors (center), and insulators (right) at 0 K. Adapted from [79].

2.1.2 Charge Carrier Concentration

Electrons are fermions and obey the *Fermi-Dirac* statistics in solid crystals. Starting from the Pauli exclusion principle and by considering the electrons in the solid as indistinguishable particles, statistical arguments lead to the *Fermi-Dirac distribution function* $f(E)$, which gives the probability that an electron occupies an available state with energy E at equilibrium temperature T :

$$f(E) = \frac{1}{e^{\frac{E-E_F}{k_B T}} + 1}, \quad (2.1)$$

where the quantity E_F is called the *Fermi level* and represents the energy for which the occupation probability is 0.5, while k_B is the Boltzmann constant. Figure 2.3 shows the Fermi-Dirac distribution function.

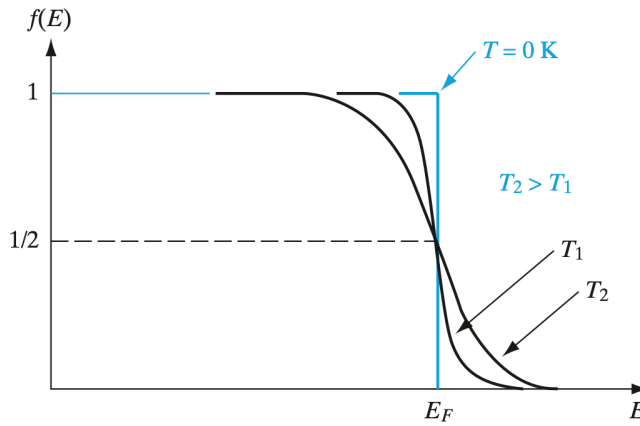


FIGURE 2.3: Fermi-Dirac distribution function for different temperatures: in light blue $T = 0$ K, in black T_1 and T_2 , with $T_2 > T_1$. Taken from [79].

When the temperature is absolute zero, the distribution is a step function, which implies that, at 0 K, all the available energy states below E_F are filled with electrons, and all the energy states above E_F are empty. When the temperature increases, the distribution gets smoother, and some energy states above E_F start to be filled with electrons with probability $f(E)$.

We can use the Fermi-Dirac distribution function to calculate the concentrations of electrons and holes in a semiconductor. For example, free electrons in intrinsic silicon are present in the conduction band, hence with energies $E_C < E < \infty$. Knowing the density of states $N(E)$ (in cm^{-3}) for the energy range dE , the concentration of electrons n in the conduction band and at equilibrium is given by:

$$n = \int_{E_C}^{\infty} f(E) N(E) dE. \quad (2.2)$$

With the approximation that the Fermi level lies at least several $k_B T$ below the conduction band, the Fermi-Dirac distribution can be approximated, and the concentration of electrons in the conduction band can be calculated, giving:

$$n = N_C e^{-\frac{E_C - E_F}{k_B T}}, \text{ with } N_C = 2 \left(\frac{2 \pi m_n^* k_B T}{h^2} \right)^{3/2}, \quad (2.3)$$

where the quantity m_n^* is called the density-of-states effective mass for electrons, while h is the Planck constant.

Similarly, the concentration of holes in the valence band is calculated, knowing that holes can only be present in the valence band, hence with $-\infty < E < E_V$, giving:

$$p = N_V e^{-\frac{E_F - E_V}{k_B T}}, \text{ with } N_V = 2 \left(\frac{2 \pi m_p^* k_B T}{h^2} \right)^{3/2}, \quad (2.4)$$

where the quantity m_p^* is called the density-of-states effective mass for holes.

The expressions found are valid for an intrinsic semiconductor (without impurities) and a doped semiconductor, given that the thermal equilibrium condition is satisfied. The concentration of electrons in the conduction band n_i is equal to the concentration of holes in the valence band p_i in intrinsic semiconductors at thermal equilibrium, which translates to $n_i = p_i$. This equation can be solved to find the Fermi level for an intrinsic semiconductor E_{Fi} :

$$n_i = p_i \Rightarrow E_{Fi} = \frac{E_V + E_C}{2} + \frac{k_B T}{2} \ln \left(\frac{N_V}{N_C} \right). \quad (2.5)$$

At room temperature, the second term is much smaller than the bandgap; hence, the Fermi level for an intrinsic semiconductor lies very close to the midpoint of the energy gap.

2.1.3 Intrinsic Semiconductor and Doping

In addition to the intrinsic carriers generated thermally, charge carriers can be created in a semiconductor by introducing impurities into the crystal. This process, called doping, can alter a crystal to predominate either electrons (*n*-type doping) or holes (*p*-type doping). When impurities are introduced into an intrinsic crystal, additional levels are created in the energy band structure, usually within the band gap.

Atoms from the third column of the periodic table (like B, Al, Ga, and In) introduce energy levels near the valence band in Si. These types of impurity levels are called *acceptor levels*, and they have a hole concentration p much greater than the conduction band electron concentration n . This type of material is called *p*-type material.

On the contrary, atoms from the fifth column of the periodic table (like P, As, and Sb) introduce energy levels near the conduction band in Si. These levels are called *donor levels*, with an electron concentration n much greater than the valence band hole concentration p . This type of material is called *n*-type material.

Figure 2.4 shows an example of doping with acceptor (B) and donor (As) atoms in the covalent bonding model of a silicon crystal.

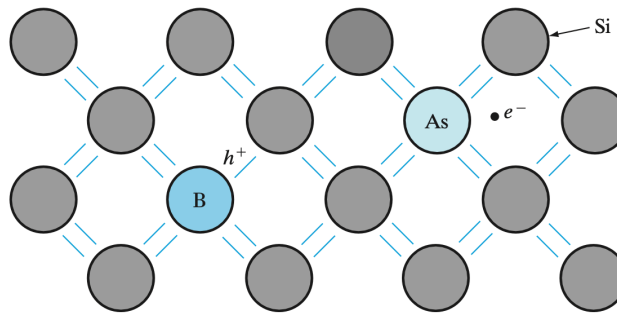


FIGURE 2.4: Doping with acceptor (B) and donor (As) atoms in the covalent bonding model of a silicon crystal. Taken from [79].

The introduction of impurities changes the Fermi level in the energy band diagram. Applying the equations 2.3 and 2.4, the concentrations of electrons and holes can be calculated. Figure 2.5 shows a scheme for intrinsic and doped semiconductors, where it is clear that *n*-type doping creates a prevalence of electrons in the conduction band. In contrast, *p*-type doping creates a prevalence of holes in the valence band.

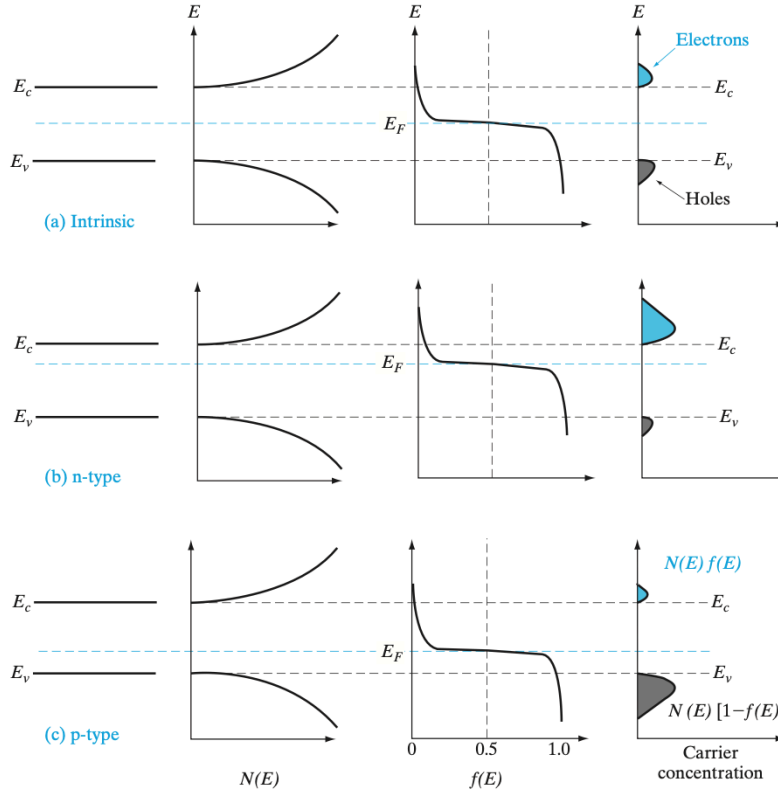


FIGURE 2.5: Schematic band diagram, density of states, Fermi–Dirac distribution, and the carrier concentrations for (a) intrinsic, (b) *n*-type, and (c) *p*-type semiconductors at thermal equilibrium. Notice how *n*-type doping creates a prevalence of electrons in the conduction band, while *p*-type doping creates a prevalence of holes in the valence band. Taken from [79].

The product of the concentration *n* and *p* is a constant for a particular material and temperature, even if the doping is varied. Assuming that *n_i* is the concentration of electrons in the conduction band for intrinsic semiconductors (equal to the concentration of holes in the valence band *p_i*). It holds:

$$n_i p_i = (N_C e^{-\frac{E_C - E_{Fi}}{k_B T}})(N_V e^{-\frac{E_{Fi} - E_V}{k_B T}}) = N_C N_V e^{-\frac{E_g}{k_B T}} = n p. \quad (2.6)$$

Apart from showing that the product of the concentration *n* and *p* is a constant, Equation 2.6 ($n p = n_i p_i = n_i^2$), usually referred to as the *Mass Law Equation*, suggests a convenient way to write the concentration of electrons in the conduction band and holes in the valence band, which will be used in the following chapters. These expressions are:

$$n = n_i e^{\frac{E_F - E_{Fi}}{k_B T}}, \quad p = n_i e^{\frac{E_{Fi} - E_F}{k_B T}}. \quad (2.7)$$

Using Equation 2.6, together with the expression for *N_C*, and *N_V* (from Equation 2.3 and Equation 2.4), we can calculate for intrinsic silicon at room temperature that $n_i \simeq 1.2 \times 10^{10} \text{ cm}^{-3}$.

2.1.4 Mobility, Drift Velocity, and Resistivity

Understanding the mechanisms that regulate the movement of charge carriers in a semiconductor is essential for calculating the current flow when an electric field is present. Even without an electric field, the charge carriers in a solid are constantly in motion due to their thermal energy. Considering an n -type semiconductor sample with uniform donor concentration and at thermal equilibrium, the conduction electrons in the conduction band are essentially free particles, with the lattice influence incorporated in a so-called conductivity effective mass for electrons m_n (different from the density-of-states effective mass m_n^* of Equation 2.3). The Boltzmann theorem for equipartition of energy holds:

$$\frac{1}{2}m_n v_{th}^2 = \frac{3}{2}k_B T, \quad (2.8)$$

where v_{th} is the average thermal velocity of electrons. At a temperature $T = 300$ K, the thermal velocity is of the order of 10^7 cm/s (or, in a more suitable unit of measurement, $100 \mu\text{m}/\text{ns}$). It's important to notice that when describing the silicon detectors in the rest of the thesis, more suitable units of measurement are μm for distances and ns or ps for time, hence the unusual decision of expressing the velocity of charge carriers in $\mu\text{m}/\text{ns}$ and the electric fields in $\text{V}/\mu\text{m}$. The electrons undergo a succession of random scattering with lattice atoms, with a mean free path of $\lambda_c \sim 0.1 \mu\text{m}$ and a mean free time between collisions $\tau_c \sim 1$ ps.

The movement created by thermal agitation is entirely random, and there is no net flow of electrons in a specific direction. On the other hand, when an external electric field $\mathcal{E}$ is applied to the doped semiconductor, there is a net flow of electrons in the opposite direction to the electric field (because electrons are negatively charged). The velocity the electrons gain from the electric field is called the drift velocity v_e . All the momentum gained by a single electron between two consecutive collisions, which is $-q\mathcal{E}\tau_c$, is lost to the lattice collision. Hence, we can equate the momentum gain to $m_n v_e$ and get an expression for the drift velocity for electrons:

$$v_e = -\frac{q\tau_c}{m_n}\mathcal{E} \Rightarrow v_e = -\mu_n\mathcal{E}, \quad (2.9)$$

having defined the mobility for electron $\mu_n \equiv \frac{q\tau_c}{m_n}$.

The current density J_n resulting from the net drift of electrons is just the number of electrons crossing a unit area per unit time (nv_e) multiplied by the charge of the electron ($-q$):

$$J_n = -qn v_e. \quad (2.10)$$

Together with the previous Equation 2.9, we have:

$$J_n = qn\mu_n\mathcal{E} \Rightarrow J_n = \sigma\mathcal{E}, \quad (2.11)$$

having defined the conductivity for electron $\sigma \equiv qn\mu_n$. We've already found the conductivity σ to classify if a solid behaves like a metal, a semiconductor, or an insulator. For an n -doped semiconductor, the conductivity σ is directly proportional to the density of electrons in the conduction band n , which implies that the behavior of the solid under an external electric field can be altered by adding impurities, also known as doping. Equation 2.11 is called the microscopic Ohm's law. We can define

the resistivity of the material as:

$$\rho = \frac{1}{\sigma} = \frac{1}{qn\mu_n}. \quad (2.12)$$

The electron mobility μ_n is tabulated with great precision for many common semiconductors and slightly depends on doping. The semiconductor resistivity ρ and the density of electrons n are inversely proportional. Resistivity is a crucial parameter when discussing silicon detectors based on a p - n junction, as it determines several functional properties, as we will see in the following sections.

We took an n -doped semiconductor and investigated the velocity properties of the intrinsic charge carriers, specifically electrons in the conduction band. Of course, everything is valid also in the case of a p -doped material, whose charge carriers are holes in the valence band, or an intrinsic material, whose charge carriers are both electrons in the conduction band and holes in the valence band. In this more general case, Equation 2.11 becomes:

$$J_n = q(n\mu_n + p\mu_p)\mathcal{E}. \quad (2.13)$$

p is the density of holes in the valence band, and μ_p is their mobility. The conductivity σ and the resistivity ρ modify accordingly.

The proportionality between the electron drift velocity v_e (or hole drift velocity v_h) and the electric field $\mathcal{E}$ is valid only for velocities smaller than the thermal velocities, because it still holds the assumption that the mean free time between collisions τ_c is independent of the electric field. The drift velocity approaches a saturation value for high electric fields, and increasing the electric field does not correspond to an increase in the drift velocity. The resulting relation between electron drift velocity v_e and electric field E can be modeled empirically, for example using the so-called *Canali model* [81]:

$$|v_e| = \frac{\mu_e \mathcal{E}}{[1 + (\mu_e \mathcal{E} / v_{\text{sat},e})^{\beta_e}]^{1/\beta_e}}, \quad (2.14)$$

where μ_e is the low-field mobility, $v_{\text{sat},e}$ is the electron saturation velocity, and β_e is a parameter found empirically. In the limit for low electric field ($\mathcal{E} \ll v_{\text{sat},e} / \mu_e$), the relation from Equation 2.9 holds, while, in the limit for high electric field ($\mathcal{E} \gg v_{\text{sat},e} / \mu_e$), the electron drift velocity tends to its saturation value. The same relation also holds when the charge carriers are holes. In the following chapters, especially when discussing simulations, the drift velocity will be a key parameter to consider, as it determines the shape of the induced current. The values used for the parametrization at room temperature of $T = 300$ K are found in Table 2.1, which are the same as those used from the default models in Sentaurus Device [82], and are taken from [83].

Carrier Type i	μ_i [$\mu\text{m}^2 \text{V}^{-1} \text{ns}^{-1}$]	$v_{\text{sat},i}$ [$\mu\text{m ns}^{-1}$]	β_i
Electrons	141.7	107.0	1.109
Holes	47.1	83.7	1.213

TABLE 2.1: Parameters used to calculate the drift velocity in silicon from the electric field according to Equation 2.14 (*Canali model*). These values are used as the default model in Sentaurus Device [82] and are taken from [83].

The values of Table 2.1 are used to produce the plot of drift velocity for electrons and holes as a function of the electric field shown in Figure 2.6.

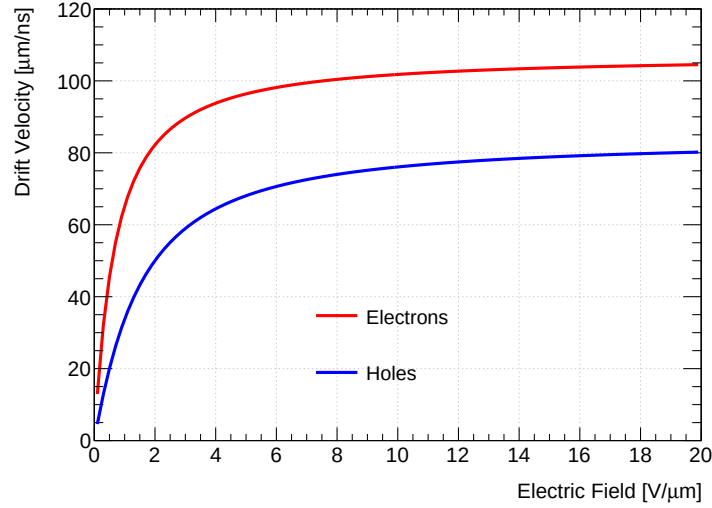


FIGURE 2.6: Drift velocity in silicon for electrons (in red) and holes (in blue) as a function of the electric field, using the Equation 2.14 with the values of Table 2.1.

Figure 2.6 shows that, already at an electric field of $5 \text{ V}/\mu\text{m}$, the drift velocity is within 10 % of the saturation value. To improve the time resolution of silicon detectors (as explained in Section 3.5), it is fundamental to ensure that the drift velocity is close to the saturation value, which translates to a condition on the electric field: $\mathcal{E} > 5 \text{ V}/\mu\text{m}$.

2.1.5 Avalanche Due to Impact Ionization

An interesting phenomenon occurs when the electric field exceeds a certain threshold: the impact ionization. The electric field accelerates the charge carrier between two consecutive collisions and can acquire enough energy to produce an ionization of the lattice atom at the moment of the scattering. The ionization creates an electron and a hole, which are also accelerated by the electric field and can form other electron-hole pairs. This process can be virtually infinite, limited only by the constraints of the circuit.

Let's consider again a piece of n -doped silicon. When the electric field is high, each electron drifting towards the low-potential terminal has a probability P of producing an electron-hole pair during its path. Given n_{in} initial electrons, these will produce $n_1 = Pn_{\text{in}}$ secondary electrons after Pn_{in} ionization events. These secondary electrons will produce $n_2 = Pn_1 = P^2n_{\text{in}}$ tertiary electrons and so on. We can define the gain M as the ratio of the final n_{out} over the initial n_{in} electrons:

$$M = \frac{n_{\text{out}}}{n_{\text{in}}} = \frac{n_{\text{in}} + n_1 + n_2 + \dots}{n_{\text{in}}} = 1 + P + P^2 + \dots = \frac{1}{1 - P}. \quad (2.15)$$

The probability P depends on the peak and the shape of the electric field inside the semiconductor. The first models expressed the ionization rate α to which P depends [84]. It's clear that for P approaching unity, the gain M approaches infinity, meaning an uncontrollable avalanche, for example, the case of a junction breakdown. However, the impact ionization avalanche can also be used to produce a beneficial effect, for instance, in the case of the Picosecond Avalanche Detector (see Chapters 4 and 6), where a smartly engineered electric field produces a shallow gain of 20 to 100 in a narrow region. This gain multiplication aims to increase the number of charge carriers after an ionizing particle passes, thereby enhancing the primary signal.

2.2 The p - n Junction

We now consider a situation involving a single-crystal semiconductor containing both p - and n -doped regions to form a p - n junction. The relevance of the p - n junction is twofold, on one hand, an inversely biased p - n junction serves as the particle detector, since an incoming ionizing radiation creates charge carriers inside the silicon, and, on the other hand, the p - n junction is the key building block to build both *bipolar junction transistors* (see *BJT* in Section 2.3.1) and the *metal-oxide-semiconductor field effect transistors* (see *MOSFET* in Section 2.3.3).

2.2.1 Electrostatic of the p - n Junction

The simplest case to treat to get insight into a p - n junction is when we consider a *step junction*, meaning a region with a uniform p doping, then a sharp edge, and a uniform n doping on the other side. This model accurately represents epitaxial junctions, particularly when dopant diffusion during the fabrication process is minimal. However, some corrections are required when treating graded junctions. Also, we now consider a junction at equilibrium, i.e., with no external bias, and at room temperature $T = 300$ K.

Figure 2.7 shows a way (not a practical one) to make a p - n junction. We start with two blocks of semiconductor, the first one being n -doped with a density of donor atoms N_D , the second one being p -doped with a density of acceptor atoms N_A . Each block has its own Fermi level E_F , which lies in a different position between the edges of the valence band E_V and the conduction band E_C . Free charge carriers diffuse when the two blocks are connected due to the concentration difference. In particular:

- some free electrons in the conduction band of the n -type silicon will diffuse into the p -type silicon, leaving uncompensated positive donor ions N_D^+ ;
- some free holes in the valence band of the p -type silicon will diffuse into the n -type silicon, leaving uncompensated negative acceptor ions N_A^- .

The diffusion makes the Fermi level uniform between the two blocks, and the junction reaches equilibrium. It follows that there will be a space region with uncompensated charges, usually referred to as *Space Charge Region*, with width W that will produce an electric field $\mathcal{E}$ directed from the n -type to the p -type. Figure 2.7 also shows the built-in electrostatic potential V_0 produced from the junction.

The electric field $\mathcal{E}$ creates a movement of charge carriers via drift, and equilibrium is reached when the net current of drift and diffusion is zero for each type of carrier. The built-in electrostatic potential, $V_0 = V_n - V_p$, can be calculated using this intuitive property for a single charge carrier. Let's consider the electrons in the conduction band: their concentration will be $n_n \simeq N_D$ in the n -type region, and $n_p \simeq n_i^2/N_A$ in the p -type region (according to the mass law from Equation 2.6). The diffusion equation describes the diffusion current:

$$J_{n,\text{diff}} = q D_n \frac{dn(x)}{dx}, \text{ with } D_n = \frac{\mu_n k_B T}{q}, \quad (2.16)$$

where D_n is called the electron diffusion coefficient. Equation 2.11 gives an expression for the drift current, so the sum of drift and diffusion currents equaling zero can be written as:

$$J_{n,\text{drift}} + J_{n,\text{diff}} = 0 \Rightarrow n(x) \mu_n \mathcal{E}(x) + \mu_n k_B T \frac{dn(x)}{dx} = 0. \quad (2.17)$$

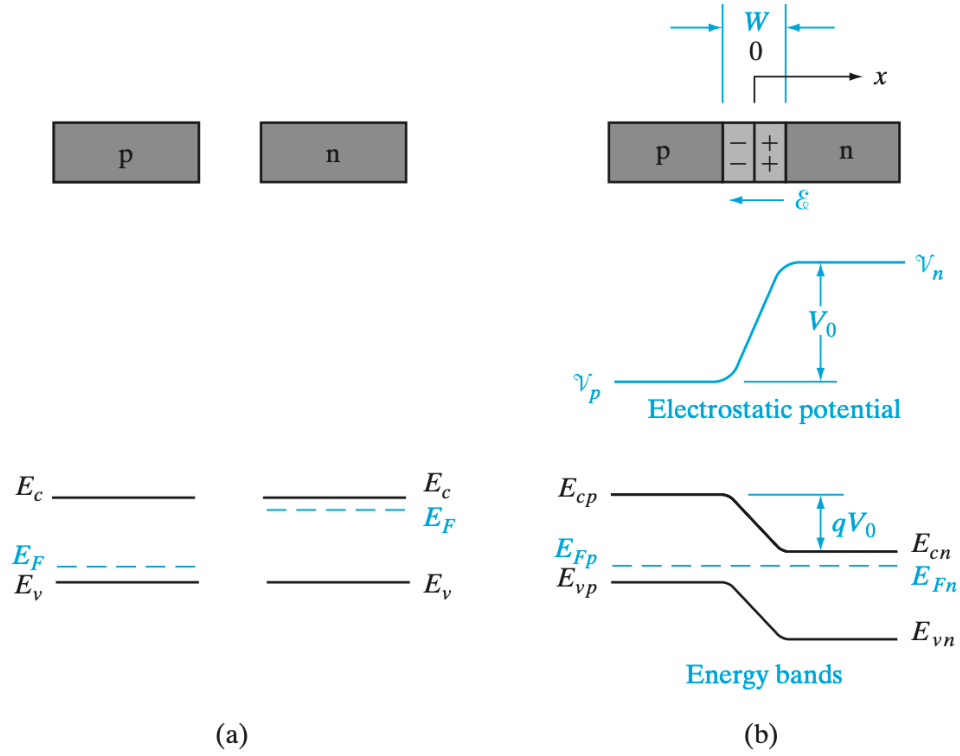


FIGURE 2.7: (a) Isolated p and n -type doped semiconductor with a schematic diagram of their band energy and Fermi level. (b) A p - n junction is formed when the two materials are brought together. A space charge region with thickness W is formed at the interface, giving rise to a built-in electrostatic potential V_0 , and an electric field $\mathcal{E}$. Taken from [79].

Since we are considering the current density as unidimensional along the x -axis, we can write the electric field $\mathcal{E}(x)$ as the derivative of the potential $\mathcal{E}(x) = -dV(x)/dx$. After some math, we find that:

$$\int_{V_p}^{V_n} dV = \frac{k_B T}{q} \int_{n_p}^{n_n} \frac{dn}{n} \Rightarrow V_0 = \frac{k_B T}{q} \ln \frac{N_D N_A}{n_i^2}. \quad (2.18)$$

Equation 2.18 shows that the built-in electrostatic potential V_0 at the equilibrium depends on the concentration of donor N_D and acceptor N_A .

2.2.2 Electric Field and Built-in Potential of the p - n Junction

We can assume the depletion approximation, i.e., the space charge region has no free charge carriers, all the dopants are ionized, and the charge is created uniquely by uncompensated charges $N_D^+ \equiv N_D$ and $N_A^- \equiv N_A$ left after the diffusion of electrons and holes. To treat the general case, we can also assume that the concentrations of donors, N_D , and acceptors, N_A , are different, i.e., $N_A \neq N_D$. The space charge region extends by x_{n0} inside the n -type semiconductor and $-x_{p0}$ inside the p -type semiconductor. Figure 2.8 shows the width W of the space charge region, and the charge density and electric field $\mathcal{E}(x)$ at the interface.

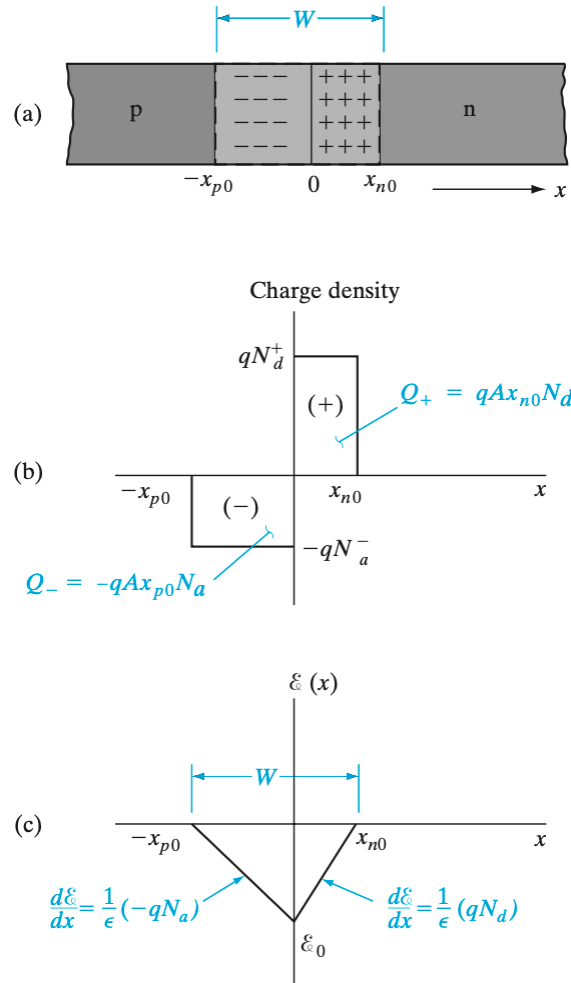


FIGURE 2.8: (a) Detail of the width W of the space charge region, which extends by x_{n0} inside the n -type semiconductor and $-x_{p0}$ inside the p -type semiconductor. (b) Space charge inside the junction is created by uncompensated donor and acceptor atoms. (c) Electric field $\mathcal{E}(x)$ across the junction. Taken from [79].

To calculate the electric field $\mathcal{E}(x)$, we can use the Poisson equation, with the charge density $\rho(x)$ given by uncompensated atoms:

$$\frac{d\mathcal{E}(x)}{dx} = \frac{\rho(x)}{\epsilon}, \quad (2.19)$$

where ϵ is the product of the vacuum dielectric constant ϵ_0 times the relative dielectric constant of the semiconductor ϵ_r (for silicon $\epsilon_r = 11.7$). As visible from Figure 2.8, there are two separate regions, the first from $-x_{p0}$ to 0, and the second from 0 to x_{n0} . Hence, we can write the Poisson equation twice:

$$\frac{d\mathcal{E}(x)}{dx} = \frac{1}{\epsilon}(-qN_A), \quad -x_{p0} < x < 0, \quad (2.20)$$

$$\frac{d\mathcal{E}(x)}{dx} = \frac{1}{\epsilon}(qN_D), \quad 0 < x < x_{n0}. \quad (2.21)$$

By integrating, we can find the analytical expression of the electric field $\mathcal{E}(x)$. Still, we have to ensure the condition of continuity, i.e., for $\mathcal{E}(0^+) = \mathcal{E}(0^-) = \mathcal{E}_0$ from both

sides. We find that:

$$\mathcal{E}(x) = -\frac{q N_A}{\epsilon}(x + x_{p0}), \quad -x_{p0} < x < 0, \quad (2.22)$$

$$\mathcal{E}(x) = \frac{q N_D}{\epsilon}(x - x_{n0}), \quad 0 < x < x_{n0}. \quad (2.23)$$

The condition of continuity defines the maximum achievable electric $\mathcal{E}_0$ field for certain dopings:

$$\mathcal{E}_0 = -\frac{q}{\epsilon} N_A x_{p0} = -\frac{q}{\epsilon} N_D x_{n0}. \quad (2.24)$$

This equation also confirms the intuitive condition to have charge neutrality in the equilibrium condition, which is $N_A x_{p0} = N_D x_{n0}$. Recalling that the electric field $\mathcal{E}(x)$ is the derivative of the potential $\mathcal{E}(x) = -dV(x)/dx$, one can find the analytical expression for the potential $V(x)$ starting from the electric field. The relevant quantity is the built-in electrostatic potential V_0 , which is found to be:

$$V_0 = V_n - V_p = \frac{q}{2\epsilon} (N_A x_{p0}^2 + N_D x_{n0}^2) \Rightarrow V_0 = \frac{q}{2\epsilon} \frac{N_A N_D}{N_A + N_D} W^2. \quad (2.25)$$

The second expression emphasizes the dependence of the built-in electrostatic potential V_0 on the square of the space charge region width W .

2.2.3 Depletion Depth of the p - n Junction

The width of the space charge region W , sometimes called the *depletion depth*, is obtained inverting Equation 2.25:

$$W = \sqrt{\frac{2\epsilon V_0}{q} \left(\frac{1}{N_D} + \frac{1}{N_A} \right)}. \quad (2.26)$$

Equation 2.18 gives an alternative expression for the built-in electrostatic potential V_0 that depends solely on the dopings of the two semiconductor regions. Replacing V_0 from Equation 2.18 into Equation 2.26, it's clear that the depletion depth at equilibrium also depends solely on the dopings of the two semiconductor regions. The depletion depths x_{n0} and x_{p0} inside the two types of semiconductors are different, as we assume that the concentrations of donors N_D and acceptors N_A are different. At all times, it must hold the condition of charge neutrality $N_A x_{p0} = N_D x_{n0}$, which gives:

$$x_{p0} = \frac{W}{1 + N_A/N_D}, \quad x_{n0} = \frac{W}{1 + N_D/N_A}. \quad (2.27)$$

2.2.4 Depletion Capacitance of the p - n Junction

A p - n junction at equilibrium has a space charge region with a width W that contains no free charge carriers. This situation is often compared to a parallel plate capacitor, in which an area with no free charges separates two charged plates. Starting from the capacitance for a parallel plane capacitor, we can calculate the capacitance for the p - n junction with a transversal section with area A , and made of a semiconductor with dielectric constant ϵ :

$$C = \epsilon \frac{A}{W} = A \sqrt{\frac{q\epsilon}{2V_0} \frac{N_A N_D}{N_A + N_D}}. \quad (2.28)$$

In this case, the capacitance of the p - n junction solely depends on the dopings of the two sides of the semiconductor. We can also see that the junction capacitance is inversely proportional to the square of the built-in electrostatic potential.

2.2.5 External Bias to the p - n Junction

So far, we have treated a p - n junction under the equilibrium condition without external bias. The space charge region has no free carriers, a crucial requirement for building semiconductor detectors. At room temperature, for intrinsic silicon, we calculated a concentration of electron-hole pairs equal to $n_i \simeq 1.2 \times 10^{10} \text{ cm}^{-3}$. An ionizing particle inside a silicon cube with dimensions $1 \text{ cm} \times 1 \text{ cm} \times 1 \text{ cm}$ produces a concentration of electron-hole pairs of about $6 \times 10^5 \text{ cm}^{-3}$: it is clear that the charges generated by the incoming particle cannot be detected, as the intrinsic charge carriers of the silicon suffocate them.

A space charge region with no free charge carriers, as provided by a p - n junction at equilibrium, makes the charges generated by the incoming particle detectable. But, with typical concentrations of dopants $N_{A(D)}$ from $\simeq 10^{13} \text{ cm}^{-3}$ to $\simeq 10^{17} \text{ cm}^{-3}$, the depletion depth w varies from $\simeq 3 \mu\text{m}$ to $\simeq 0.03 \mu\text{m}$. An external reverse bias applied to the p - n junction increases the depletion depth w , thereby increasing the number of charges generated by an ionizing particle and making the signal even more detectable.

A reverse bias given to the p - n junction means that the n -doped side is in contact with a higher potential, while the p -doped side has a lower potential. The reverse bias voltage V sums up with the built-in electrostatic potential V_0 , so the relevant equations are modified accordingly. Assuming standard dopings, the built-in potential V_0 is less than 1 V, while the reverse bias voltage can be as high as a few hundred V. A good approximation is to consider only the reverse bias voltage V and modify Equation 2.26 and Equation 2.28:

$$W = \sqrt{\frac{2\epsilon V}{q} \left(\frac{1}{N_D} + \frac{1}{N_A} \right)}, \quad C = A \sqrt{\frac{q\epsilon}{2V} \frac{N_A N_D}{N_A + N_D}}. \quad (2.29)$$

An important parameter used to characterize a p - n junction used as a particle detector is the full depletion voltage V_{FD} , which is the reverse bias voltage that is necessary to deplete the thickness d of the device completely. At the full depletion voltage V_{FD} and with reverse bias voltages $V > V_{FD}$, the capacitance C_{FD} is the lowest, and it is called the full depletion capacitance:

$$V_{FD} = \frac{q d^2}{2\epsilon} \frac{N_A N_D}{N_A + N_D}, \quad C_{FD} = \frac{\epsilon A}{d}. \quad (2.30)$$

Apart from a reverse bias applied to a p - n junction, a forward bias can also be applied to the p - n junction. The built-in potential is overcome, and charge carriers flow freely between the n and p sides of the p - n junction. Forward bias is used in various applications, including rectifiers made with diodes, light-emitting diodes (LEDs), and bipolar junction transistors (BJTs). Still, it is not used to create sensors for particle detection, so it will not be discussed further.

2.2.6 The p - n Junctions for the MONOLITH Project

We can now apply the basic concepts and formulas on the p - n junction that we found to a real physical object. For example, we take the ASIC of the second MONOLITH prototype without a gain layer and examine some values. We use the information given in Section 1.2.6 to calculate the depletion depth W , the depletion voltage V_{FD} , the depletion capacitance C_{FD} , and the shape of the electric field. Figure 2.9 shows a transversal cut of the prototype and the shape of the electric field profile for a bias voltage HV of 200 V from a TCAD simulation.

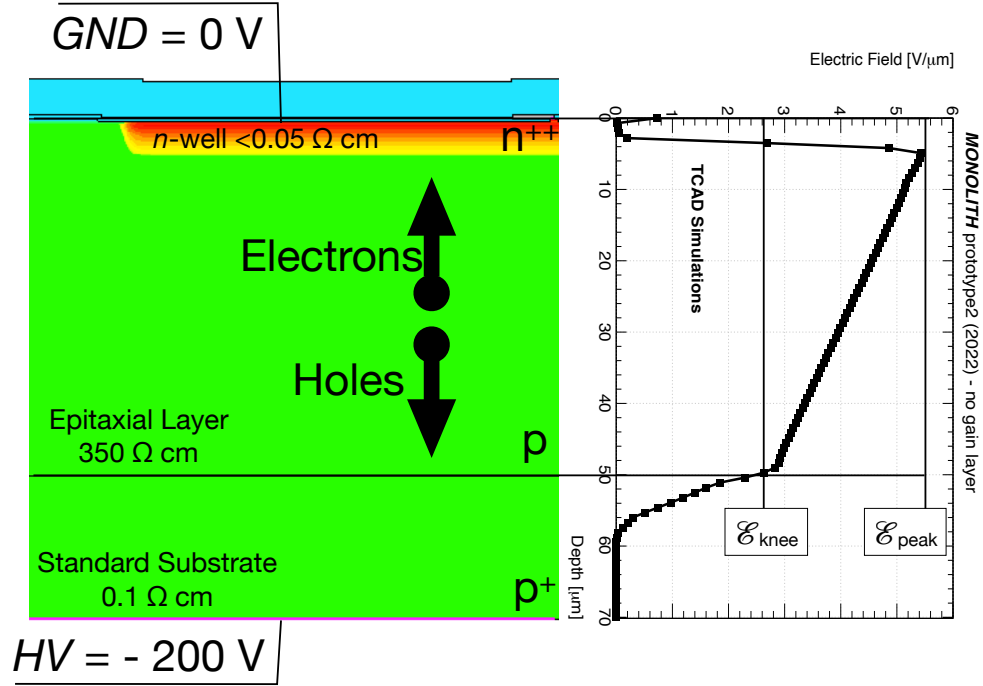


FIGURE 2.9: (Left) Transversal cut of the second MONOLITH prototype without a gain layer. The different dopings and materials are represented as follows: red represents n -type silicon, green represents p -type silicon, and blue represents silicon oxide. The given HV is a reverse bias to the p - n junction, enough to over-deplete the sensor. (Right) Electric field profile for a bias voltage HV of 200 V from a TCAD simulation.

The p - n junction of the second MONOLITH prototype, without a gain layer, is realized using an epitaxial p -type layer and an implanted n^{++} -type layer. The meaning of n^{++} is that the concentration of n dopants is much higher than the concentration of p dopants, so the junction is unbalanced ($N_D \gg N_A$). The donor atoms in the n^{++} -type region have a peak concentration of $N_D = 1 \times 10^{17} \text{ cm}^{-3}$, while the acceptor atoms in the p -type region have a uniform concentration of $N_A = 4.5 \times 10^{13} \text{ cm}^{-3}$. Using Equation 2.12, we can calculate the resistivities of $\rho_p \simeq 350 \Omega \text{ cm}$ for the p -type layer, and $\rho_n \simeq 0.05 \Omega \text{ cm}$ for the n -type layer, as indicated also on the left image of Figure 2.9.

The HV difference of 200 V is given as a reverse polarization, enough to over-deplete the junction. The donor dopants were implanted in the epitaxial p layer so that we can consider the thickness of the space charge region as equal to the thickness of the epitaxial p region ($d = 50 \mu\text{m}$). Using Equation 2.30 and the thickness d of the

space charge region, we can calculate the full depletion voltage V_{FD} :

$$V_{FD} = \frac{q d^2}{2 \epsilon} \frac{N_A N_D}{N_A + N_D} \simeq \frac{d^2}{2 \epsilon} q N_A = \frac{d^2}{2 \epsilon \mu_h \rho_p} \simeq 73 \text{ V}. \quad (2.31)$$

Equation 2.31 shows an interesting way of writing down the depletion voltage V_{FD} with a clear dependence on the resistivity ρ_p of the p -type layer and the hole mobility μ_h . This expression is valid in the approximation of a strongly unbalanced p - n junction, as is the case for the MONOLITH Prototype 2, where $N_D \gg N_A$.

The capacitance C_{FD} of the depleted p - n junction is calculated using Equation 2.30. The thickness of the space charge region is $d = 50 \mu\text{m}$, and the area of an hexagonal pixel with side $s = 65 \mu\text{m}$ is $A = 1.1 \times 10^4 \mu\text{m}^2$. Hence the full-depletion capacitance $C_{FD} = 23 \text{ fF}$. However, the total capacitance of the detector is dominated by the inter-pixel regions, rather than by the capacitance of the junction. The six sides of the hexagon contribute about $C_{\text{side}} \simeq 10 \text{ fF}$ each side, and the total capacitance of a single pixel is $C_{\text{det}} = C_{FD} + 6 C_{\text{side}} \simeq 83 \text{ fF}$.

Finally, we can use Equation 2.24 to calculate the maximum electric field $\mathcal{E}_{\text{max}}$ for a bias voltage equal to the full-depletion voltage V_{FD} . In approximating a strongly unbalanced p - n junction, all the space charge region lies inside the p -doped material, hence $x_{p0} \simeq d$. We can calculate $|\mathcal{E}_{\text{max}}| \simeq 2.9 \text{ V}/\mu\text{m}$ from:

$$|\mathcal{E}_{\text{max}}| \simeq \frac{q}{\epsilon} N_A d = \frac{d}{\epsilon \mu_h \rho_p} = \frac{2 V_{FD}}{d}. \quad (2.32)$$

When the reverse bias voltage HV is enough to over-deplete the p - n junction, all the excess voltage $HV - V_{FD}$ in the thickness of the space charge region d gives a constant electric field with height $|\mathcal{E}_{\text{knee}}| = (HV - V_{FD})/d \simeq 2.6 \text{ V}/\mu\text{m}$. The value of $|\mathcal{E}_{\text{knee}}|$ is visible in the right image of Figure 2.9, and matches the expectation from the simple calculation. The excess voltage $HV - V_{FD}$ also increases the maximum value of the electric field $\mathcal{E}_{\text{peak}} = \mathcal{E}_{\text{max}} + \mathcal{E}_{\text{knee}} \simeq 5.5 \text{ V}/\mu\text{m}$. This value of $|\mathcal{E}_{\text{peak}}|$ is also visible in the right image of Figure 2.9, and matches the expectation from the simple calculation.

Equation 2.32 shows an interesting link between the maximum electric field $\mathcal{E}_{\text{max}}$ and the full-depletion voltage V_{FD} . When the p - n junction is overdepleted, there is a simple expression to find $\mathcal{E}_{\text{peak}}$ and $\mathcal{E}_{\text{knee}}$:

$$|\mathcal{E}_{\text{peak}}| = \frac{HV + V_{FD}}{d}, \quad |\mathcal{E}_{\text{knee}}| = \frac{HV - V_{FD}}{d}. \quad (2.33)$$

2.3 Transistors

Transistors (whose name means *transfer resistors*) are multi-junction semiconductor devices that control the flow of electrical current. They are the fundamental building blocks of modern electronics, including computer chips and memory, each of which can contain up to billions of miniaturized transistors.

In modern experimental particle physics, transistors are commonly used to build the electronics for particle detectors, enabling the amplification and shaping of signals, as well as digital operations. The most commonly used type of transistors in this field is:

- *Bipolar Junction Transistors*, shortly BJTs, which consist of multiple p - n junctions, and control the flow of current via a current;
- *Metal-Oxide-Semiconductor Field Effect Transistors*, shortly MOSFETs, consisting of a metal-oxide-semiconductor junction, and control the current flow via an electric field.

The MONOLITH project utilizes both types of transistors, with some technical modifications to make them more suitable for a monolithic ASIC capable of high precision in space and time measurements. The BJTs are modified with a grading of germanium in the base, giving rise to *Heterojunction Bipolar Transistors*, commonly referred to as HBTs, and they are used for amplifying analog signals. The MOSFETs are incorporated into the monolithic process in a specialized variant that extends the voltage handling capability of the transistors, resulting in *High-Voltage CMOS*, commonly abbreviated as HV-CMOS. They are used for circuit biasing and digital logic. The integration of SiGe HBTs and HV-CMOS is done using the special SiGe BiCMOS 130 nm technology, called *SG13G2*, from the IHP foundry. This technology is the basis for the MONOLITH project, as already discussed in Section 1.2.

2.3.1 Bipolar Junction Transistors

Bipolar Junction Transistors (BJTs) are constructed by sandwiching a layer of one type of semiconductor between two layers of the opposite kind, resulting in either an n - p - n or p - n - p structure. The three terminals are referred to as the emitter, base, and collector. Figure 2.10 shows a p - n - p transistor with the emitter-base junction forward biased and the base-collector junction reverse biased. Under this condition, holes are injected from the emitter into the base. Some holes can diffuse through the base and are swept into the collector due to the reverse bias. The result is a collector current that is highly sensitive to the much smaller base current, allowing the device to function as an amplifier.

An ideal p - n - p transistor must satisfy two main requirements: almost all the holes injected through the base must diffuse into the collector, and there must be no back injection of electrons from the base to the emitter. For the first requirement, the n -type region must be narrow compared to the diffusion length of the holes: $W_b \ll L_p$, where W_b is the width of the base, and $L_p = \sqrt{D_p \tau_p}$ is the diffusion length of the holes (D_p is the diffusion coefficient, and τ_p is the hole lifetime). In this way, the holes have the time to survive the diffusion into the base and reach the collector intact. For the second requirement, the base must be lightly doped for the emitter, so that back injection of electrons occurs in a negligible manner compared to the injection of holes.

To understand the high-frequency performance of the BJT, we consider the holes collected by the base, which will see a base resistance r_b during their path in the base.

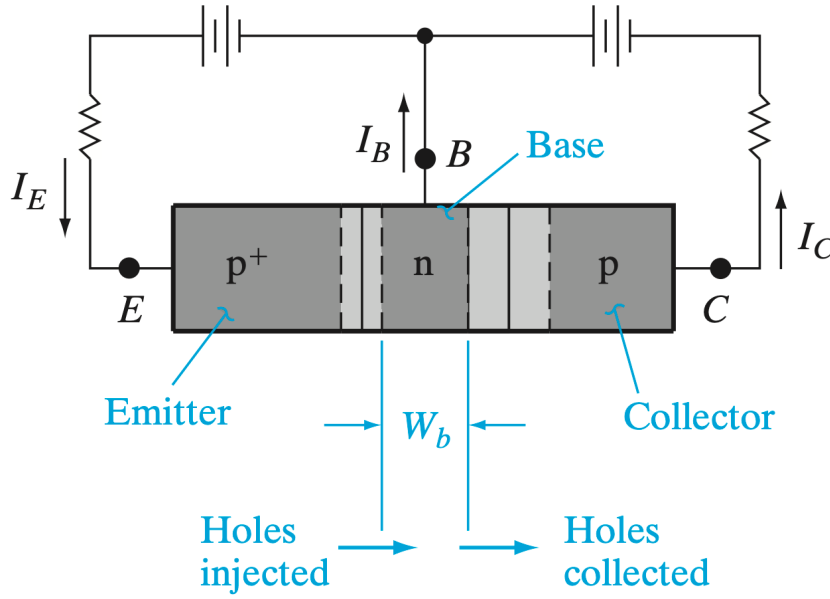


FIGURE 2.10: (Top) Band diagram of a p - n - p transistor with a Si base (solid black line) and with a SiGe graded base (dotted blue line). (Bottom) Concentration of acceptor in the base of a p - n - p transistor with a Si base (solid black line) and concentration grading of Ge in a SiGe graded base (dotted blue line). Taken from [79].

The base resistance r_b depends on the resistivity of the base ρ_b , the length of the path l_b , and the cross section of the base A_b , which is directly proportional to the width of the base: $A_b \propto W_b$. Therefore, according to the microscopic Ohm's law:

$$r_b \simeq \rho_b \frac{l_b}{A_b} \propto \frac{\rho_b}{W_b}. \quad (2.34)$$

Both the requirements for an ideal p - n - p transistor go in the direction of increasing the base resistance r_b , either by reducing the width of the base W_b , or by lowering the donor doping of the base N_D (which means increasing the resistivity ρ_b).

It can be demonstrated that the BJT does not behave like a perfect amplifier at high frequencies. Internal parasitic capacitances and resistances start to limit the response, acting like a low-pass filter with a time constant τ . The most relevant contributions are due to the base resistance r_b and the base-emitter capacitance C_{be} , creating effectively a low-pass filter with a cutoff frequency f_T :

$$f_T \simeq \frac{1}{2\pi\tau} = \frac{1}{2\pi r_b C_{be}}. \quad (2.35)$$

To improve high-frequency behaviour, the cutoff frequency f_T must be high, but it is limited by the minimum base resistance that is reachable given the constraints on the base width W_b and the base resistivity ρ_b .

Another important aspect of the BJTs is their ability to amplify current. We define the common-emitter gain β as the ratio of collector current I_C divided by base current I_B , when the BJT is biased in a common-emitter configuration: $\beta = I_C / I_B$. The n -type base region must maintain electrostatic neutrality, so the presence of excess holes in transit from emitter to collector calls for compensating excess electrons from the base contact. The average excess hole spends τ_t seconds inside the base, defined as the transit time from the emitter to the collector. At the same time, the average electron

waits τ_p seconds for recombination. Since the base width $W_b \ll L_p$, the hole transit time τ_t is much less than the average hole lifetime τ_p in the base. Hence, for each electron entering from the base contact, τ_p/τ_t holes can pass from the emitter to the collector. The ratio of collector current to base current can be written as:

$$\frac{I_C}{I_B} = \beta = \frac{\tau_p}{\tau_t}. \quad (2.36)$$

To have a high current amplification, a short hole transit time τ_t is needed, which is reached as long as $W_b \ll L_p$ is satisfied. But a base width W_b that is too short increases the base resistance r_b , reducing the cutoff frequency f_T , thus limiting the high-frequency performance. If we consider the bandwidth of the BJT as proportional to the cutoff frequency f_T , it is clear that there is a trade-off between the amplification and the bandwidth.

2.3.2 Heterojunction Bipolar Transistors

Both requirements for a BJT, which are summarized as a thin base and a low doping in the base, go in the direction of slowing down the transistor. The BJT, as it is, is not used for applications that require good performance at high frequencies. To address the problems and improve the high-frequency performance, it is possible to introduce a grading of germanium in the base of the BJT, creating a *Heterojunction Bipolar Transistor*, abbreviated as HBT.

The grading of germanium inside the base of the HBT engineers the band diagram of the transistor, creating a small electric field within the base that causes the majority carriers to move via drift, rather than diffusion, as in a traditional bipolar junction transistor (BJT). Figure 2.11 shows the band diagram comparison for an *n-p-n* BJT and an HBT, together with the detail of the concentration of germanium in the base, which is graded increasing towards the collector.

The primary difference between bipolar junction transistors (BJTs) and heterojunction bipolar transistors (HBTs) lies in the transport mechanism within the base: in BJTs, it is primarily dominated by diffusion, whereas in HBTs, it is primarily dominated by drift. The drift-based transport in HBTs significantly reduces the transit time τ_t , allowing for a higher gain following Equation 2.36. Moreover, since the transit time is reduced, a thicker base is possible without losing efficiency in collecting the injected electrons. Equation 2.34 shows that increasing the width of the base W_b makes the base resistance r_b smaller. Additionally, a higher doping level in the base is also possible. Hence, the cutoff frequency f_T increases according to Equation 2.35. For this reason, HBTs are particularly suited for high-speed applications, such as radio-frequency amplifiers and digital logic circuits operating at GHz frequencies.

The advantages of Ge grading in the base extend even beyond those already mentioned. Low-noise amplification is particularly relevant in experimental particle physics, as it yields a higher signal-to-noise ratio and lower electronic time jitter. These results translate into high efficiency, low noise-hit rate, and precise time resolution particle detectors. To understand this effect, we define the series *Equivalent Noise Charge*, shortly ENC_{series} , which is the output voltage noise of the amplifier in terms of equivalent electrons at the input. It is possible to demonstrate that [85, 86]:

$$ENC_{\text{series}} \propto \sqrt{k_1 \frac{C_{\text{tot}}^2}{\beta} + k_2 r_b C_{\text{tot}}^2}. \quad (2.37)$$

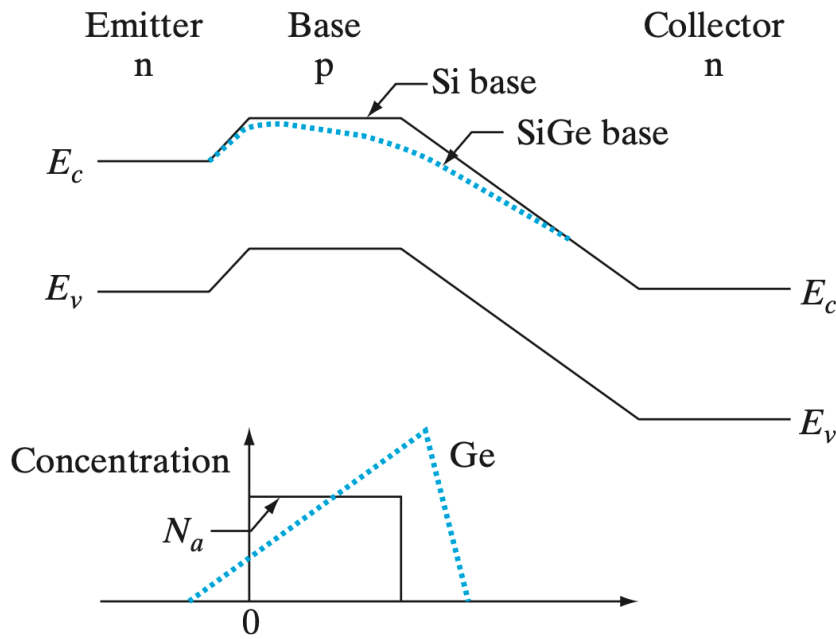


FIGURE 2.11: (Top) Band diagram of an n - p - n transistor with a Si base (solid black line) and with a SiGe graded base (dotted blue line). (Bottom) Concentration of acceptor in the base of an n - p - n transistor with a Si base (solid black line) and concentration grading of Ge in a SiGe graded base (dotted blue line). Taken from [79].

The coefficients k_1 and k_2 depends on physical parameters of the transistor, C_{tot} is the total capacitance considering the input and parasitic capacitances, r_b is the base resistance, and β is the common-emitter current gain. An HBT has a more minor base resistance r_b and a higher current gain β than a traditional BJT, both of which contribute to lowering the ENC_{series} .

Furthermore, HBTs exhibit robust performance at cryogenic temperatures. At low temperatures, traditional bipolar junction transistors (BJTs) often experience carrier freeze-out, which results in degraded performance. HBTs, however, benefit from the band engineering of the heterojunction, maintaining good carrier injection and gain even under cryogenic conditions. Finally, HBTs are attractive for fast, low-noise front-end electronics in scientific instrumentation, such as particle detectors operating in extreme environments.

2.3.3 Metal-Oxide-Semiconductor Field Effect Transistors

The term *Field-Effect Transistor*, shortly FET, indicates generally a device that can control a flow of current via an electric field. One way to build an FET is through a *Metal-Oxide-Semiconductor* junction, giving rise to the so-called *Metal-Oxide-Semiconductor Field-Effect Transistor*, shortly MOSFET. A typical MOSFET consists of a source, a drain, and a gate, all of which are constructed on a silicon substrate. The gate is separated from the channel by a thin layer of insulating oxide, usually silicon dioxide. Applying a voltage to the gate terminal generates an electric field that modulates the channel's conductivity between the source and drain, thereby controlling the current flow. Depending on the type of semiconductor doping, MOSFETs can be classified as n -type or p -type. Figure 2.12 shows a cross section of an n -type MOSFET.

Figure 2.12 also shows that below the contacts of the source and drain, and the channel of charge carriers, there is a depleted region with no charge carriers, as it is

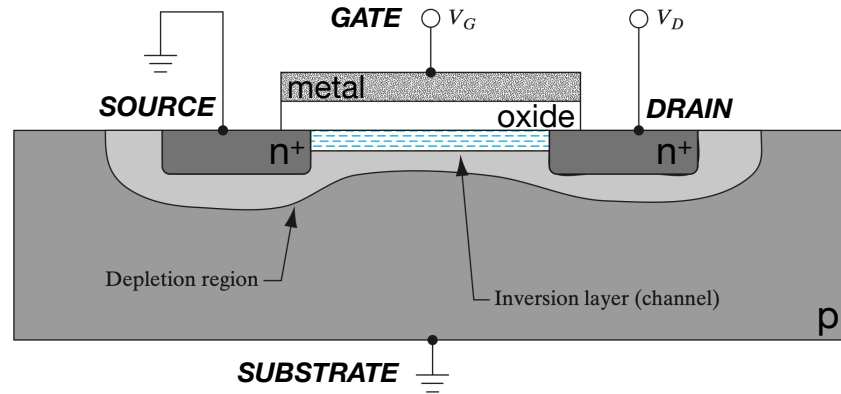


FIGURE 2.12: Cross section of an n -type MOSFET. When the voltage between gate and substrate V_G is higher than a certain threshold voltage V_T , an electron channel forms between the source and the drain, making the flow of current possible. The voltage V_G controls the current flow. Adapted from [79].

for a p - n junction in reverse bias. Since a space charge region is required to detect particles, this suggests that a MOSFET can effectively be used as a particle detector.

When both n -type and p -type MOSFETs are integrated inside a single n -well, the device takes the name of *Complementary MOS*, shortly CMOS. CMOS technology forms the basis of most integrated circuits in digital and analog systems, from micro-processors and memory devices to image sensors and analog front-end electronics.

2.3.4 High Voltage CMOS Devices

The depleted region in Figure 2.12 suggests that a standard MOSFET can be used as a particle detector. However, two main problems arise when dealing with the detector's efficiency and fast response. The drain-substrate voltage is usually of the order of a few V and not more (otherwise the MOSFET is irreversibly damaged), which makes the depleted region very thin, and the electric field inside the depleted region very low. A small depleted region results in a low probability of conversion for punctual interactions, such as X-rays, or a small number of primary electron-hole pairs for charged particles. In both cases, this reduces the detector's efficiency. On the other hand, the small electric field is not enough to overcome diffusion, which is the main mechanism of primary charge transport. As a result, charge transport and the detector's response are slow.

High Voltage CMOS, shortly HV-CMOS, devices extend standard CMOS technology by enabling operation at significantly higher voltages, typically in the tens to hundreds of volts, while maintaining compatibility with standard low-voltage logic. The applied high voltage creates a strong electric field in the depleted region. This field enables the rapid drift of electrons, which have high mobility, resulting in faster charge collection than diffusion-dominated mechanisms and improving the fast timing performance. Moreover, a higher electric field increases the depletion region, leading to a corresponding increase in efficiency. Figure 2.13 shows the cross section of an HV-CMOS.

In particle detection and fast-timing applications, HV-CMOS is particularly valuable for implementing monolithic devices, where the sensor and readout electronics are integrated on the same substrate. The integration simplifies system design, reduces material budget and costs, and improves timing performance by minimizing parasitic capacitance and shortening the drift path of charge carriers. The high electric

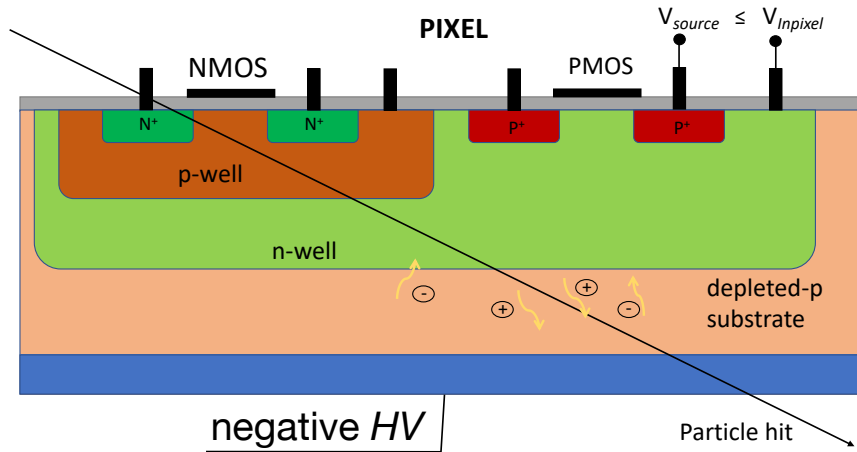


FIGURE 2.13: Cross section of an HV-CMOS. The charged particle crossing the detector creates some charge carriers that drift toward the collection electrode, the n -well, under the electric field generated by the negative high voltage ($HV < 0$ V). Adapted from [87].

field also enhances the device's radiation tolerance, a strong requirement for future particle detector applications.

HV-CMOS is a promising technology for applications that require precise time resolution, such as time-of-flight systems and 4D tracking detectors in high-energy physics. The SiGe 130 BiCMOS technology, developed by the IHP foundry for the MONOLITH project, is indeed a monolithic integration of an HBT in an HV-CMOS process.

Chapter 3

Silicon-Based Particle Detectors

This chapter explains the mechanisms that correlate particle-matter interactions with electrical signal production and extraction in silicon-based silicon detectors. It provides the conceptual and physical basis for understanding how primary ionization leads to detectable pulses and how their characteristics shape a detector's time and energy resolution.

Section 3.1 is a collection of the milestones in the evolution of silicon-based particle detectors, from initial surface-barrier detectors to present-day monolithic pixel technologies. The advances led to the high-resolution, radiation-hardened detectors used today in particle physics experiments.

Section 3.2 addresses the elementary processes that are responsible for energy deposition within the detector material. Charged particles inelastically scatter primarily from atomic electrons by ionization and excitation. The average energy loss is described by the Bethe-Bloch formula, and the Landau theory describes the fluctuations. Photons are treated separately, and the photoelectric effect is the only phenomenon included.

Section 3.3 defines the theoretical framework for relating detector moving charges to induced currents on electrodes. The instantaneous current due to drifting charge carriers is explained by the Shockley-Ramo theorem, which is crucial for modeling detection responses to particle interactions. Localized energy depositions and homogeneous ionization distributions are only a few of the present signals, which transform physical processes into electrical signals processed by the readout chain.

Section 3.4 provides the overall form of the readout electronics, i.e., amplification and shaping. At the same time, we discuss the dominant sources of electronic noise, i.e., thermal, shot, and $1/f$ noise, and their impact on the detector performance.

Section 3.5 addresses the numerous contributions to the time resolution of a detection system. Landau fluctuations resulting from energy deposition, time walk resulting from amplitude dependence, and the impact of noise are all treated completely. Other effects, including inhomogeneities caused by sensor geometry or digitization jitter, are also addressed to provide a comprehensive understanding of what limits temporal accuracy in modern particle detectors.

Section 3.6 provides the mathematical tools to calculate the space resolution of a pixel detector, with particular emphasis on hexagon-shaped pixels as in the detectors for the MONOLITH project.

Finally, Section 3.7 examines the importance of minimizing power consumption in silicon detectors, especially in the context of large-scale tracking systems for future collider experiments. The section focuses on the analog front-end, using the MONOLITH prototypes as a case study.

3.1 Brief History of Silicon-Based Particle Detectors

Particle detectors are devices that measure the properties of elementary particles, such as their energy, momentum, position, and time. Particle detectors are key devices in experimental particle physics, as they provide the interface between physical processes and the measurable signals that originate from them. A particle detector performs two general tasks: generating a primary charge when an incident particle interacts with the material, and collecting and processing the signal into an electrical output that is readable and can be digitized. The comparatively high density and low bandgap of silicon result in a larger number of primary charges compared to gases and other semiconductors. Additionally, there are numerous advantages to fabricating electronics in silicon; therefore, it comes as no surprise that silicon has become one of the primary technologies for particle detectors.

The evolution of particle detectors in silicon began in the 1950s and 1960s, when the semiconductor industry began producing high-purity, crystal-controlled doping profiles. One of the first attempts at a semiconductor particle detector was by McKay in 1951 [88] using germanium to detect α particles. Silicon applications would follow shortly, as scientists discovered that p - n junctions and silicon's response to ionizing radiation could be used to detect ionizing particles with newfound resolution. The first silicon detectors were surface barrier type, created in the early 1960s for nuclear spectroscopy [89, 90]. After the formation of a depletion layer in a metal-semiconductor junction, the devices developed exhibited improved energy resolution and lower volumes, making them suitable for use in particle physics.

In the late 1970s and early 1980s, silicon detectors became significant applications in nuclear physics and space instrumentation through the advent of planar technology. Thus, silicon strip detectors could now be produced, making precise position detection possible by splitting the detector surface into thin readout channels. The NA32 Experiment [91] at CERN was the first beamline experiment using a silicon microstrip detector. Later, silicon microstrip detectors were also employed in particle accelerator experiments for the reconstruction of real-time charged particle tracks, e.g., CERN's UA2 [92] at the SPS collider, enabling modern theoretical and experimental tracking systems.

The 1990s and 2000s were periods of revolutionary change, marked by the introduction of pixel detectors and monolithic technologies, particularly for large-scale experiments at colliders like ATLAS [93] and CMS [94] at the Large Hadron Collider (LHC). Silicon sensors evolved from passive devices to complex, integrated systems with low-noise, fast readout electronics. Silicon detectors are now ubiquitous in particle physics experiments, ranging from the innermost layers of collider trackers to precision timing detectors.

3.2 Interaction of Particles with Matter

When a particle passes through matter, it can interact with the material's atoms, producing detectable signals that can be used in particle detectors. The fundamental forces of nature govern these interactions (see Section 1.1.1), but for most practical applications in detectors, the dominant processes are electromagnetic in nature.

The electromagnetic force is significantly more potent than the weak or gravitational forces at the scale of individual particles. For instance, the electromagnetic interaction is about 10^{36} times stronger than gravity and many orders of magnitude stronger than the weak nuclear force over atomic distances. Although much more potent at sub-femtometer ranges, the strong nuclear force is short-ranged and typically relevant only within atomic nuclei. Therefore, electromagnetic processes are the most pertinent for detecting high-energy particles in solid-state materials, such as silicon.

The primary objective of the MONOLITH project is to track charged particles that can interact electromagnetically with medium atoms through ionization and energy loss. When ionization occurs, sufficient energy is transferred to bound electrons, which can be ejected from atoms, creating free electron-hole pairs in silicon. At room temperature, the average energy required to produce one electron-hole pair in silicon is approximately 3.6 eV.

Electrically neutral photons interact primarily through the photoelectric effect, Compton scattering, and pair production. Although these processes are less relevant to the final objectives of the MONOLITH project, they are often used in the laboratory to characterize the ASIC's performance. Moreover, given the intrinsic advantages in signal-to-noise ratio and time resolution of the SiGe BiCMOS technology, the detection of photons has become an interesting future development of the MONOLITH project (see Section 1.2.9).

The primary reference for the interaction of particles with matter is the Annual Review of the Particle Data Group, available at [10]. The review compiles many ideas developed over the years, beginning with the works by Bethe [95] and Landau [96].

3.2.1 Interaction of Charged Particles with Matter

Charged particles traversing matter interact primarily via the electromagnetic force with the atomic electrons of the material. As they pass through, they lose energy by ionizing atoms along their path, creating electron-hole pairs. The mean rate of this energy loss per unit distance is described by the Bethe-Bloch formula, which, for a particle of charge z and velocity βc on a medium with atomic number Z (number of electrons per atom) and mass number A (number of nucleons), reads:

$$\left\langle -\frac{dE}{dx} \right\rangle = Kz^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \left(\frac{2m_e c^2 \beta^2 \gamma^2 W_{\max}}{I^2} \right) - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right], \quad (3.1)$$

K is a dimensional constant, I is its mean excitation potential, $W_{\max}$ is the maximum energy transferable in a single collision. Equation 3.1 has a good validity for silicon in a range $10^{-1} \lesssim \beta\gamma \lesssim 10^3$, and this is the range normally involved in experiments with the prototypes for the MONOLITH project. Figure 3.1 shows the average energy loss per unit distance dE/dx as a function of $\beta\gamma$ of the incident charged particle.

Figure 3.1 shows that the Bethe-Bloch formula has a minimum around $\beta\gamma \simeq 3.5$. A particle with energy such that its energy loss is near the minimum of the Bethe-Bloch curve is called a *Minimum Ionizing Particle*, shortly MIP. For all the $\beta\gamma > 3.5$, the average energy loss is within 20 % of that of a MIP, so, with a good approximation, a

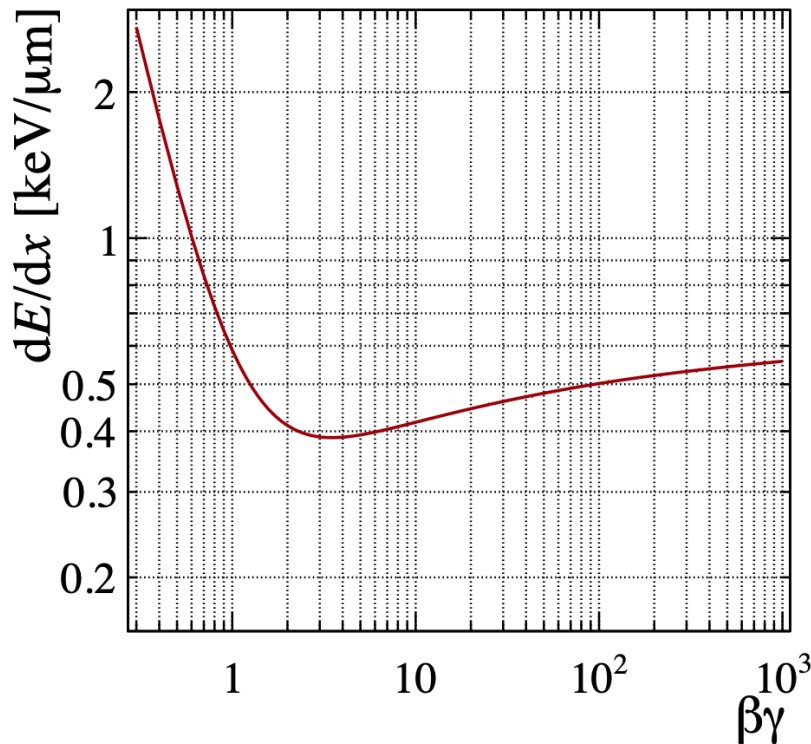


FIGURE 3.1: Average energy loss per unit distance dE/dx as a function of $\beta\gamma$ of the incident charged particle. Created via a Garfield++ simulation, using the Garfield++ software [97], and the Heed toolkit [98].

MIP is almost always considered. As visible from Figure 3.1, a MIP typically loses about $0.4 \text{ keV}/\mu\text{m}$ in silicon.

This formula provides the average energy loss but does not describe the fluctuations in energy loss observed on a track-by-track basis. Since the energy loss per unit length Δ/x is a stochastic process, we must expect a probability distribution function $f(\Delta/x)$. The function $f(\Delta/x)$ is usually called a *straggling function* and is derived considering the electromagnetic cross section. Many derivations exist, each with its own approximations and degree of complexity. The Landau theory can derive the shape of the straggling function using some approximations [96], but the result differs from the data when the silicon is thin ($\lesssim 1 \text{ mm}$). Despite the Landau theory being less precise than other theories for thin sensors (for which the derivations by Bischel [99] or the photo-absorption ionization model by Allison and Cobb [100] are often preferred), it is frequently used as a baseline to discuss the fundamental physics. Figure 3.2 shows some calculated straggling functions for 500 MeV pions in silicon with different thicknesses: the straggling functions become more and more similar to a Landau functional form when the thickness of the silicon increases.

Figure 3.2 also shows the long high-energy tail caused by rare but large energy transfers in single collisions (delta electrons), which implies that the *most probable value*, shortly MPV, is lower than the mean energy loss rate given by the Bethe-Bloch formula in Equation 3.1. The MPV is often used instead of the mean for detector design and calibration. A significant value is the MPV of electron-hole pairs generated per μm $\tilde{q}_{\mu\text{m}}$: Figure 3.2 shows that this number varies between $\simeq 70$ e-h pairs per μm for a $80 \mu\text{m}$ -thick detector, to $\simeq 80$ e-h pairs per μm for a $640 \mu\text{m}$ -thick detector.

The passage of a high-energy charged particle in a silicon medium results in the formation of ionization clusters, as several nearby atoms are ionized within a

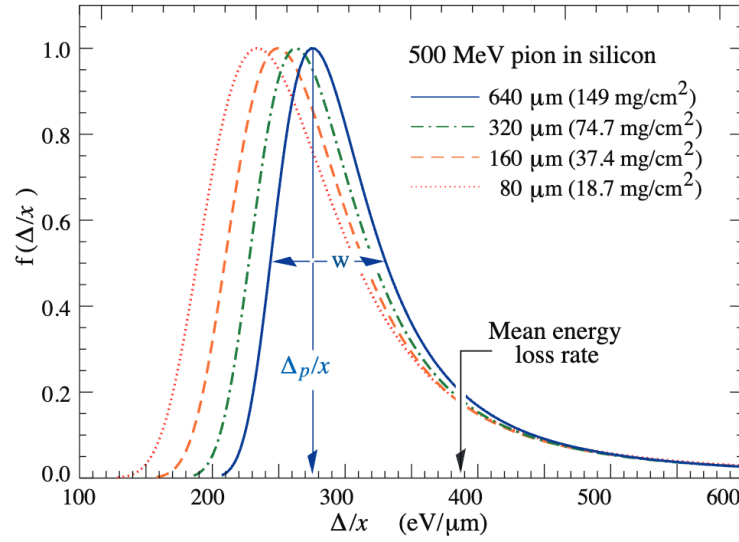


FIGURE 3.2: Straggling functions for 500 MeV pions in silicon with different thicknesses, normalized to unity at the most probable value Δ_p/x . The MPV is lower than the mean energy loss rate given by the Bethe-Bloch formula in Equation 3.1 because of the strong asymmetry to the right. Adapted from [10].

small volume. These primary clusters are typically a few nanometers in size, and the distance λ between consecutive clusters depends on the particle's energy. For particles with $\beta\gamma \gtrsim 1$, the average distance between these primary interactions is $\lambda = 0.212 \mu\text{m}$ in silicon. The number of electrons in a single cluster depends on the electromagnetic cross section. Figure 3.3 shows the probability density function of having a certain number of electrons in a cluster.

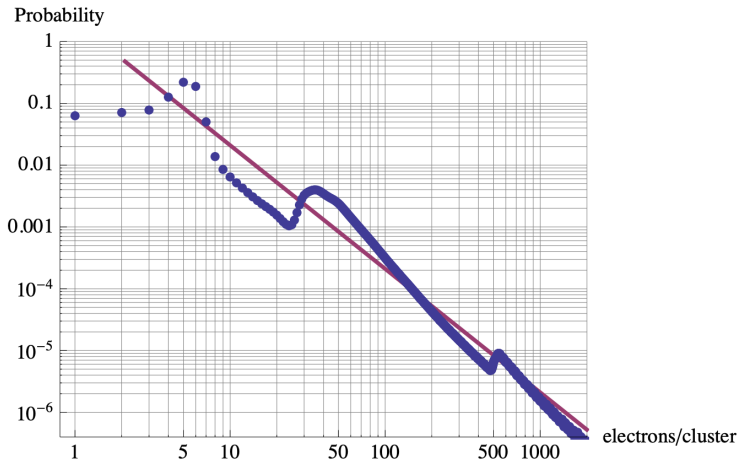


FIGURE 3.3: Probability density function of the number of electrons per cluster. The red line is the basis for the Landau distribution, while the blue points correspond to the photo-absorption ionization (PAI) model. Taken from [101].

Figure 3.3 also compares the Landau and the PAI theories, with the PAI model being the more accurate one, considering the silicon shells. The meaning of the three peaks in the electron number n per cluster is: the peak at $n \simeq 5$ electrons is the plasmon peak for collective valence electrons oscillation, the peak at $n \simeq 40$ electrons is the resonance with the L_{23} edge, and the peak at $n \simeq 600$ electrons is the resonance with the K edge (see [102] at page 88 for more details).

3.2.2 Interaction of Photons with Matter

Photons interact with matter through discrete quantum processes rather than continuous energy loss. The likelihood of a photon interacting via a specific interaction process is characterized by a cross section, which varies with the photon's energy and the atomic number Z of the target material. Figure 3.4 shows the cross section for photons on a light material like carbon as a function of the photon energy.

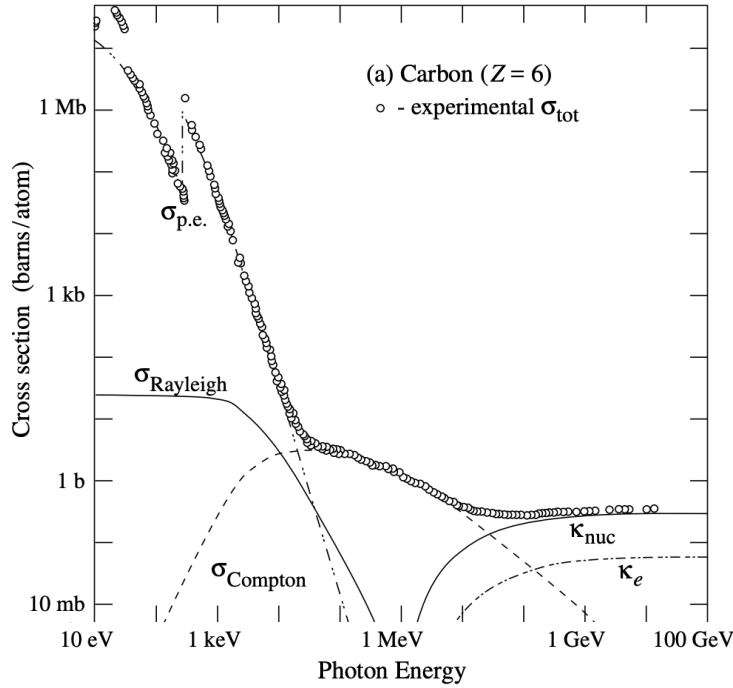


FIGURE 3.4: Cross section as a function of the photon energy for photons on carbon atoms. The dots represent the experimental data, and the various lines represent the analytical calculations for the different types of processes. Adapted from [10].

The meaning of the different cross sections in Figure 3.4 is:

- $\sigma_{p.e.}$ is the atomic photoelectric effect, where the photon is absorbed and the electron is ejected;
- σ_{Rayleigh} is the Rayleigh coherent scattering, where the atom is neither ionized nor excited;
- σ_{Compton} is the Compton incoherent scattering, where the atom is ionized or excited;
- κ_{nuc} is the pair production in the proximity of a nuclear field;
- κ_e is the pair production in the proximity of an electron field.

For light elements like carbon and silicon, Figure 3.4 shows that all relevant photon interactions are electromagnetic and occur across different regions of the electromagnetic spectrum. Photons in the infrared and visible ranges (energies below a few eV) do not interact significantly with the silicon. Their energies are too low to eject bound electrons from atoms via band-to-band transitions. However, they can still ionize electrons via other mechanisms, such as phonon-assisted transitions or multiphoton processes, although with a very low probability. The photoelectric

effect becomes relevant for ultraviolet photons and soft X-ray photons (energies from 10 eV to 10 keV), which are efficiently absorbed in silicon. Hard X-ray and low-energy Gamma ray photons (energies from 10 keV to 1 MeV) interact via both photoelectric absorption and Compton scattering. Finally, high-energy Gamma ray photons (energy > 1.022 MeV) interact mainly via pair production, as the energy threshold of the combined rest mass of the electron and positron becomes available.

From the cross section, it is possible to calculate the average absorption length of a photon in silicon, which depends strongly on the photon's energy. Figure 3.5 shows a summary of the absorption length in silicon as a function of the photon energy limited to the spectrum relevant for this thesis. The experimental points are the sum of two datasets: one taken from [103] for low energies and the other from [104] for high energies.

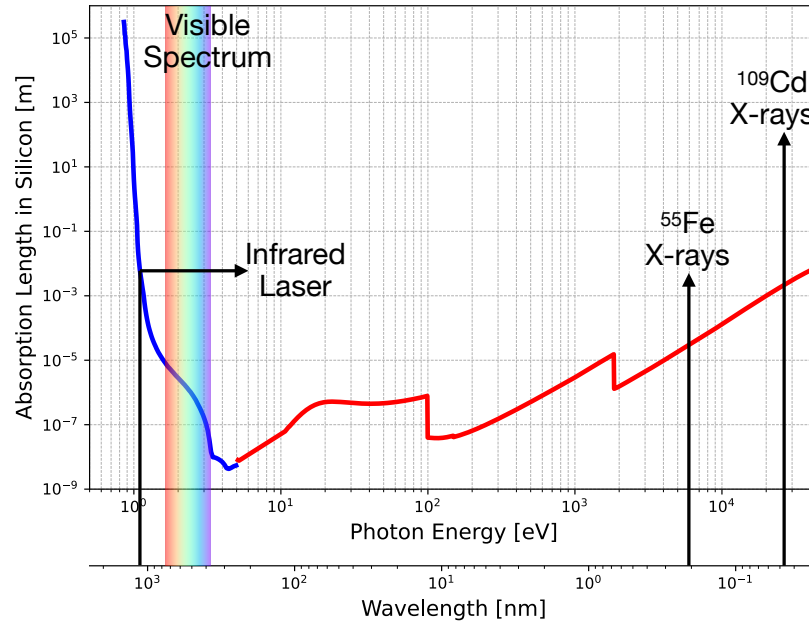


FIGURE 3.5: Absorption length in silicon as a function of the photon energy (or, equivalently, the photon wavelength). The blue dataset values are taken from [103], while the red dataset values are taken from [104]. The energies corresponding to the main sources of photons used in the rest of the thesis are indicated.

Figure 3.5 shows the wavelength λ corresponding to the photon energy E . The conversion happens using the formula:

$$\lambda = \frac{hc}{E} \Rightarrow \lambda [\text{nm}] = \frac{1240}{E [\text{eV}]}, \quad (3.2)$$

h is Planck's constant, and c is the velocity of light in vacuum.

Figure 3.5 also shows the energies corresponding to the main sources of photons used during the ASICs' characterization in the following chapters. In particular:

- an infrared laser with wavelength $\lambda \simeq 1050$ nm shows an absorption length of few mm;
- an ^{55}Fe radioactive source that emits X-rays with average energy of 5.9 keV shows an absorption length of few tens of μm ;
- a ^{109}Cd radioactive source that emits X-rays with average energy of 25 keV shows an absorption length of few mm.

The MONOLITH project uses silicon detectors with a few tens of μm of active silicon depending on the ASIC. Therefore, all three sources of photons are detectable, despite having different intrinsic probabilities due to their varying absorption lengths.

The infrared laser has photons with energy lower than the silicon bandgap, so it must create phonon-assisted transitions or multiphoton processes to ionize electrons. Since the probability of these processes is very low, a high number of photons is needed. Typically, ionization is assumed to occur uniformly within a cylinder corresponding to the laser beam. On the other hand, the ^{55}Fe and ^{109}Cd radioactive sources produce a single X-ray photon that interacts with an atom via the photoelectric effect. All the photon's energy is transferred to the ionized electron that behaves like a charged particle in the medium (see previous Section 3.2.1). The relativistic mechanics states that:

$$\beta\gamma = \frac{p}{mc} = \sqrt{\frac{E^2}{(mc^2)^2} - 1}, \quad (3.3)$$

where $\beta\gamma$ is the relativistic factor of Figure 3.1, E , p , m are respectively the energy, momentum, and mass of the particle, and c is the velocity of light in vacuum.

Using this Equation 3.3, we find that ^{55}Fe produces an electron with $\beta\gamma \simeq 0.13$, and ^{109}Cd produces an electron with $\beta\gamma \simeq 0.32$. Since $\beta\gamma$ is low, Figure 3.1 tells us that these primary electrons lose all their energy in a short path (less than $1\text{ }\mu\text{m}$), creating a few clusters localized in the short path. Finally, while for a MIP-like charged particle, ionization occurs along the entire track with a discrete number of electron-hole pairs per μm , for X-ray photons, there is a single cluster with several electron-hole pairs, depending on the energy of the incoming X-ray. The ^{55}Fe radioactive source produces $\simeq 1640$ localized electron-hole pairs, and ^{109}Cd produces $\simeq 6944$ localized electron-hole pairs, corresponding respectively to a charge of 0.26 fC and 1.11 fC .

3.3 The Shockley-Ramo Theorem

When a charge moves in an electric field, it induces an instantaneous current $i(t)$ on an electrode; the Shockley-Ramo theorem provides a mathematical framework for calculating this current. This signal, $i(t)$, is then fed into the electronic chain, where it is amplified and digitized before being analyzed. The theorem was introduced in two independent papers by Shockley [105] and Ramo [106], and it states that the induced current $i(t)$ is given by:

$$i(t) = -q \vec{v}_d(\vec{x}, t) \cdot \vec{E}_w(\vec{x}). \quad (3.4)$$

The three quantities that contribute to the induced current $i(t)$ are: the electric charge q of the moving charge, the instantaneous drift velocity $\vec{v}_d(\vec{x}, t)$, and the weighting field $\vec{E}_w(\vec{x})$. The MONOLITH project aims to reach a ps-level time resolution, so each ingredient must be perfectly tailored. One wants to have the highest possible induced current $i(t)$, which means that a lower amplifier gain can be used without limiting the bandwidth and high-frequency performance.

The first ingredient is the weighting field $\vec{E}_w(\vec{x})$. While the electric field defines the drift motion of a charge (see Equation 2.9), the weighting field $\vec{E}_w(\vec{x})$ describes the coupling between the charge and the respective electrodes. Is it mathematically determined by applying unit potential $V = 1\text{ V}$ to the measurement electrode, zeroing all the other electrodes, and solving the Poisson equation. The solution depends solely on the sensor's geometry. For example, a parallel plate geometry with an electrode of width w in a substrate with thickness d (with $w/d \gg 1$) gives a weighting potential $\phi_w = 1 - x/d$, where x is the position along the thickness of the detector.

The weighting field becomes:

$$|\vec{E}_w(\vec{x})| = -\frac{d\phi_w}{dx} = \frac{1}{d}. \quad (3.5)$$

The weighting field is optimized when it is as uniform and high as possible. The ratio w/d must be as high as possible to achieve uniformity along the sensor thickness, so the parallel plate capacitor approximation holds. The thickness d must be as small as possible to maximize the weighting field. However, both requirements increase the sensor capacitance, while one aims to minimize it to reduce noise (see Section 3.4.2 below). Hence, the geometry chosen is a compromise between a better weighting field and a smaller capacitance.

The second ingredient is the drift velocity $\vec{v}_d(\vec{x}, t)$. The drift velocity is optimized when it is as uniform and high as possible. To achieve uniformity along the sensor thickness, the electric field must be uniform; hence, the ratio w/d must be as high as possible. Moreover, Equation 2.33 suggests minimizing the full-depletion voltage to make the $\mathcal{E}_{\text{peak}}$ and $\mathcal{E}_{\text{knee}}$ similar. This effect can be achieved by increasing the doping ρ_p or reducing the thickness d , as stated in Equation 2.31. To maximize the drift velocity, the electric field must be high enough to make the drift velocity reach the saturation value.

The third and final ingredient is the electric charge q of the moving charge. There is no trivial way related to geometry to increase this charge q . One may think that it is possible to increase the charge q deposited by a MIP by increasing the thickness d , since a MIP creates $\tilde{q}_{\mu\text{m}}$ electron-hole pairs per μm , so $q = \tilde{q}_{\mu\text{m}}d$. However, the weighting field is $E_w = 1/d$, so the increase of the thickness d will increase the charge q , but it will accordingly decrease the weighting field E_w , with no net effect on the induced current $i(t)$. A solution not related to geometry is introducing a gain layer, as was done for the prototypes with the PicoAD sensor. The gain layer effectively multiplies the primary charge q_p by a factor G : $q = Gq_p$, hence a higher drifting charge, therefore a higher induced current.

The induced current can be calculated for two trivial cases with the two approximations just discussed: a parallel plate geometry to have a uniform weighting field $E_w = 1/d$ and a high electric field to have the drift velocity saturated $v_d = v_{\text{sat,e(h)}}$ for electrons (or holes). We use $d = 50 \mu\text{m}$ and an $n^+ - p$ junction, which are the characteristics of the second MONOLITH prototype without a gain layer. The first trivial case is a single cluster of charge q generated at a position x along the sensor thickness, similar to the charge deposition caused by an X-ray photon. The second trivial case is a uniform charge deposition with charge per unit of length $\tilde{q}_{\mu\text{m}}$. A uniform charge deposition is similar to how a MIP interacts, but without the typical charge fluctuations described by the Landau theory in Figure 3.3. Figure 3.6 illustrates the induced current signals for the two trivial examples.

In the case of a single charge cluster, the electrons drift towards the collection electrode with a velocity $v_{\text{sat,e}}$. They drift for a distance x and are then collected. Hence, after a time $t = x/v_{\text{sat,e}}$, they disappear. For the holes, the distance is $d - x$ and the saturated drift velocity is $v_{\text{sat,h}}$. The induced current for a single cluster $i_{\text{SC}}(x, t)$ reads:

$$i_{\text{SC}}(x, t) = \frac{q v_{\text{sat,e}}}{d} \Theta\left(\frac{x}{v_{\text{sat,e}}} - t\right) + \frac{q v_{\text{sat,h}}}{d} \Theta\left(\frac{d-x}{v_{\text{sat,h}}} - t\right), \quad (3.6)$$

with $\Theta(t)$ the standard Heaviside function, equal to 1 when $t > 0$.

In the case of a uniform charge deposition, the charge profile is a collection of

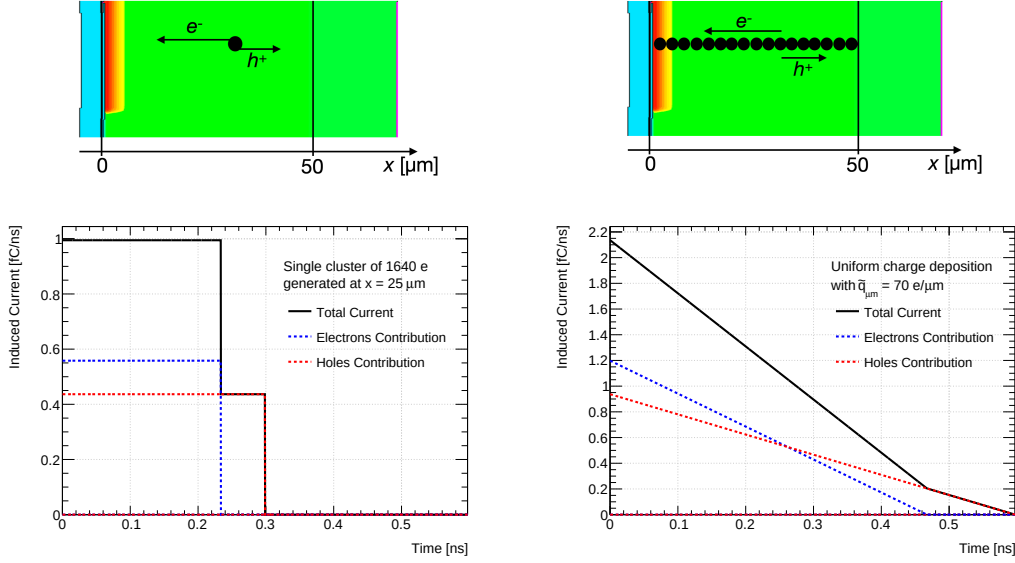


FIGURE 3.6: Examples of the induced current signals for a single charge cluster (bottom left) and for a uniform deposition not affected by Landau fluctuations (bottom right). The top figures schematize the charge deposition inside the 50 μm -thick silicon detector considered.

single clusters of electrons and holes along the path of the MIP. At the beginning of the drift, all the charges $\tilde{q}_{\mu\text{m}} d$ induce current, but, as long as the electrode collects clusters, they stop contributing to the induced current. The induced current for a MIP-like uniform deposition $i_{\text{MIP}}(t)$ reads:

$$i_{\text{MIP}}(t) = \tilde{q}_{\mu\text{m}} \left[v_{\text{sat},e} \left(1 - \frac{v_{\text{sat},e}}{d} t \right) \Theta \left(\frac{d}{v_{\text{sat},e}} - t \right) + v_{\text{sat},h} \left(1 - \frac{v_{\text{sat},h}}{d} t \right) \Theta \left(\frac{d}{v_{\text{sat},h}} - t \right) \right]. \quad (3.7)$$

3.4 Readout Electronics

Readout electronics are crucial for bridging the gap between the weak, transient signals generated in a detector and the robust, digital information required for data analysis. For the scope of the MONOLITH project so far, the ASIC's output signal is analog, which is then read out and digitized by a fast oscilloscope outside of the ASIC. This section considers the analog front-end electronics, and the subsequent sections address signal amplification and shaping. Of course, a complete ASIC must directly give as output a digital signal, because this is nowadays the standard to work with in any application, but this is beyond the scope of this thesis.

Signal amplification and shaping must be, at their core, linear signal processing techniques to preserve the shape and timing information of the original signal, which means that a given input pulse $i_{\text{in}}(t)$ results in an output pulse shape $v_{\text{out}}(t)$ which is independent of the signal amplitude. In other terms, $v_{\text{out}}[k i_{\text{in}}(t)] = k v_{\text{out}}[i_{\text{in}}(t)]$ for any value of the constant c , which is practically impossible since the linearity has a specific range of validity. For example, if the input charge is too large, the transistors begin to saturate and cease amplification. We will see an example of a non-linearity effect in Section 5.3.

A powerful tool for electric circuit analysis is the Fourier transform $F(\omega)$ of a signal $f(t)$, defined as:

$$\mathcal{F}[f(t)] = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt = F(\omega), \quad (3.8)$$

where $\omega = 2\pi f$, and f is the frequency. The Fourier transform is generally a complex number, i.e., it has a real and an imaginary part.

The utility of the Fourier transform is evident when considering the output signal of a detector. The input current pulse is $i_{in}(t)$ and the output voltage pulse is $v_{out}(t)$. A single function, called the delta response $h(t)$, can always summarize the entire electronics chain. The output signal $v_{out}(t)$ is related to the input signal $i_{in}(t)$ and the delta response $h(t)$ via the mathematical operation of convolution in the time domain. Defining the Fourier transforms $V_{out}(\omega)$, $I_{in}(\omega)$, and $H(\omega)$ of, respectively, the output signal, the input signal, and the delta response, one finds that a convolution in the time domain is a multiplication in the frequency domain:

$$v_{out}(t) = \int_0^t h(t-t') i_{in}(t') dt' \Rightarrow V_{out}(\omega) = H(\omega) I_{in}(\omega). \quad (3.9)$$

The delta response $h(t)$ is the electronics' response to a Dirac delta $\delta(t)$ as the input signal. The Fourier transform of a Dirac delta $\delta(t)$ is equal to one: $\mathcal{F}[\delta(t)] = 1$, which means that when we feed a Dirac delta $\delta(t)$ to the electronics, we find an output signal in the frequency domain $V_{out}(\omega) = H(\omega) \mathcal{F}[\delta(t)] = H(\omega)$. A common way to characterize the electronic response is to feed the input a Dirac delta function $\delta(t)$, then measure the output signal; this gives the transfer function of the circuit, which is the delta response.

3.4.1 Signal Amplification and Shaping

Signal amplification and shaping involve modifying the temporal profile of the detector signal to optimize resolution and noise rejection. Lastly, it consists of defining a specific delta response. A common approach in modern systems consists of using unipolar signal shaping, which means defining an analytical formulation for the delta response in the time domain and, consequently, in the frequency domain, as follows:

$$h(t) = \left(\frac{t}{t_p}\right)^n e^{n(1-t/t_p)} \Theta(t) \Rightarrow H(\omega) = \frac{t_p e^n n!}{(n + i\omega t_p)^{n+1}}, \quad (3.10)$$

where t_p is called the peaking time of the delta response and is the time corresponding to the highest value of the delta response, n is called the degree of the delta response, and $\Theta(t)$ is the Heaviside function.

A unipolar signal shaping with degree n is realizable by cascading $n + 1$ identical RC filters separated by identical voltage buffers [107]. A standard amplifier with a transistor and a silicon detector as an input capacitance has a delta response approximated with a unipolar shaper. For example, this is the case for the second MONOLITH prototype without a gain layer.

Figure 3.7 shows the delta response for the analog front-end electronics of the second MONOLITH prototype for different detector capacitances. The delta response was obtained through a simulation using the Virtuoso Studio tool by Cadence [108], which involved feeding an input delta response to the analog front-end electronics. The analytical expression from Equation 3.10 with $t_p = 235$ ps and $n = 1$ is also shown: it approximates quite well the shape of the simulated delta response.

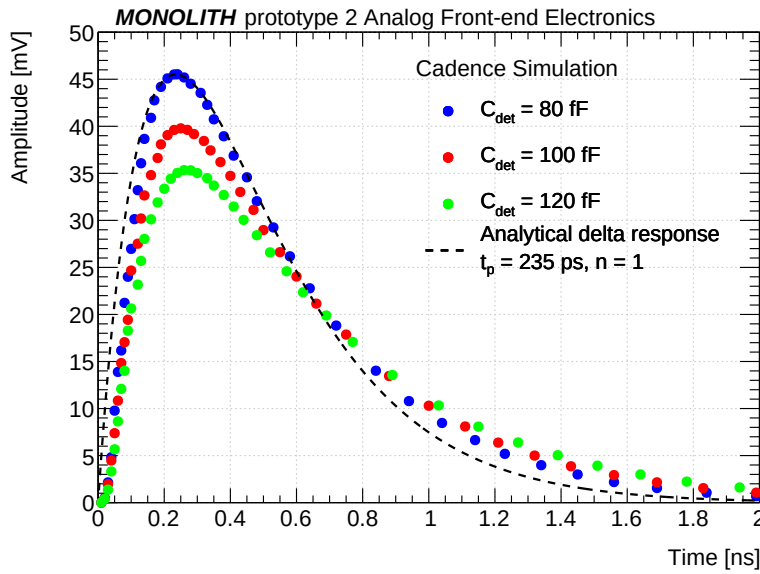


FIGURE 3.7: Delta response in the time domain for the analog front-end electronics of the second MONOLITH prototype with different detector capacitance C_{det} . The delta responses were obtained via a simulation using the Virtuoso Studio tool by Cadence [108] by feeding an input Dirac delta to the analog front-end electronics. The analytical expression from Equation 3.10 with $t_p = 235$ ps and $n = 1$ is also shown: it approximates quite well the shape of the simulated delta response.

The detector capacitance C_{det} for the second MONOLITH prototype without a gain is about 80 fF when the substrate thickness is $d = 50$ μm (see Section 2.2.6). The second MONOLITH prototype with a gain layer has thicknesses varying from $d = 18$ μm to $d = 30$ μm , so the total detector capacitance varies between $\simeq 120$ fF to $\simeq 100$ fF. Figure 3.7 shows how increasing the detector capacitance has the effect of slowing down the delta response, but also decreasing the amplitude. This behaviour corresponds to decreased bandwidth and gain, reducing the signal-to-noise ratio. Therefore, the performance is better from the electronics perspective when the detector capacitance is low.

3.4.2 Noise Sources

Noise is unavoidable in any readout electronics and can severely impact a detection system's sensitivity and resolution. Thus, understanding and reducing noise are crucial aspects of electronics design.

Various sources contribute to the total noise budget. Among the most relevant, we remember:

- the *shot noise* arises from the discrete nature of charge carriers and is inherently linked to the signal current. This noise is *white noise*, meaning that the noise spectrum has no frequency dependency, and all the frequencies have equal noise power.
- The *thermal* (or *Johnson*) *noise* is generated by the random motion of electrons in the resistive elements of the circuit. This noise is also white.
- The *flicker noise* (or *1/f noise*) is related to the condition of non-equilibrium situations, for example, when a bias voltage is applied. It becomes significant at

low frequencies, so it is intrinsically different from the *shot noise* and the *thermal noise*.

- The *amplifier noise* includes contributions from both the input transistor and other circuit elements, further degrading the performance.

Effective design of the readout chain involves minimizing these noise sources through careful selection of low-noise components, optimized biasing conditions, and thoughtful circuit layouts. The goal is to maximize the Signal-to-Noise Ratio (SNR) so that weak signals corresponding to minimal energy depositions are reliably detected and processed. However, the SNR has an intrinsic limit, as derived in [107]. An input signal, i.e., the induced current, with a Fourier transform $F(i\omega)$ and a superimposed noise having a power spectrum $w(\omega)$ is processed by an amplifier with transfer function $H(i\omega)$. The SNR must satisfy:

$$SNR^2 \leq \frac{2}{\pi} \int_0^\infty \frac{|F(i\omega)|^2}{w(\omega)} d\omega. \quad (3.11)$$

The maximum SNR value is reached when the transfer function $H(i\omega)$ satisfies:

$$H(i\omega) = \frac{1}{k} \frac{F^*(i\omega)}{w(\omega)} e^{-i\omega t_m}, \quad (3.12)$$

with an arbitrary scaling constant factor k and an arbitrary time delay t_m . This transfer function represents the optimum transfer function that maximises the SNR. It can be determined completely by the shape of the input signal and the power spectrum of the noise, which is why knowing both the induced current and the noise sources is so relevant.

Recalling Equation 2.37, we see that the equivalent noise charge has a dependence on the total capacitance C_{tot} , which contains the contribution of the detector C_{det} . Therefore, detector capacitance should be minimized to minimize the noise.

3.5 Time Resolution

Achieving good time resolution is one of the primary goals of the MONOLITH project, so it's worth focusing on how time resolution can be improved in silicon detectors. We can combine the crucial contributions that we discussed in the previous sections.

Time resolution refers to the error that occurs when performing a temporal measurement. In silicon detectors with analog output, such as the prototypes for the MONOLITH project, the analog signal resulting from the ionization of the interacting particle is typically digitized by an oscilloscope. The signal is then reconstructed, and the time at which this signal reaches a certain voltage threshold V_{th} is called the time of arrival ToA. The standard deviation of the ToA distribution is referred to as the time resolution of the detector σ_t .

Figure 3.8 shows a collection of analog signals from a CERN SPS testbeam with 120 GeV pions on the second MONOLITH prototype without a gain layer. The signals have been horizontally shifted using a precise time reference.

The distribution of ToA is not always perfectly Gaussian, especially for thin sensors or when some nonuniformities (weighting field or drift velocity) are present. The meaning of the tails will be explained in Section 5.5. However, a Gaussian fit is usually performed on the distribution, excluding a percentage of the tails to

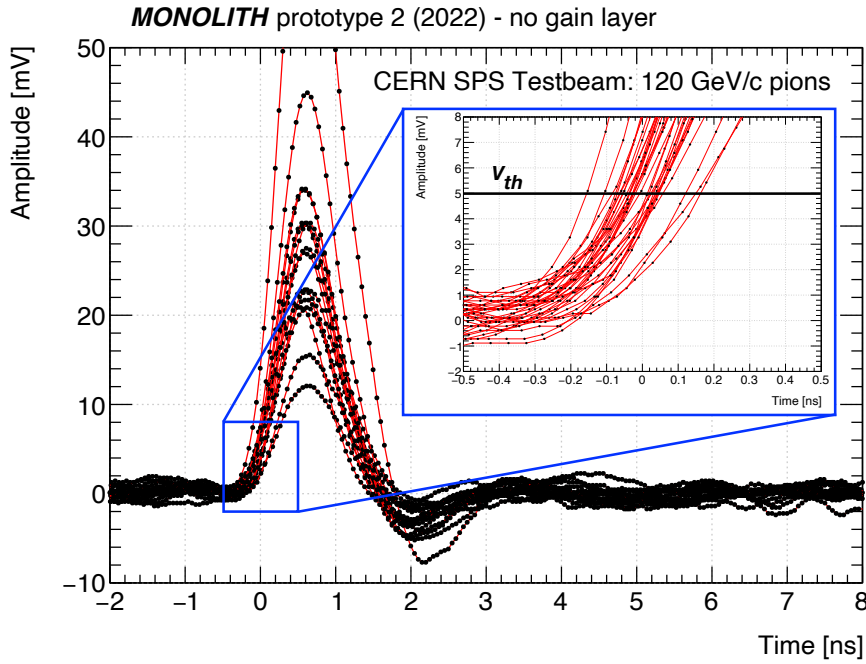


FIGURE 3.8: Collection of analog signals from a CERN SPS testbeam with 120 GeV pions on the second MONOLITH prototype without a gain layer. The insert shows a zoom around the signals' voltage threshold V_{th} .

determine a Gaussian standard deviation, which corresponds to the time resolution of the detector, σ_t .

According to [61], the contributions to the time resolution σ_t can be broadly divide into four categories:

- physics of the interaction, that shapes the energy deposition by the particle, determining the irregularities of the signal σ_{Landau} due to the statistics of the signal formation, and the overall amplitude variations $\sigma_{Time Walk}$. The two contributions must be discussed separately because Landau noise represents an intrinsic limit for the detector, while the time walk can be corrected.
- Electronics, mostly due to the noise and the amplifier slew rate σ_{jitter} .
- Signal distortion, due to non-uniform weighting field and varying charge carrier drift velocity $\sigma_{Distortion}$.
- Digitization, driven by the TDC uncertainties σ_{TDC} .

Considering all the contributions as independent, so considering the correlations as negligible (which is not always the case), the overall time resolution σ_t takes the form:

$$\sigma_t^2 = \sigma_{Time Walk}^2 + \sigma_{Landau}^2 + \sigma_{jitter}^2 + \sigma_{Distortion}^2 + \sigma_{TDC}^2. \quad (3.13)$$

We can now break down Equation 3.13 into the single contributions to the time resolution and analyze them.

3.5.1 Landau Fluctuations Contribution to the Time Resolution

The Landau noise (also known as charge-collection noise) is related to signals originating from the same total charge, but with different charge distributions. These charge fluctuations cannot be eliminated and create a noise σ_{Landau} that contributes to the time resolution σ_t . This contribution to the time resolution represents the final intrinsic limit of the detector.

To understand what is the Landau noise contribution to the time resolution (σ_{Landau}), we focus on a simple case in which the amplifier peaking time t_p is larger than the total duration T of the induced current signal $i_{\text{in}}(t)$, following the discussion given in [101].

First, we define the centroid time τ_{cur} of an induced current signal $i_{\text{in}}(t)$ as:

$$\tau_{\text{cur}} = \frac{\int_0^\infty t i_{\text{in}}(t) dt}{\int_0^\infty i_{\text{in}}(t) dt} = \frac{\int_0^\infty t i_{\text{in}}(t) dt}{Q}. \quad (3.14)$$

In the approximation where the amplifier peaking time is greater than the total duration of the induced current signal, it can be demonstrated that the output voltage signal $v_{\text{out}}(t)$ is the delta response $h(t)$ of the front-end electronics translated by the centroid time τ_{cur} and scaled by the integrated charge Q . Using Equation 3.9, we get:

$$v_{\text{out}}(t) = \int_0^t h(t-t') i_{\text{in}}(t') dt' \simeq Q h(t - \tau_{\text{cur}}). \quad (3.15)$$

Now, we set a threshold for the output signal $v_{\text{out}}(t)$ and examine the spread of the threshold arrival times. Equation 3.15 is crucial, since the shape of the amplifier is well-defined and always equal to the delta response, so it doesn't contribute to fluctuations in the arrival time. Therefore, the spread of the distribution of the arrival time is given by the spread of the distribution of the centroid time τ_{cur} . In other words, the standard deviation of the distribution of the centroid time is the Landau contribution to the time resolution we were looking for: $\sigma_\tau \equiv \sigma_{\text{Landau}}$.

To find the standard deviation of the distribution of the centroid time τ_{cur} , we recall that $\sigma_\tau^2 = \langle \tau_{\text{cur}}^2 \rangle - \langle \tau_{\text{cur}} \rangle^2$. Once the shape of the induced current $i_{\text{in}}(t)$ is defined, the centroid time can be calculated, and, therefore, its standard deviation. We begin with the simple case of a single cluster generated at a depth x within the detector's active layer. From the shape of the induced current of Equation 3.6, and using 3.14, we find that:

$$\tau_{\text{cur, SC}} = \frac{1}{2d} \left[\frac{x^2}{v_{\text{sat,e}}} + \frac{(d-x)^2}{v_{\text{sat,h}}} \right], \quad (3.16)$$

with x being the depth of the single cluster, and d the total thickness of the detector. The electrons move towards the collection electrode at $x = 0 \mu\text{m}$, and the holes move in the opposite direction.

The single cluster can be generated at a random position x uniformly distributed between 0 and d . A certain depth x creates a certain shape of the induced current, with its own centroid time τ_{cur} . The calculation of the standard deviation of the distribution of the centroid time τ_{cur} gives:

$$\sigma_{\tau, \text{SC}}^2 = \frac{4}{180} \frac{d^2}{v_{\text{sat,e}}^2} - \frac{7}{180} \frac{d^2}{v_{\text{sat,e}} v_{\text{sat,h}}} + \frac{4}{180} \frac{d^2}{v_{\text{sat,h}}^2}. \quad (3.17)$$

For the second MONOLITH prototype without a gain layer (the active thickness is $d = 50 \mu\text{m}$), the standard deviation of the centroid time for a single cluster is $\sigma_{\tau, \text{SC}} \simeq 44 \text{ ps}$. This value represents the contribution to the time resolution for single-cluster deposition, such as for ^{55}Fe from ^{109}Cd radioactive sources, due to the unknown depth at which the single cluster is generated.

Moving to a MIP-like charge deposition, the charge clusters profile can be seen as a succession of single clusters, whose charge fluctuates according to the photo-absorption ionization (PAI) theory. Again [101] found that the standard deviation of the distribution of the centroid time τ_{cur} in this case has a relatively simple expression:

$$\sigma_{\text{Landau}} \equiv \sigma_{\tau, \text{MIP}} = w(d) \sigma_{\tau, \text{SC}}, \quad (3.18)$$

$w(d)$ is a universal function of the detector thickness d , completely defined by the type of theory used.

This rather simple expression links the standard deviation for a single cluster with that for an MIP. The universal function $w(d)$ is shown in Figure 3.9 for different theoretical models.

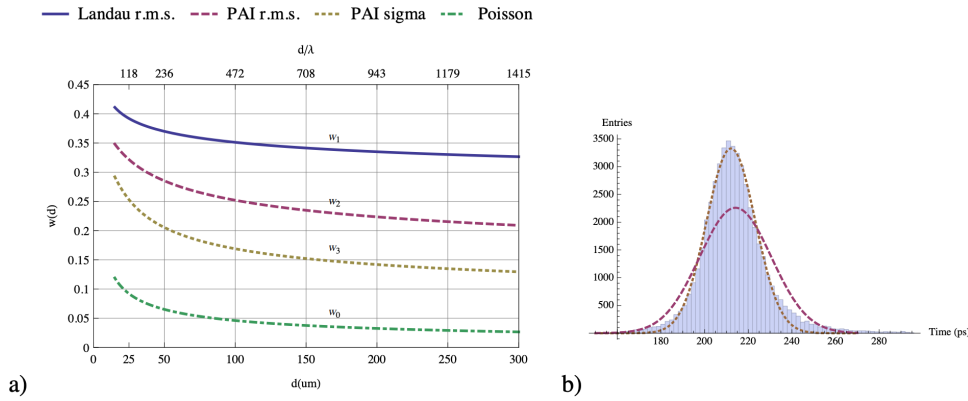


FIGURE 3.9: (Left) Universal function $w(d)$ as a function of the silicon thickness d for different theoretical models: w_1 represents the Landau theory, w_2 the PAI model, w_3 the PAI model with a Gaussian fit instead of taking the r.m.s., w_0 the Poisson statistics. (Right) Histogram of the centroid time $\tau_{\text{cur, MIP}}$ distribution for a detector with $d = 50 \mu\text{m}$ with the PAI model. The dashed curve is a Gaussian with w_2 , and the dotted one with w_3 . Taken from [101].

Figure 3.9 shows that the universal function $w(d)$ is always lower than 1. Hence, the contribution to the time resolution for an MIP is lower than that for a single cluster. For the second MONOLITH prototype without a gain layer ($d = 50 \mu\text{m}$), we can use a value from Figure 3.9 which lies somewhere in between the PAI Gaussian fit w_3 and the PAI r.m.s. value w_2 . We find that the Landau noise contribution to the time resolution for a MIP is $\sigma_{\text{Landau}} \equiv \sigma_{\tau, \text{MIP}} \simeq 10 \text{ ps}$.

3.5.2 Time Walk Contribution to the Time Resolution

When the total primary charge is different, the voltage signals have different heights and time of arrival, which introduces noise into the time resolution. This effect is called time walk, and correcting the signals for it is possible, essentially bringing σ_{TimeWalk} to zero.

To understand the concept of time walk and the necessary corrections, let's consider a straightforward scenario involving signals that start at an initial value of zero and exhibit a linear increase in value. These signals reach their maximum amplitude,

denoted as $\max$, at a specific time denoted as $t_{\max}$. A crucial prerequisite for effective signal correction against time walk is that the time $t_{\max}$ must remain unaffected by the total primary charge. In this instance, the linear rise of the signal is governed by the equation: $V(t) = \frac{\max}{t_{\max}}t$. The time of arrival ToA at a certain threshold V_{th} is therefore $ToA = \frac{V_{th}}{\max}t_{\max}$. Hence, there is an inverse proportionality between the time of arrival ToA and the maximum amplitude $\max$.

Figure 3.10 shows an example of signals with a linear rise to understand the time-walk effect. Figure 3.10 also shows the analytical dependence of the time of arrival ToA on the maximum amplitude of the signals $\max$.

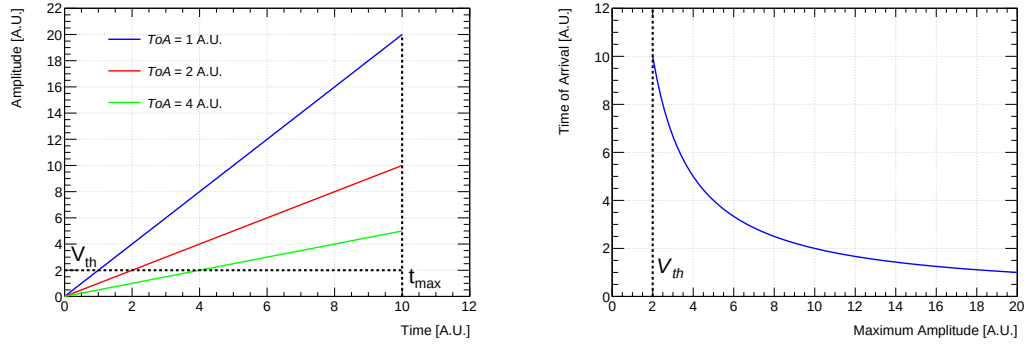


FIGURE 3.10: (Left) Example of three signals with linear rise and the same time $t_{\max}$ to reach the maximum amplitude. The time walk is the effect that makes the ToA for the three signals differ. (Right) Analytical dependence of the arrival time ToA on the maximum amplitude of the signals $\max$. The signal must reach the threshold V_{th} to define a time of arrival ToA.

The process of correcting for the time-walk effect consists in fitting an analytical function to the distribution of ToA versus $\max$ (as the one in Figure 3.10), and then subtract to the ToA of each signal the value of the fitted function corresponding to the maximum amplitude of the signal itself. Another method often used to correct the time walk effect is to take the ToA not at a specific voltage threshold, but instead at a threshold that is a particular fraction of the signal height. Ideally, either correction reduces the time-walk contribution to the time resolution to zero ($\sigma_{\text{Time Walk}} \simeq 0$).

3.5.3 Noise Contribution to the Time Resolution

The noise discussed in Section 3.4.2 creates a voltage noise σ_V on the output signal $V(t)$. When setting a threshold V_{th} and measuring the time of arrival ToA of a signal, this voltage noise smears the ToA measurement, adding a contribution to the time resolution σ_t . This contribution is usually called time jitter, and it is directly related to the voltage noise and the slope of the signal through:

$$\sigma_{\text{jitter}} = \frac{\sigma_V}{dV/dt}. \quad (3.19)$$

To minimize the time jitter σ_{jitter} , it is necessary to reduce the noise σ_V and maximize the slope dV/dt . To reduce noise, one must act on the sources listed in Section 3.4.2 during the design of the detector and amplifier. The main contribution is still due to the detector capacitance C_{det} , so one wants to minimize it. To maximize the slope dV/dt , the amplifier's gain must be high, and the peaking time must be as short as possible. However, there is usually a trade-off between gain and bandwidth.

The slope dV/dt varies at different heights of the signal. In particular, the slope is low when the signal first begins to develop. Most charges have already been collected close to the peak, and the slope is low again. In an intermediate region, the slope is higher. Therefore, a simple trick to minimize time jitter when measuring the time resolution σ_t is to set a threshold V_{th} around the maximum slope, which typically occurs at the midpoint of the rise.

3.5.4 Other Contributions to the Time Resolution

The remaining contributions are the effects of signal distortion, due to non-uniform weighting field and charge carrier drift velocity $\sigma_{\text{Distortion}}$, and the effects of the digitization σ_{TDC} .

As already mentioned in Section 3.3, the ratio w/d must be maximized to make the weighting field and the drift velocity constant, and the thickness d must be minimized to achieve a high weighting field. Moreover, the electric field must be high enough to saturate the drift velocity everywhere. When these conditions are met, the detector's geometry is approximated with a parallel plate geometry, and the contribution to the time resolution $\sigma_{\text{Distortion}}$ becomes negligible.

In a real case situation, the ratio w/d is limited by the constraints on the detector capacitance, hence $\sigma_{\text{Distortion}}$ is not always negligible. For the MONOLITH prototype 2 ASIC, the ratio w/d has a value of $w/d \simeq 100 \mu\text{m}/50 \mu\text{m} = 2$. Considering the specified thickness, the ratio is even higher for the prototypes with the PicoAD sensor. According to [101], when $w/d \gtrsim 1$, the distortion introduced by the weighting field is negligible.

The other contribution is related to the effect of the digitization σ_{TDC} . When testing the prototypes for the MONOLITH project, digitization was performed using an oscilloscope with a sampling rate of 40 GS/s. This means a time binning Δ_{bin} of 25 ps. Assuming that the error on the time binning is related to the uniform probability of sampling within a certain time bin, we have $\sigma_{\text{TDC}} = \Delta_{\text{bin}}/\sqrt{12} \simeq 7$ ps. These values are not negligible when trying to measure a time resolution varying between 10 ps and 20 ps, but with a simple linear interpolation between consecutive points, the contribution drops down to $\simeq 3.5$ ps. Hence, σ_{TDC} starts to be negligible.

3.6 Spatial Resolution

Another primary goal of the MONOLITH project is achieving good spatial resolution, which is crucial for precisely tracking charged particles in modern high-energy physics experiments. Spatial resolution refers to the detector's ability to pinpoint the exact position where a particle crosses the sensor. This capability directly affects the precision with which particle trajectories, momenta, and decay vertices can be reconstructed.

Several factors influence the spatial resolution in silicon detectors, including the pitch of the sensor pixels, the sensor thickness, and the charge sharing between adjacent electrodes. Fine segmentation, achieved by reducing the pitch, enables a more granular readout and, consequently, a more accurate localization of the particle hit. Additionally, thinner sensors can reduce charge diffusion, which helps confine the signal to fewer channels. However, some degree of charge sharing is beneficial, as it enables interpolation techniques that improve resolution beyond the binary limit defined by the pitch.

As discussed in the previous sections, the thickness of the prototypes for the MONOLITH project is thin (below $50\text{ }\mu\text{m}$), so the diffusion of charge is negligible. Moreover, despite charge sharing between adjacent electrodes being beneficial for spatial resolution, it degrades time resolution, as the jitter term, which depends directly on the amount of primary charge in a charge-sensitive amplifier, deteriorates for less charge. Therefore, the charge sharing must be reduced as much as possible. Finally, the dominant contribution to the spatial resolution of the prototypes for the MONOLITH project is the pitch of the sensor pixels.

Neglecting diffusion from the thick sensor and charge sharing, the spatial resolution depends solely on the pixel pitch. To quantify the spatial resolution of a detector with known dimensions, one must think of particles that cross the detector with a 2D uniform distribution of their coordinates. From that, it's possible to define two normalized *Probability Density Functions* (PDF) $f(x)$ and $f(y)$, one for each axis, which give the probability of finding a particle at a specific position.

Once the PDF $f(x)$ is known for a specific x -axis, the mean μ_x and the variance σ_x^2 for the x -axis are defined by:

$$\mu_x = \int_{-\infty}^{\infty} x f(x) dx; \quad (3.20)$$

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x - \mu_x)^2 f(x) dx. \quad (3.21)$$

The spatial resolution of the detector in the dimension associated with the x -axis is quantified as the standard deviation σ_x .

3.6.1 Rectangular Pixel

Let's calculate the spatial resolutions σ_x and σ_y in the two dimensions x and y for the simple case of rectangular pixels with base dimension b and height dimension h . Figure 3.11 shows a set of four rectangular pixels, but the detector may extend for multiple rows and columns other than these four pixels.

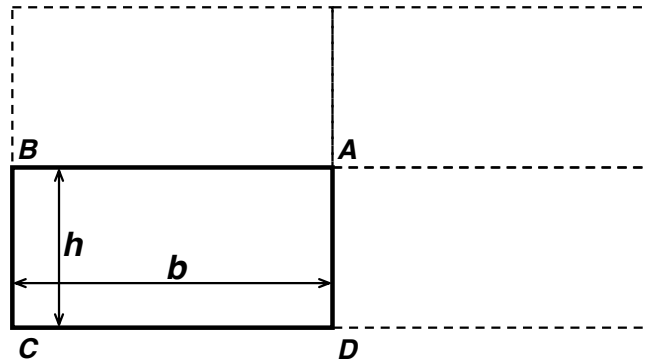


FIGURE 3.11: Detector made of four rectangular pixels with base b and height h . Each pixel is a sensitive element and gives an electric signal when it detects the passage of a particle.

Another assumption that we are making is that the particles are detected with a 100 % efficiency, which means that there is no dead area, especially in the inter-pixel regions. With this assumption and a choice of a Cartesian coordinate system with the x -axis parallel to the base and the y -axis parallel to the height, the spatial distribution of the events is uniform in the x - y plane across all pixels. Therefore, the PDF for each axis is a uniform distribution with constant height spanning over the detector

dimensions. Starting from Equation 3.20 with a uniform PDF, one finds that the expectation value of the PDF on the x -axis is $\mu_x = b/2$, and the standard deviation, i.e., the spatial resolution on the x -axis, is $\sigma_x = b/\sqrt{12}$.

3.6.2 Hexagonal Pixel

The pixels of the prototypes for the MONOLITH project are not rectangular, but they are hexagonal. This shape was chosen because the 120° corner angles produce a much more uniform electric-field distribution than the 90° corners of square pixels. In a square pixel, sharp right-angle corners concentrate electric field lines, creating localized regions of high field interspersed with broader, low-field "pockets" that slow charge drift and degrade timing uniformity. By contrast, the obtuse angles of a regular hexagon yield gentler field gradients at every boundary, minimizing these low-field zones and ensuring that charge carriers experience a consistently strong drift field regardless of where they are generated within the pixel. This geometric benefit directly translates into faster and more uniform charge collection, reducing time jitter and improving the detector's spatial and temporal resolution.

It is possible to repeat the same steps from the previous section for a rectangular pixel, but for a hexagonal pixel, to determine its spatial resolution. Figure 3.12 shows a set of four hexagonal pixels, with a characteristic dimension represented by the side s . The detector may extend beyond these four pixels in multiple rows and columns.

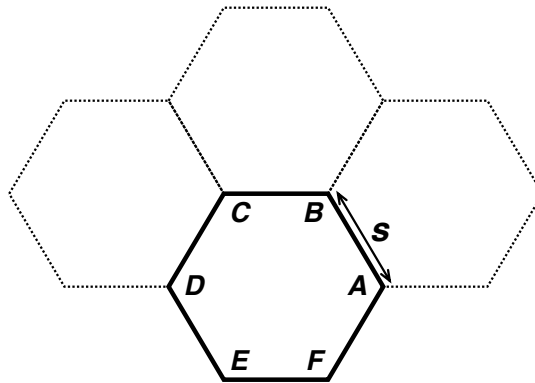


FIGURE 3.12: Detector made of four regular-hexagonal pixels with side s . Each pixel is a sensitive element and gives an electric signal when it detects the passage of a particle.

Again, assuming a 100 % efficiency, even in the inter-pixel regions, the spatial distribution of the events is uniform in the x - y plane across the entire hexagonal pixel. We define a simple Cartesian reference system associated with the hexagon, originating in the center of the pixel and the x -axis along the direction dictated by the segment DA of Figure 3.12.

Pure geometrical arguments make it possible to find the shape of the probability density functions along the x and y axes. Figure 3.13 shows the PDFs on the x and y -axis.

These piecewise functions are defined by the constants K and K' , which determine the absolute height. The values of the constants are arbitrary and do not change the geometrical arguments. They are fixed by normalizing the PDFs. After having found the mathematical expressions for the PDFs, it is possible to use Equation 3.20 to find that $\mu_x = 0$ and $\mu_y = 0$. This result is a trivial confirmation that the Cartesian reference system has the origin coinciding with the center of the hexagon, as we fixed.

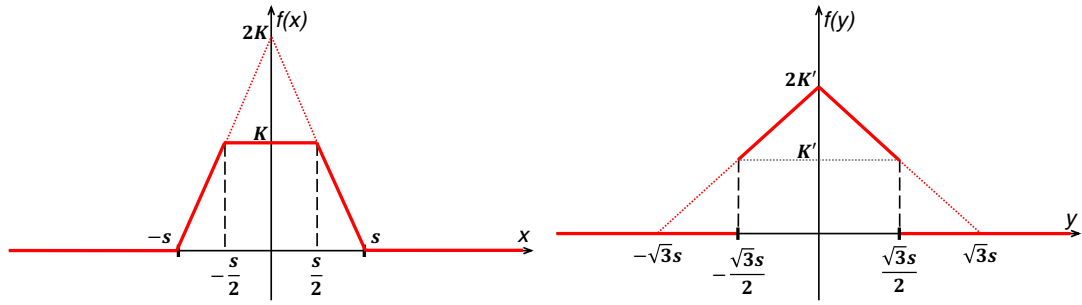


FIGURE 3.13: In solid red line, the probability density function for a regular hexagonal pixel with side s . (Left) On the x -axis. (Right) On the y -axis. The values of the constants K and K' are fixed, ensuring the probability density function is normalized.

Finally, it is possible to calculate the standard deviation of the PDFs on each axis, finding:

$$\sigma_x^2 = \sigma_y^2 = \frac{5s^2}{24} \Rightarrow \sigma_x = \sigma_y = \sqrt{\frac{5}{24}}s \quad (3.22)$$

The PDF is a uniform distribution for a rectangular pixel on the x and y axes. The PDFs on the x and y axes have different shapes for a hexagonal pixel. Despite that, the standard deviation on each axis is the same and amounts to $\sigma_x = \sigma_y = \sqrt{\frac{5}{24}}s$.

3.6.3 Comparison between Squared and Hexagonal Pixels

We can compare the two geometries to see if one offers better spatial resolution. It is reasonable to ask which shape has better spatial resolution when the area is equivalent. Therefore, we consider a square pixel with a side of l , whose area is the same as that of a regular hexagon with a side of s . The spatial resolution for a square pixel is the same on the two principal axes and is calculated as done for a rectangular pixel: $\sigma_x = \sigma_y = \frac{l}{\sqrt{12}}$.

A hexagon with side s has an area of $A_{\text{hexagon}} = \frac{3\sqrt{3}s^2}{2}$. A square with side l has an area of $A_{\text{square}} = l^2$. The two surfaces are equal when the following condition between the hexagon side s and the square side l holds:

$$l = \sqrt{\frac{3\sqrt{3}}{2}}s. \quad (3.23)$$

Considering the x -axis as a reference, it is possible to calculate the ratio between the two spatial resolutions of a hexagon and a square of the same area. Using the Equation 3.23, it holds:

$$\frac{\sigma_{\text{hexagon}}}{\sigma_{\text{square}}} = \frac{\sqrt{\frac{5}{24}}s}{\frac{l}{\sqrt{12}}} = \sqrt{\frac{5}{3\sqrt{3}}} \simeq 0.98. \quad (3.24)$$

This means that $\sigma_{\text{hexagon}} < \sigma_{\text{square}}$. Thus, when considering the same area, the spatial resolution of a hexagonal pixel is intrinsically better than that of a square pixel.

Although the spatial resolution of a hexagonal pixel is intrinsically better than that of a square pixel with the same area, there is a simple drawback to consider: the fill factor. Tiling a finite rectangular plane with pixels, as is done when building a

particle detector made of pixels, leads to a full fill factor when the pixels are squares, but results in some dead area when the pixels are regular hexagons. In this last case, the geometrical efficiency is lower. Figure 3.14 shows the different tiling between squares and hexagons.

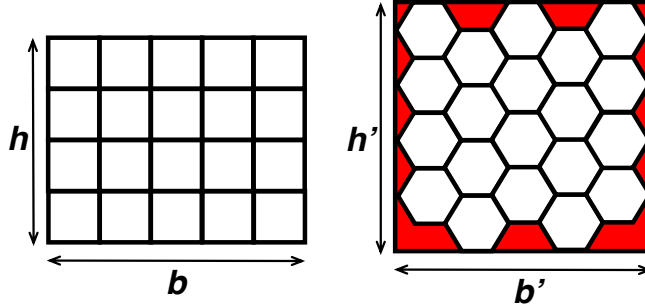


FIGURE 3.14: Comparison between tiling a finite plane with four rows and five columns of squared pixels (left) or hexagonal pixels (right). On the right, the red area represents the dead zone, which is insensitive to particles passing through it.

The comparison between squares and hexagons in Figure 3.14 shows how to tile a finite plane with base b and height h . It can be tiled with $m \times n$ squared pixels of side l , with $m = \frac{b}{l}$ and $n = \frac{h}{l}$. If b and h are such that m and n are integer numbers, then the squared pixels cover the plane exactly. In this case, the fill factor is 100 %.

On the other hand, tiling a finite plane with base b' and height h' with regular hexagonal pixels with side s requires $m' \times n'$ pixels. If b' and h' are such that m' and n' are integer numbers greater than or equal to two, it can be found that:

$$n' = \frac{h'}{\sqrt{3}s} - \frac{1}{2} \text{ and } m' = \frac{2b'}{3s} - \frac{1}{3}. \quad (3.25)$$

Even in the case of m' and n' integer numbers greater than or equal to two, the hexagonal pixels do not tile the plane perfectly. Still, some dead areas are closer to the corner, as visible from the red area of Figure 3.14. In this case, it can be found that the fill factor is:

$$FF = \frac{A_{active}}{A_{total}} = \frac{m'n'A_{hexagon}}{b'h'} = \frac{1}{1 + \frac{1}{3m'} + \frac{1}{6n'} + \frac{1}{6m'n'}}. \quad (3.26)$$

Assuming the same number of rows and columns, one gets $m' = n' \equiv N$. Figure 3.15 shows the evolution of the fill-factor as a function of the number N of rows and columns of regular hexagonal pixels when tiling a finite plane.

Figure 3.15 shows that a number N of rows and columns of regular hexagons greater than 9 is already sufficient to have a fill-factor greater than 95 %. To reach a fill-factor of 99 %, the number N of rows and columns of regular hexagons must be at least 55. In the case of the prototypes for the MONOLITH projects, the ASICs are made of matrices of 12 rows times 12 columns of regular hexagonal pixels. Therefore, the fill-factor is close to 96 %.

An important observation to add is that the dead area of the detector is concentrated along the border of the matrix. Hence, its presence is not a problem for particle detection. In any ASIC, a dead area close to the border is always present, typically for bias lines and guard rings. Moreover, for monolithic detectors, this dead area can be used to host some readout electronics. The one within the inter-pixel regions is the important fill-factor to consider to avoid a drop in efficiency. As demonstrated by the

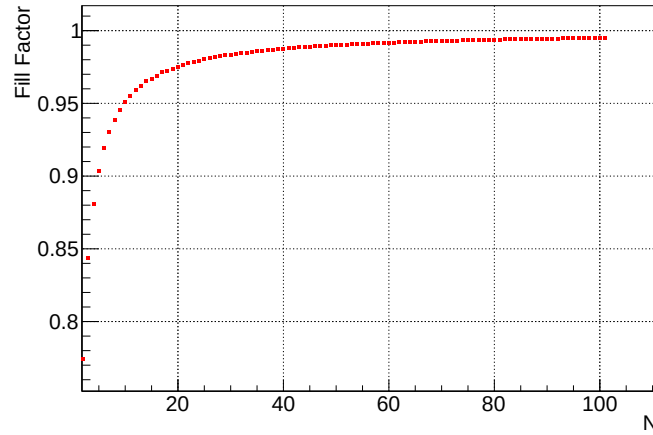


FIGURE 3.15: Evolution of the fill factor as a function of the number N of rows and columns of regular hexagonal pixels needed to tile a plane with the constraints of Equation 3.25.

efficiency measurements in the following sections, the fill factor for the MONOLITH prototypes is close to 100 % everywhere, including in the inter-pixel regions.

3.7 Power Consumption

Among the multitude of challenges that the design of silicon detectors for future colliders must address, power consumption is undoubtedly a critical parameter. As the size and complexity of detector systems increase, so does the cumulative power required to operate the vast arrays of front-end electronics integrated with high-resolution sensors. Hence, limiting the power consumption of each single detector is essential to ensure the feasibility of the detector infrastructure and enable efficient and reliable thermal management across the system.

One of the primary motivations for reducing power consumption is simplifying the cooling system. Every milliwatt of dissipated power must be effectively removed to prevent thermal runaway, preserve performance, and avoid damage to sensitive components. In high-granularity silicon detectors, where millions of pixels or strips are closely packed, the local heat density can become significant, especially if each channel includes active amplification or signal processing electronics. High power densities necessitate complex cooling strategies, such as active liquid cooling with precise flow control or advanced microchannel systems. These technologies, while effective, introduce material budgets, mechanical constraints, and potential failure points that are undesirable in a precision tracking system.

Moreover, maintaining low operating temperatures is particularly important in radiation environments, where the leakage current of silicon sensors increases with dose and temperature. A lower power budget reduces the thermal load, thereby alleviating pressure on the cooling system and making it easier to sustain low detector temperatures, suppress leakage currents, and prolong the sensor's operational lifetime. From a system-level perspective, reduced power requirements also ease the constraints on power delivery infrastructure, decreasing the need for bulky cables and power converters, which in turn leads to a lower material budget, which is critical for minimizing multiple scattering and improving tracking performance.

In summary, maintaining low power consumption in silicon detectors is a foundational requirement for their deployment in large-scale, high-precision experiments,

such as future colliders. It impacts thermal management, radiation tolerance, mechanical design, and achievable tracking resolution.

In the case of the MONOLITH prototypes, the amplification stage dictates the power consumption. The driver stages contribute to power consumption, but they are necessary because analog signals are required for proper operation. In a future implementation of the ASIC with only digital outputs, the driver stages will not be present; thus, they must not be considered for a fair comparison with other technologies. Moreover, the rest of the digital electronics make a negligible contribution to power consumption, as their consumption is instantaneous, whereas the amplification stage has continuous power consumption.

Figure 3.16 shows a zoom on the amplification stage of the MONOLITH prototypes, which is the same among the three prototypes presented in this thesis.

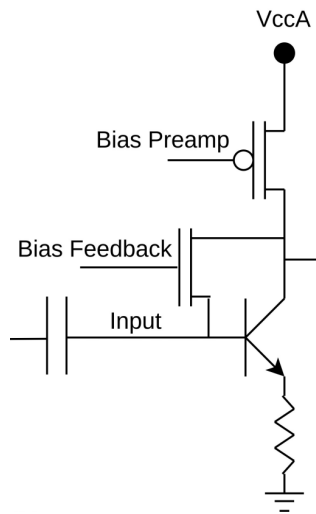


FIGURE 3.16: Zoom of the amplification stage of the MONOLITH prototypes. Two MOSFETs provide the bias feedback current and the bias preamplifier current, enabling the correct polarization of the HBT in common-emitter mode, which is used as an amplifier. The capacitance at the input of the HBT decouples the detector.

The two MOSFETs shown in Figure 3.16 give the bias feedback current I_{fbk} and the bias preamplifier current I_{preamp} to correctly polarize the HBT in common-emitter mode used as an amplifier. The preamplifier current I_{preamp} is responsible for supplying the bias current to the collector of the HBT and setting the gain-bandwidth product of the HBT. On the other hand, the feedback current I_{fbk} is responsible for supplying the bias current to the base of the HBT and for setting the working point in the gain-bandwidth curve determined by the value of I_{preamp} . Thus, it changes the preamplifier's negative feedback, regulating its gain and speed. These two bias currents entirely define the power consumption of the ASIC.

The power consumption density of the analog front-end electronics is given by:

$$P_{\text{density}} = (I_{\text{preamp}} + I_{\text{fbk}}) \cdot V_{\text{CCA}} \cdot N_{\text{pixel}}, \quad (3.27)$$

where V_{CCA} is the analog voltage given to the preamplifier and N_{pixel} is the number of pixels per square cm.

For example, for the second MONOLITH prototype without a gain layer, typically $V_{\text{CCA}} = 1.8 \text{ V}$, $I_{\text{preamp}} = 150 \mu\text{A}$, and I_{fbk} is less than $1 \mu\text{A}$, thus negligible compared to I_{preamp} . We know the geometric dimensions of the hexagonal pixel; hence, for these values, we can calculate a power density of 2.7 W/cm^2 .

Part II

Characterization of the Prototypes

Chapter 4

PicoAD Proof-of-Concept Prototype

This chapter presents the combined approach of detailed TCAD simulations, laboratory measurements, and test-beam studies used to validate and optimize the silicon sensor design of the PicoAD proof-of-concept prototype. The ASIC design and the gain layer layout were discussed in Section 1.2.5.

Section 4.1 describes the Technology-Computer Aided Design (TCAD) framework used to model device electrostatics, charge transport, and signal formation for the PicoAD proof-of-concept prototype. Synopsys Sentaurus was used to simulate doping profiles, depletion regions, electric-field distributions, and weighting potentials for the PicoAD pixel geometry.

Section 4.2 covers laboratory measurements performed before beam tests. In particular, we detail probe station techniques for current–voltage (I–V) scans to gain insight into the process’s operation. These measurements determine leakage currents, the depletion voltage of the gain layer, and breakdown voltages. Then, we describe energy-calibration runs with a ^{55}Fe radioactive source, which emits 5.9 keV X-rays that generate well-known charge peaks in silicon. The resulting spectra are analyzed to extract the electronic noise and gain of the analog front-end chain.

Section 4.3 presents the test-beam campaign, where sensors are exposed to minimum-ionizing particles at a CERN SPS test beam. A high-resolution silicon-pixel telescope provides precise track references, while a fast-sampling oscilloscope reads the Device Under Test (DUT). The DUT signal amplitudes are analyzed to extract the charge-deposition distribution and to determine the dependence on the track position. Then, the efficiency is presented as a function of bias voltage and threshold, focusing on variations in efficiency across the pixel cell and revealing edge and corner effects. We study time-resolution performance by measuring the threshold crossing time relative to a fast reference. We then discuss the details of applying time-walk corrections to improve the temporal precision. Finally, we synthesize these results, demonstrating that the PicoAD proof-of-concept prototype meets the timing and efficiency requirements for fast-timing applications. However, some improvements are still possible.

Some of the results presented in this chapter have been published previously. In particular, the TCAD simulations and laboratory characterization can be found in [1]. The test beam results are presented in [2] and the gain measurements using a ^{55}Fe radioactive source are reported in [3].

4.1 TCAD Simulations

A study was conducted to examine the expected behavior of the PicoAD proof-of-concept prototypes using Technology Computer-Aided Design (TCAD) simulations with the Synopsis Sentaurus tool [82]. The results refer to a 3D geometry model that consists of two adjacent pixels, each with a volume of one-quarter of a full hexagonal pixel. In this way, the inter-pixel region is fully characterized. The simple geometry is then mirrored twice to have one full hexagonal pixel. Figure 4.1 shows the 3D model of the simple geometry.

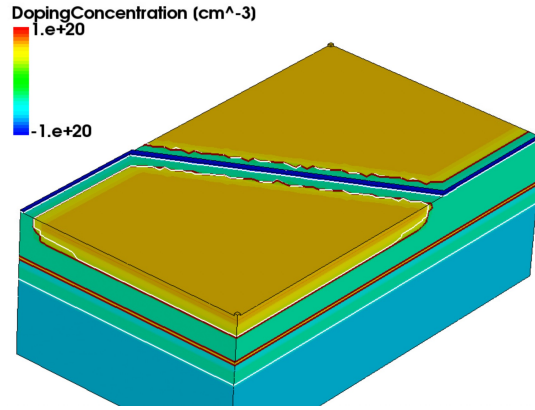


FIGURE 4.1: TCAD 3D model geometry used to simulate the PicoAD gain with a fully depleted sensor. The volume simulated corresponds to one-quarter of two adjacent hexagonal pixels.

A quasi-stationary ramp-up of the sensor bias simulated the electric field in the sensor after it had fully depleted both the epitaxial and gain layers. The bottom of Figure 4.2 shows the expected electric field intensity for a fully depleted PicoAD sensor at the highest implantation dose of the gain layer (dose 4). The top of Figure 4.2 shows the position of the 1D cuts used to study the electric field.

At complete depletion, the maximum value of the electric field is observed in the gain layer, with a slight drop in the inter-pixel region due to the limited thickness of the drift region and the presence of the p^+ -implant (p -stop) used to stop the formation of a conductive channel under the shallow trench isolation. However, the field in the drift region, despite showing a gradient due to the relatively low resistivity of the epitaxial layer, is high enough to guarantee saturation of the electron drift velocity over a significant fraction of the pixel cell while, as is the case for this prototype, it is close to saturation even in the region under the p -stop.

Transient simulations were performed to estimate the gain value of avalanche multiplication inside the sensor at a bias voltage of 2.5 V below the breakdown voltage of the gain layer. For this study, a primary charge of 63 electrons, corresponding to the most probable charge deposited by a MIP in one micrometer of silicon, was deposited on segments with length 1 μm at different depths within the depletion region, and the gain was evaluated as the ratio between the collected charge and the primary charge. Figure 4.3 shows the results of the simulation at the center and the edge of the pixel (positions C4 and C1 of Figure 4.2, respectively).

According to design, the PicoAD gain layer amplifies the signal of electrons produced in the absorption region, known as electron gain. At the same time, it makes little to no gain when the primary charge is generated in the drift region, known as hole gain. It must be noted that the slight differences in the electric field between the edge and the center of the pixel observed in this prototype (see Figure 4.2)

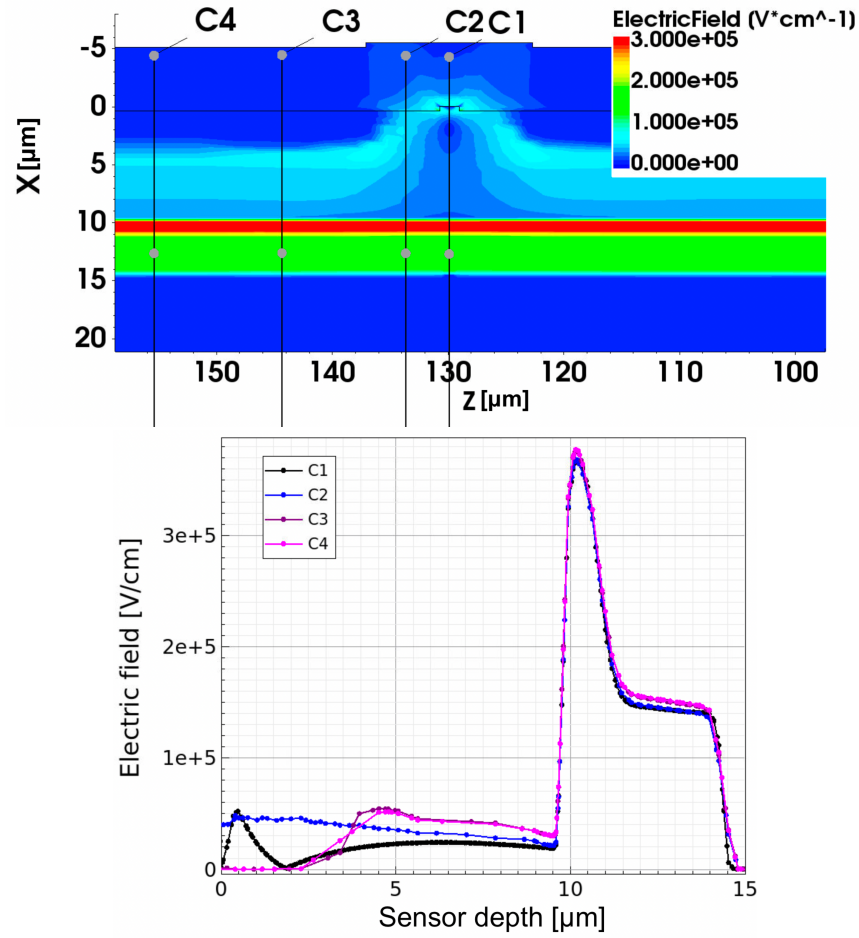


FIGURE 4.2: (Top) 2D cut of the intensity of the electric field in the sensor volume after full depletion at a bias voltage of 2.5 V before breakdown. The vertical lines C1, C2, C3, and C4 represent the regions used for the 1D cuts used for the study of the electric field: C1 is at the center of the inter-pixel region; C2 at the edge of the collection electrode; C3 is close to the side of the collection electrode; C4 is at the center of the collection electrode. (Bottom) 1D cut of the modulus of the electric field in the four regions of the sensor volume.

produce a significant difference in gain. The impact of these electric-field variations is more severe when the sensor is operated close to breakdown. Despite this, the simulation confirms the presence of an intense electric field in the gain layer under the pixel and in the inter-pixel region. An intense electric field in the inter-pixel region is crucial to achieving a 100 % gain in fill factor. Reducing the ratio between the inter-pixel distance and the epitaxial layer thickness above the gain layer in future PicoAD prototypes will allow for better uniformity of the sensor gain.

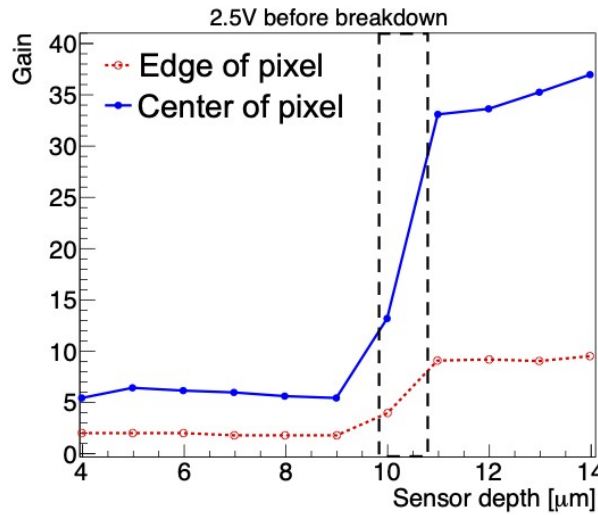


FIGURE 4.3: Simulation of the gain produced by 63 electrons in a dose 4 PicoAD proof-of-concept prototype as a function of the initial position of the primary charge. The simulation is performed for 1 μm segments of a MIP in two of the cross sections reported in Figure 4.2: C1, edge of the pixel (red dotted line); C4, center of the pixel (blue line). In this prototype, the gain layer is placed at a depth of 10 μm . The black dashed lines indicate the high field region produced by the gain layer. For a primary electron-hole pair positioned between 0 and 10 μm , the avalanche is initiated by holes. Between 10 μm and 14 μm , electrons initiate the avalanche.

4.2 Laboratory Measurements

This section presents laboratory measurements of the PicoAD proof-of-concept prototype. The measurements were conducted in the clean room laboratory at the University of Geneva from the second half of 2021 to the first half of 2022 to determine if the gain layer works as expected, by comparing laboratory results with simulations.

First, we describe the current–voltage scan (IV scan) of the PicoAD proof-of-concept prototype, which provides the values of depletion voltage of the gain layer, leakage current, and breakdown voltage of the detector. The IV curves are fundamental for gaining insights into the process’s operation, such as checking resistivity and the implanted dose. Additionally, the IV scans can clarify whether the overall design of the sensor works as expected, particularly the guard rings and the gain layer mask, which prevent current injection from the sides of the pixel matrix.

Second, we describe energy-calibration runs using a ^{55}Fe radioactive source. The source generates well-known charge peaks in silicon, which are used to estimate the gain provided by the gain layer for both electron-initiated and hole-initiated avalanches. The PicoAD sensor concept is fascinating because it allows for the simultaneous study of hole and electron gain multiplication on a single sensor. This feature is usually impossible with other technologies that rely on multiplying only a type of charge carrier.

4.2.1 Probe Station Measurements

The IV characteristics of the four gain-layer doping doses in the PicoAD proof-of-concept prototype were measured in the University of Geneva’s clean rooms. The

ASIC was contacted using the probe station *Cascade Microtech CM300* [109] to the power device analyzer *Keysight B1505A* [110].

The substrate was connected to a negative bias voltage HV for the measurements, while the deep n -wells that formed the pixel matrix and housed the electronics were grounded. The probe station measured four independent currents, using four independent needles, each connected to its module on the device analyzer. The chosen currents were: the total current, called the substrate current; the current entering a sub-matrix of pixels, called the pixel current; the current to the electronic n -wells, called power current; the innermost guard ring current, called Guard Ring current. Figure 4.4 shows the location on the ASIC surface of the different n -wells with the associated color code that will be used below.

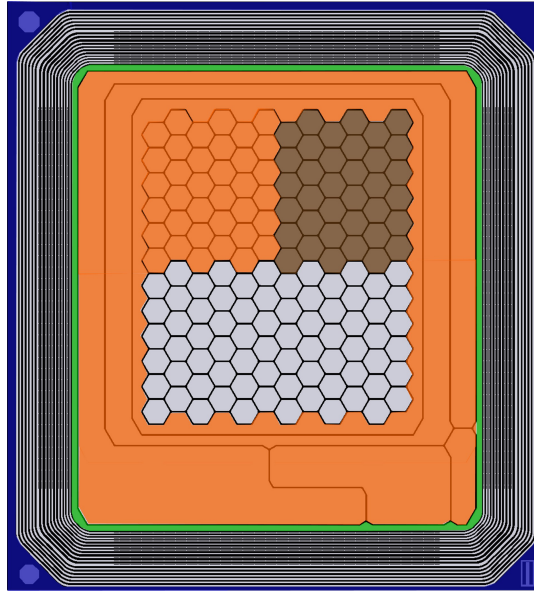


FIGURE 4.4: Regions of the chip probed in the IV scans. Each color corresponds to a different region for which the current was measured independently. The brown region is one of the four sub-matrices. The orange region is a deep n -well that houses the electronics of the four active analog pixels, the chip's periphery, and another sub-matrix. The green region probes the current collected by the innermost guard ring. The blue region represents the substrate, which is accessible outside the guard rings on the top surface and the backside connection. The pixels in grey and the guard rings were left floating.

Figure 4.5 (left) shows the substrate current as a function of the sensor bias for a prototype chip with a dose one implant of the gain layer when a negative high voltage from 0 to 90 V was applied and subsequently ramped down. The difference between ramp-up and ramp-down currents is displayed in Figure 4.5, right.

The sensor breakdown voltage was measured to be 185 V and thus is not visible in Figure 4.5. The substrate current deviates from a typical monotonic rise, showing a steep increase at 10 V during ramp-up, followed by a reduction and a return to the original trend line at 30 V. This behavior is not present when the voltage is ramped back down.

The excess current between 10 and 30 V, which TCAD quasi-stationary simulations cannot reproduce, can be explained by free carriers inside the junction formed by the gain layer. Indeed, as soon as the high voltage is applied, the depletion layer begins to extend downward from the pixel junction into the epitaxial layer. During this phase, the epitaxial layer above and below the n -type gain layer has a negative

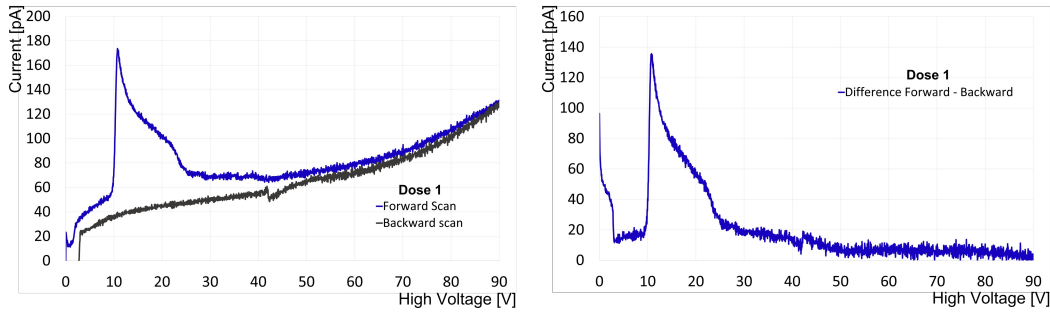


FIGURE 4.5: (Left) Substrate current as a function of the sensor bias voltage applied to the substrate of a PicoAD sensor with gain-layer implant dose 1. In blue, the current is measured during the ramp-up of the high voltage; in black, the current is measured during the ramp-down. The reason for the current peak visible between 10 V and 30 V during ramp up and not observed during ramp down of the high voltage is given in the text. Each point is obtained by averaging the current for 200 ms. (Right) The current difference between the ramp-up and ramp-down of the sensor bias voltage.

high voltage, isolating the gain implant and trapping the free charges within the non-depleted n -type gain layer. Removing these trapped charges produces the measured non-stationary current that disappears when the sensor is fully depleted. The magnitude of the extra current depends on the speed of the high-voltage ramp since removing a fixed amount of charge produces it. The measured excess current confirms that the TCAD simulations suggest how the double p - n junctions are depleted, starting from the pixel junction and isolating the gain layer, before depleting the gain layer.

Figure 4.6 shows the IV measurement performed on a sample with the highest gain-layer implant dose (dose 4) from the same wafer.

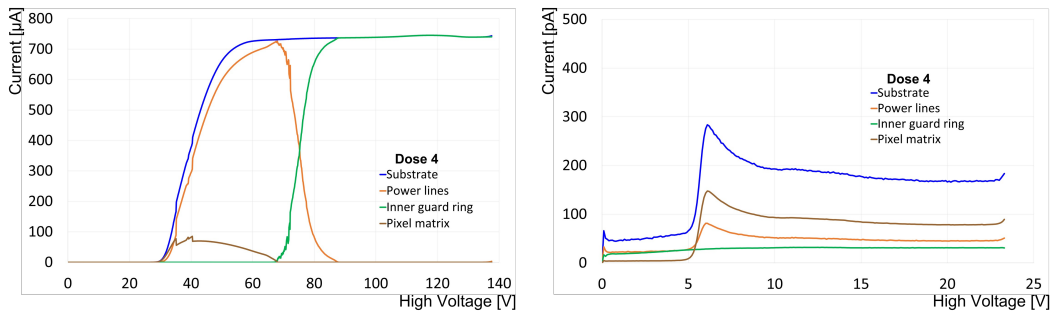


FIGURE 4.6: IV measurement of a PicoAD proof-of-concept prototype with a dose four implant of gain-layer. In blue: the total current collected by the substrate. In brown: the current collected by the pixel terminals. In orange: the current collected by the power lines. In green: the current collected by the innermost guard ring. (Left) A full scan shows how the guard ring collects the extra current when the sensor is fully depleted. (Right) Zoom before the onset of the extra current observed beyond 25 V. The effect of the depletion of the gain layer under the pixels and power lines is visible above 5 V.

In this case, a large extra current was visible after reaching the start of the depletion layer. The origin of this additional current is attributed to the region outside the guard ring, possibly near the chip edge. TCAD simulations suggest that this current could be injected into the pixels and power lines via the undepleted gain layer, which

is the case of PicoAD sensors with the gain layer extending under the guard ring. This extra current was primarily observed in samples with larger gain-layer implant doses and an early onset of floating n -layer depletion. The presence of local defects in silicon, discussed in Section 1.2.4, could contribute to activating this extra current. Figure 4.6 left also shows that, after the complete depletion of the gain layer, all the excess current is collected by the guard ring, and therefore, this current has no influence on the sensor performance. This effect can be avoided by removing the part of the gain-layer implant underneath the guard ring.

4.2.2 Measurements with a ^{55}Fe Radioactive Source

The gain response of the PicoAD proof-of-concept prototype was studied using a ^{55}Fe X-ray source. The measurements were performed in a climate chamber at an ambient temperature between -20°C and 20°C . The output of the analog pixels was sent to an oscilloscope with a 1 GHz analog bandwidth, and the waveforms were saved for offline analysis.

The ^{55}Fe radioactive source emits X-ray photons with an energy of 5.9 keV, which interact through the photoelectric effect in silicon in a point-like type of interaction (see Section 3.2.2). Each photon, if absorbed, produces a charge cluster inside the sensor that can be approximated as a point-like charge deposition of $n_i = 1640$ initial electron-hole pairs.

According to where the photon is absorbed, avalanche multiplication can be initiated by electrons or holes, resulting in different gain values. If the cluster is generated inside the absorption region, i.e., below the gain layer, the electrons drifting toward the collection electrode produce avalanche gain. On the other hand, if the cluster is generated inside the drift region, i.e., above the gain layer, the holes drifting towards the substrate produce avalanche gain. The two possibilities correspond to generating a total charge of $n_{\text{abs.}} = G_e * n_i$ or $n_{\text{drift}} = G_h * n_i$. The electron gain, named G_e , and the hole gain, named G_h , are defined as the total charge divided by the primary charge, according to the carriers that initiated the avalanche mechanisms. The impact ionization coefficient of holes and electrons is expected to be different [111], hence the gain.

Figure 4.7 (left) shows an example of the signal amplitude distributions obtained with the ^{55}Fe source from a dose 4 sample. Figure 4.7 (right) shows the same spectrum generated using the electric field profile extracted from TCAD simulations and calculating the gain using the impact multiplication formula from [112] for primary clusters uniformly distributed throughout the sensor volume, for qualitative comparison.

The peak in amplitude visible at 10 mV can be attributed to X-ray photons that convert inside the drift region. In this case, holes will drift towards the gain region, producing a hole-initiated avalanche. Conversely, the peak at 63 mV can be attributed to photon conversions within the thin absorption region. In this case, the primary electrons will drift towards the gain layer, producing the much larger electron-initiated avalanche. The flat region of the spectrum between the two peaks can be attributed to photons that convert either in the inter-pixel area or within the gain layer. In the inter-pixel area, the gain is lower, as shown from the gain calculation in Figure 4.3 obtained with TCAD transient simulation. For photons whose conversion happens inside the gain layer, some electron-hole pairs initiate an avalanche with holes, and others get a mixed avalanche whose gain stands in between hole-initiated and electron-initiated avalanche gain.

The analog front-end electronics convert the collected charge into a voltage signal according to the Shockley-Ramo theorem (see Section 3.3). We can assume a direct

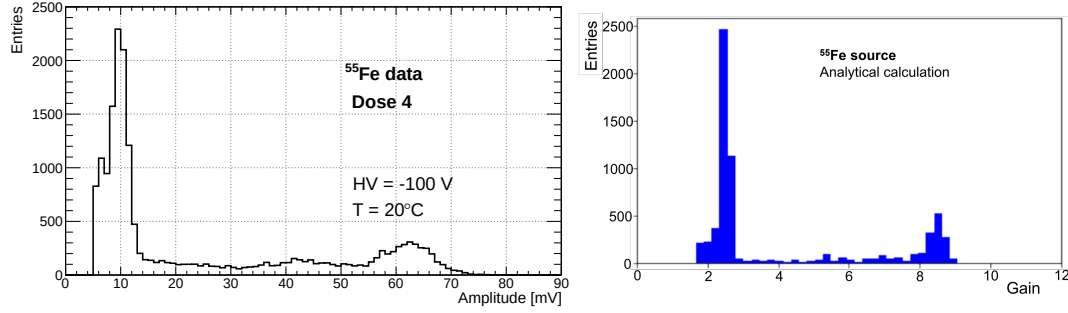


FIGURE 4.7: (Left) Signal amplitude spectrum measured with a ^{55}Fe source for a dose4 PicoAD sample at a temperature of 20°C and high voltage of 100 V. (Right) Gain spectrum was obtained using the electric field profiles from TCAD simulation and the analytical formula given in [112] for the gain obtained from impact ionization.

proportionality between the amount of charge inside the sensor, n , and the amplitude of the output signal waveform as seen by the oscilloscope, amp . The conversion factor is the gain of the amplifier G_{preamp} . The amplitude of the signal is: $amp_e = G_{preamp} * n_{abs}$ if the photon conversion happened in the absorption layer and the electrons initiated the avalanche; or $amp_h = G_{preamp} * n_{drift}$ if the photon conversion happened in the drift layer and the holes initiated the avalanche.

The preamplifier gain has been calibrated using different methods that give compatible results: a ^{55}Fe source, a ^{109}Cd source, and an infrared laser. The calibration was performed under a condition where there was no charge multiplication. This condition is fulfilled if the charge cluster is generated inside the drift region; the dose is the lowest, i.e., dose 1; the High Voltage HV is the lowest possible to have the PicoAD detector fully depleted, i.e., 85 V. The condition of no charge multiplication is supported by the spectrum, which has only one peak in these conditions. The calibration gave a preamplifier gain of $G_{preamp} = 27 \text{ mV/fC}$.

Spectra like the one in Figure 4.7 have been acquired at different voltages and temperatures for each of the four doses of the gain layer implantation. The average amplitudes for hole and electron gain have been obtained as the most probable value of the local maxima associated with each spectrum peak. Then, the preamplifier gain, G_{preamp} , was used to calculate the equivalent charge. Figure 4.8 shows the average collected charge for the hole-gain (left) and electron-gain (right) peaks of the ^{55}Fe spectrum as a function of the HV for the four gain-layer implant doses, measured at temperatures between -20 and 20°C .

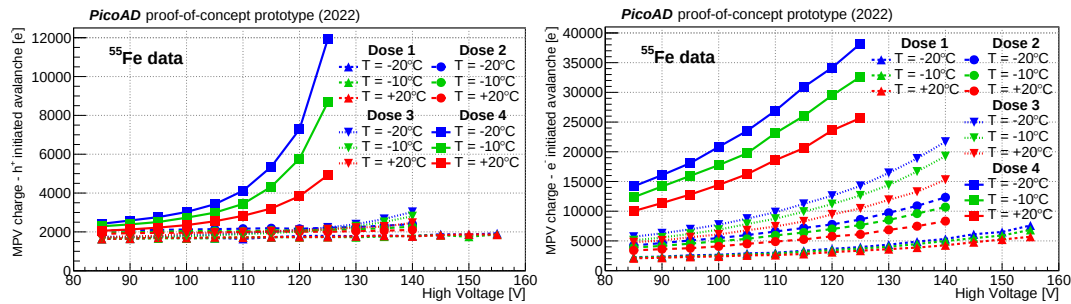


FIGURE 4.8: Average collected charge of the signals produced in the PicoAD by converting X-ray photons from a ^{55}Fe source. (Left) With a hole gain produced by photon conversion in the drift region. (Right) With electron gain produced by photon conversions in the absorption region.

All four implant doses show an increase in the mean collected charge with increasing sensor bias, indicating that electron or hole gain is present for charge clusters produced in the absorption or drift region.

To extract the sensor avalanche gain, the preamplifier gain of $G_{preamp} = 27 \text{ mV/fC}$ was used in combination with the measured mean amplitudes, recalling that the primary charge $n_i = 1640$ initial electron-hole pairs is also known. The gain was calculated according to:

$$\begin{aligned} G_e &= \frac{n_{abs.}}{n_i} = \frac{amp_e}{G_{preamp} * n_i'} \\ G_h &= \frac{n_{drift}}{n_i} = \frac{amp_h}{G_{preamp} * n_i} \end{aligned} \quad (4.1)$$

Figure 4.9 shows the result of the calculation for the electron gain, i.e., gain of the avalanche multiplication initiated by electrons. Figure 4.9 shows the electron gain as a function of the sensor bias voltage for temperatures between -20°C and 20°C .

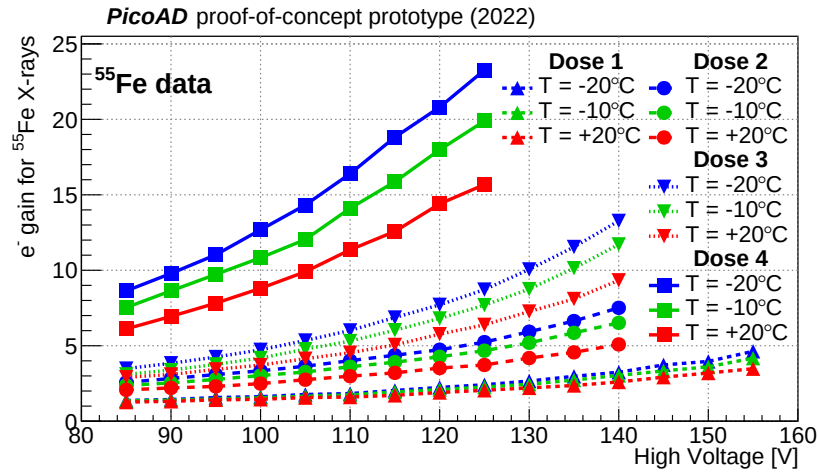


FIGURE 4.9: Gain measured for the conversion of X-ray photons from a ^{55}Fe source for electron-initiated avalanche as a function of the sensor bias voltage. For each temperature, the gain is obtained as the ratio between the mean collected charge of signals generated by photon conversion in the PicoAD absorption layer and the mean value obtained at 85 V for conversions in the drift region in the sensor with gain-layer dose 1.

A maximum electron gain of 23 was observed for the sensor with gain-layer implant dose 4, operated at the lowest ambient temperature of -20°C . Coherently with the expectation, sensors with lower gain-layer implant doses show lower electron gain at a fixed bias voltage.

The PicoAD structure offers the opportunity to study the relation between the electron and hole gain in the same junction, which can be used to infer the properties of the avalanche. Figure 4.10 shows the ratio between electron gain and hole gain as a function of the electron gain. The gain for holes was computed using the same method as the gain for electrons described above. This representation of the data removes the direct dependence of the gain on temperature and sensor bias, showing only the characteristics of impact multiplication in silicon, which in turn depend on parameters of the gain layer, such as its width or the magnitude of the electric field.

As expected, all the devices are on the same trend curve since the implantation energy used for the different doses is the same. For the sensor with dose 4, which exhibits the highest absolute electron gain, a reduction in the electron-hole gain ratio

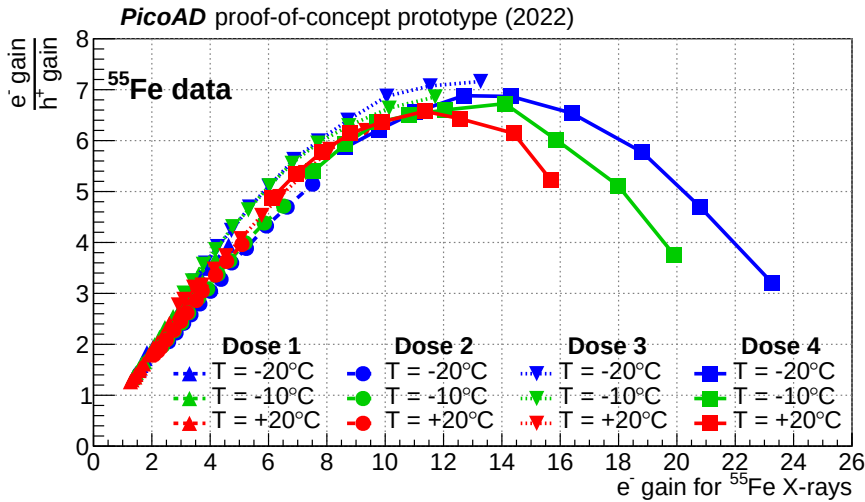


FIGURE 4.10: Ratio between the electron gain and the hole gain obtained with a ^{55}Fe source for the four gain layer implant doses as a function of the electron gain.

is observed for electron gains exceeding 12. This observation cannot be reproduced by the model of the impact ionization with any parametrization for the impact parameters [112].

Transient space-charge effects can explain the reduction in the gain ratio at high electron gain during avalanche development, which reduces the electric field in the gain layer. These space-charge effects occur primarily when the primary cluster is formed within the absorption region. The time-dependent spatial charge density varies with the total charge drifting inside the sensor, which is the product of the gain and the primary charge. The identical TCAD simulation introduced in Section 4.1 to obtain the gain profile inside the sensor was used to estimate the impact of space-charge effects on the measured electron gain. Figure 4.11 compares the gain profile expected for a primary charge deposited at different sensor depths at the center of the pixel.

Primary charges of 63 and 1640 electron-hole pairs were used, the first representing the ionization of a MIP in $1\text{ }\mu\text{m}$ and the latter the charge deposited by the photoelectric effect by an X-ray from the ^{55}Fe source. The simulation confirms that the large electric-charge density in the case of ^{55}Fe X-ray photon conversion considerably limits the gain for an avalanche initiated by electrons. At the same time, it has little to no impact when holes initiate the avalanche. This result might explain the behavior observed at high electron gain in Figure 4.10. It also suggests that the actual electron gain achievable with MIPs is much larger than the value measured with the ^{55}Fe source.

Finally, PicoAD proof-of-concept prototypes are functional. The IV characteristics behave as expected from TCAD simulations. However, large extra currents flowing to the guard ring through the gain layer can be reduced in future prototypes by refining the gain layer masks. Measurements with a ^{55}Fe X-ray source show increasing gains with increasing dopant concentration levels of the gain layer, as expected. The gain increases at lower temperatures, as expected from [113].

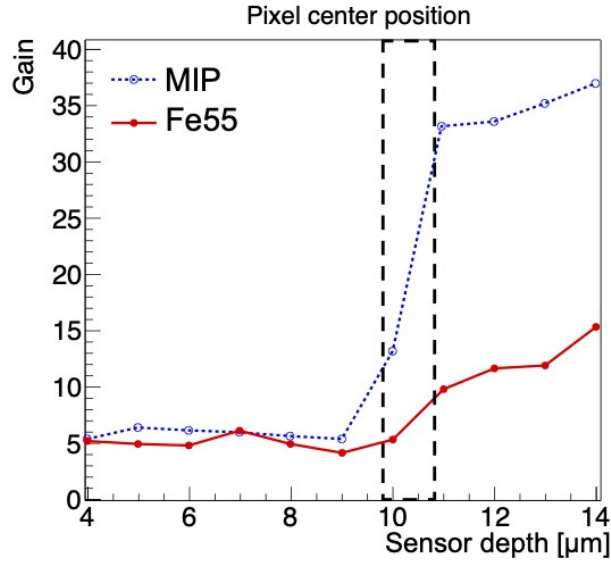


FIGURE 4.11: TCAD simulation of the gain of a point-like primary charge deposition at the center of a PicoAD pixel and at varying depth of 63 electrons (MIP, blue dotted curve, also shown in Figure 4.3) and 1640 electrons (^{55}Fe X-ray conversion, red curve). The black dashed lines indicate the high field region produced by the gain layer.

4.3 Test Beam

This section presents the measurements of the PicoAD proof-of-concept prototype performed during a CERN SPS test beam with pions with momentum $180\text{ GeV}/c$. The measurements were conducted in the experimental *North Area* at the Preveessin site of CERN during October and November 2022, where pions with the selected momentum are delivered in a beam, approximately a few mm wide, directly to the PicoAD proof-of-concept prototype under test.

A test beam aims to characterize the sensor using particles that behave like minimum ionizing particles (MIPs) to study efficiency and time resolution in an application as close to the real world as possible. As said in Section 3.2.1, the plateau of minimum energy deposition covers an extensive range of energies. Therefore, in a real case application, such as a tracker detector in a collider experiment, it is very likely to have to detect particles that behave like MIPs.

First, we present the experimental setup used for the test beam, which includes a telescope to track the incoming pions and fast detectors to provide a time reference. We discuss some insights into the selection of the track sample used to measure efficiency and time resolution, since similar discussions are also valid for the following test-beam sections for the other prototypes (see Section 5.2 and Section 6.2), although they will not be presented again.

Second, we describe the signal amplitude dependence on the position of the incoming particle, showing that the gain is lower close to the inter-pixel region and suggesting a degradation in overall performance as the particle hits the pixel far away from the center.

Then, we present the efficiency and time resolution results. The sensor performance is related to two main factors that can be modified: the bias voltage HV applied to the sensor, and the power consumption density P_{density} absorbed by the front-end electronics. Efficiency and time resolution are mapped into the available

parameter space of bias voltage (HV) and power density (P_{density}). Finally, we present the dependence on the position of the incoming particle for both efficiency and time resolution, confirming a performance degradation near the inter-pixel region and suggesting ways to improve the detector.

4.3.1 Experimental Setup

The experimental setup for the test beam comprises a PicoAD proof-of-concept prototype ASIC with the highest implant dose as the detector under test (DUT). The DUT was placed in an aluminum cooling box and operated at a temperature of -20°C . The DUT was positioned, together with two LGADs [71] (named here LGAD0 and LGAD1), in the middle of the six detection planes of the UniGe FEI4 telescope for charged-particle tracking [114], as shown in the layout in Figure 4.12. Two SPAD detectors [115] (named here SPAD0 and SPAD1) were installed downstream of the telescope to have additional time references of very high precision.

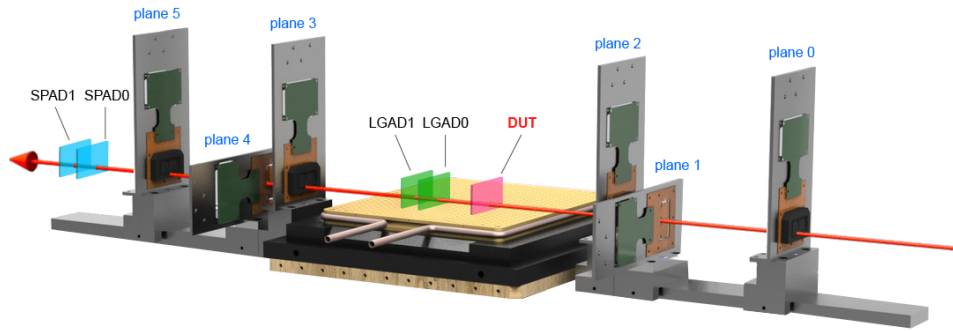


FIGURE 4.12: Schematic view of the experimental setup, showing the six planes of the UNIGE FEI4 telescope [114], the DUT, the two LGADs, and the two SPADs.

Two oscilloscopes read the output of the four analog pixels of the DUT (see Figure 1.6 for the pixels layer in the ASIC), the two LGADs, and the two SPADs, while a dedicated data acquisition system read the signals from the telescope planes. Pixel OA0 of the DUT, SPAD0, and the two LGADs were connected to a first oscilloscope with an analog bandwidth of 3 GHz and a sampling rate of 40 GSamples/s. Instead, pixels OA1, OA2, and OA3 of the DUT and SPAD1 were read by a second oscilloscope with 4 GHz analog bandwidth and a sampling rate of 20 GSamples/s. All four analog pixels were used to measure efficiency, while only pixel OA0, connected to the oscilloscope with a higher sampling rate, was used to measure time resolution.

The FEI4 telescope provided the trigger for the oscilloscopes. A region of interest of $250\ \mu\text{m} \times 600\ \mu\text{m}$ centered at the OA0 pixel of the DUT was imposed on the first telescope plane and put in coincidence with the last plane to generate the trigger. Figure 4.13 shows a typical small-amplitude waveform simultaneous to a telescope trigger, acquired for pixel OA0 with the DUT operated at preamplifier power density of $P_{\text{density}} = 2.7\ \text{W}/\text{cm}^2$ and bias voltage $HV = 125\ \text{V}$.

The waveforms acquired by the oscilloscopes, as in Figure 4.13, were recorded and used in analyzing the selected data via some codes written in modern C++ language integrated with the ROOT framework [116]. For each pixel of the DUT, the signal was delayed to the second half of the 500 ns waveform time window acquired by the oscilloscopes so that the first half of the waveform could be used to determine the voltage noise σ_V at the output of the analog front-end electronics in each given event. Since the DUT signal amplitudes were sampled at finite time intervals (25 ps for the

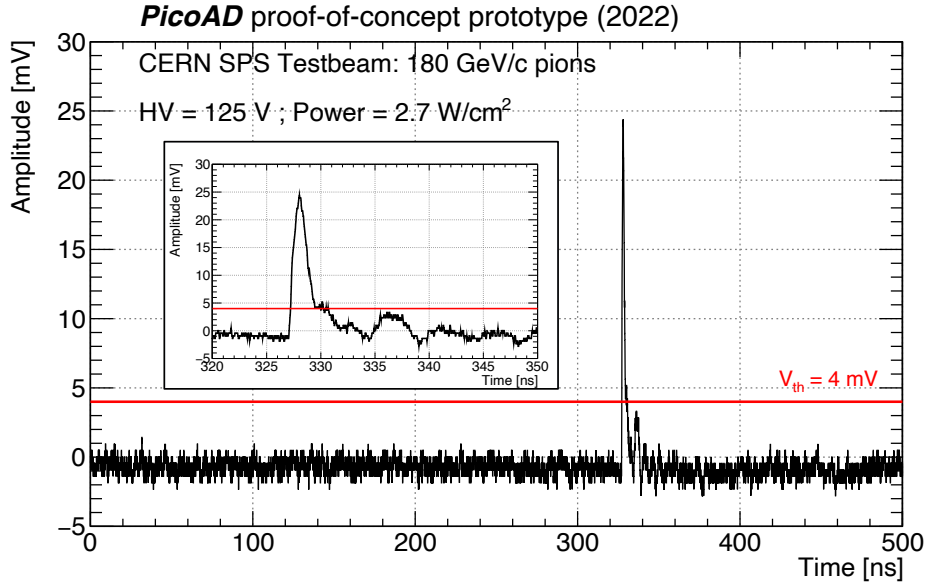


FIGURE 4.13: Example of a small-amplitude waveform acquired for pixel OA0 with the DUT operated at preamplifier power density $P_{\text{density}} = 2.7 \text{ W/cm}^2$ and bias voltage $HV = 125 \text{ V}$. The time region before 250 ns is the portion of the waveform used to extract the voltage noise σ_V . The insert shows the same signal on a magnified time scale. The full red lines indicate the discrimination threshold, V_{th} , used in the analysis.

first oscilloscope and 50 ps for the second), a linear interpolation between samplings was performed to obtain the signal characteristics.

A certain number of cuts on the quality of the tracks is necessary to have clean data samples. The requests can be roughly summarized as follows:

- the particle must have passed into the determined regions of interest of the telescope planes;
- quality cuts on the tracks, required that only one particle per event traversed all six planes, and whose track reconstruction gave a positive result, i.e., the calculated χ^2/NDF is lower than 1.6.

Two other aspects must be taken into account. First, the signal produced in the DUT pixel must be higher than a certain threshold to mark an event as efficient, and hence, used to measure efficiency and time resolution. Therefore, a threshold must be defined. Second, the signal produced in the DUT pixel must have been effectively created by a particle hitting the pixel and not a neighboring pixel that induced a cross-talk signal into the DUT pixel. Cross-talk events may only appear for efficient events, so no cross-talk cut is required for efficiency measurement. In contrast, cross-talk events must be discarded for the time resolution measurement.

After the discussed cuts, the tracks were extrapolated on the DUT, and they were found to fall inside an area defined by the four hexagonal pixels of the DUT. Therefore, the centers of the pixels could be determined, allowing for a reference when plotting amplitude, efficiency, or time resolution as a function of distance from the centers.

The following sections discuss all the cuts on the tracks that led to clean data samples and the definition of pixel centers, threshold determination, and cross-talk.

Offline Masking of the Planes

During the data acquisition, a region of interest of $250\text{ }\mu\text{m} \times 600\text{ }\mu\text{m}$ centered at the OA0 pixel of the DUT was imposed on the first telescope plane and put in coincidence with the last plane to generate the trigger. Its definition consisted of masking all the other pixels of the telescope's first plane outside the region. The region of interest corresponded to the shadow of the four analog pixels of the DUT projected orthogonally onto the first plane, and it was used to exclude useless statistics from tracks that fell outside the DUT.

No masking was set on the other telescope planes, but it is still possible to place an offline masking, since all hits on the telescope planes are also saved. Offline masking consists of discarding tracks outside a defined region in the other five planes after the first one. This operation is equivalent to setting an area of interest on all the telescope planes. It allows for the reduction of spurious tracks with abnormal behavior, tracks with normal behavior that fell outside the DUT, and tracks that crossed the DUT at a relatively large angle.

Figure 4.14 shows an overview of the cut: the 2D histogram displays the entries associated with a specific cluster position (in terms of pixel rows and columns), and the solid red lines show the cuts corresponding to the offline masking. The area of the cut is found by examining the cluster position on a specific plane and identifying a small region with a large number of entries.

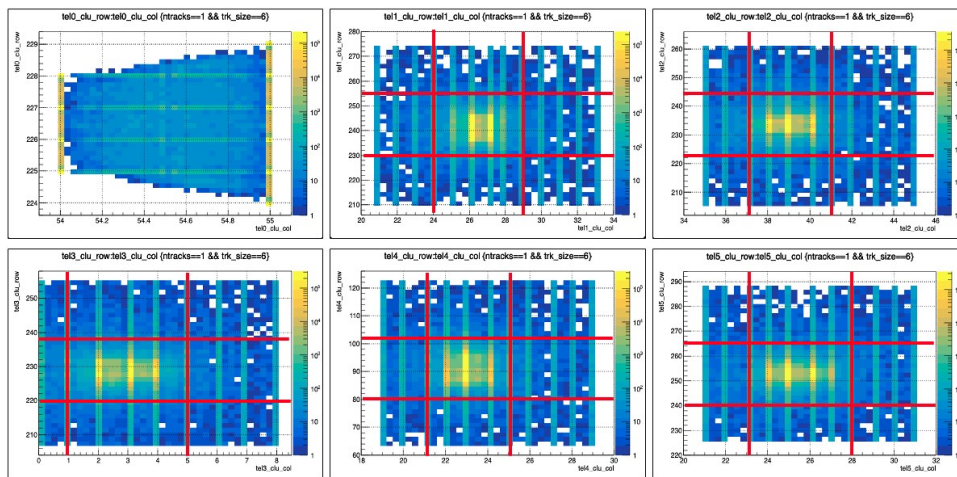


FIGURE 4.14: Overview of the cluster positions on the six planes. The solid red lines define the area of interest imposed by the offline masking on the telescope planes, excluding the first one.

Figure 4.14 shows an interesting feature of the FEI4 telescope, which will also be present during the characterization of the other prototypes: the cluster positions registered from the telescope planes tend to accumulate in vertical and horizontal lines corresponding to integer positions. These are clusters with only a single pixel fired when the particle passed through the plane, i.e., cluster size one. When the cluster size is larger than one, the reconstruction software defines the cluster position as the average of the pixel positions within the cluster, weighted by their charge, resulting in positions spaced between integer numbers. Since the telescope planes are old and irradiated, they are noisy, so the threshold during data taking must be kept high. Thus, it is more likely to find clusters of a single size. Hence, this effect of quantization is more apparent.

The offline masking of the planes removes approximately 10 % of the total statistics. Further studies show no evident correlation between the χ^2/NDF and the tracks inside or outside the offline region of interest, which implies that most of the tracks discarded by the offline masking are well-reconstructed tracks that pass outside the device under test (DUT). Therefore, they do not contribute to the efficiency or time resolution values.

Quality Cuts on Tracks

After the offline masking of the telescope planes, the tracks' positions on the telescope planes were reconstructed offline through a χ^2 minimization, calculated using the difference between the registered position (i.e., the cluster position) and the reconstructed position. The reconstructed tracks were then extrapolated to the DUT plane to determine the position where each pion crossed the sensor. A set of track-quality criteria was designed to select a sub-sample of events that maximizes the telescope's pointing resolution.

First, two particles may have crossed the same telescope plane within the same trigger, and their hits on the telescope plane may have been registered at the same timestamp. An event of this kind led to ambiguities in associating the correct hit with the tracks. To avoid this problem, the first simple quality cut was to request a single track for each trigger: events with multiple tracks were discarded from the clean data samples, which affected about 5 % of the total statistics.

A second quality cut required tracks with a hit registered in all six telescope planes. In this case, the minimization done to reconstruct the position on the telescope planes is more reliable. The number of degrees of freedom for each track is composed of two coordinates (x, y) for each plane (12 initial degrees of freedom) minus two parameters for each minimization (on the x and the y axis), which means 8 degrees of freedom. This cut affected another about 5 % of the total statistics.

The last cut to consider regarded the quality of the reconstruction via software. The parameter to consider was χ^2/NDF , since high χ^2/NDF values are correlated with significant residuals, hence a low probability of occurrence. A proper cut was set using a certain confidence level. In our case, we chose a confidence level of 90 %, which corresponds to a request that χ^2/NDF be less than 1.6.

Figure 4.15 shows a typical distribution of χ^2/NDF for a specific data set, but the same applies to other datasets. The requests to have only one track traversing all six planes are also present. The red line corresponds to $\chi^2/NDF = 1.6$.

The χ^2/NDF distribution peaks at one as expected. The legend of Figure 4.15 indicates that approximately 10 % of the total statistics is lost with this cut.

Considering the offline masking of the telescope planes and the quality cuts on the tracks, the discarded tracks accounted for approximately 30 % of the acquired statistics. For the tracks fulfilling the above selections, the telescope pointing resolution on the DUT plane was estimated to be approximately 10 μm [117].

Definition of the Threshold

We can now focus on the DUT signals with the proper data samples of cleaned tracks. Only signals exceeding a certain discrimination threshold V_{th} in any DUT pixels were considered.

Typically, the method used to define the threshold V_{th} implies taking the distribution of the standard deviation of the noise σ_V , fitting it with a Gaussian curve, and then extracting the mean: the threshold is then a certain multiple of the mean value.

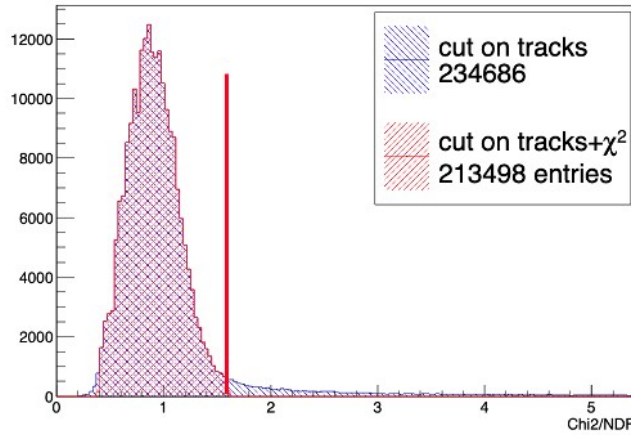


FIGURE 4.15: Typical distribution of χ^2/NDF with the cut on the tracks (one track per event traversing all six planes).

With the PicoAD proof-of-concept prototype, the background distribution was found to be highly non-Gaussian. Thus, the Gaussian fit was not applicable. Figure 4.16 shows some distributions of the RMS background for different runs, and grouped according to the preamplifier current I_{preamp} . The distributions shown are for pixel OA0.

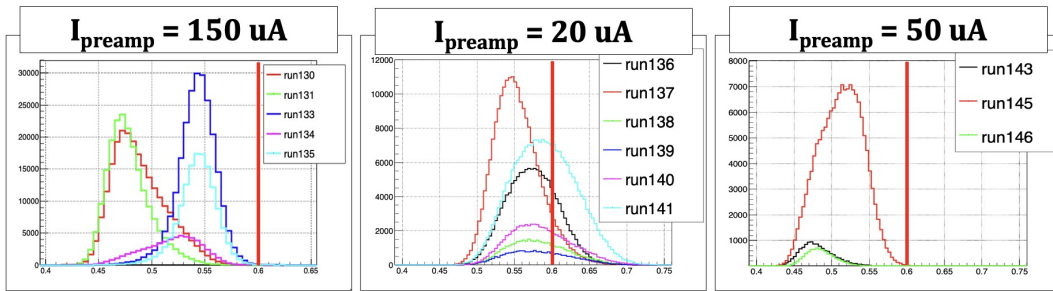


FIGURE 4.16: Some distributions of the RMS background for different runs, grouped according to the preamplifier current I_{preamp} . The red vertical line is at a 0.6 mV.

A DUT effect cannot explain the inconsistency between the distributions of the background noise within the same preamplifier current I_{preamp} . Therefore, it must be attributed to an external noise picked up by the front-end electronics of the ASIC. This noise pick-up can also explain the non-Gaussianity of the background noise for specific runs. Since a Gaussian fit was not applicable, we fixed a hard threshold for all the statistics to 4 mV. Considering a mean standard deviation of the noise $\sigma_V = 0.6$ mV (which overestimates the distribution, as visible from the red vertical line in Figure 4.16), the value of 4 mV is somewhere between 6 and 7 times the standard deviation of the noise.

Finally, a threshold $V_{\text{th}} = 4$ mV was used to mark an event as efficient. The same threshold $V_{\text{th}} = 4$ mV was used to define the time of arrival ToA, hence to measure the time resolution.

Cross-talk Cuts

Cross-talk between the four analog channels in events with significant charge deposition was already observed for the ASIC version without gain [58], which was

produced using the same masks as the PicoAD proof-of-concept prototype discussed here. As expected, this cross-talk is also present in this prototype.

The cross-talk visible in the PicoAD proof-of-concept prototype can be attributed to two possible causes: a feedback path passing through the ground of the board or noise induced by the driver pulse propagating through the power supply. The ASIC addressed both issues for the second MONOLITH prototype without a gain layer. In particular, the power supplies for the preamplifier and the driver stage were separated, as visible in Section 1.2.6.

The insert in Figure 4.13 shows the delayed cross-talk, clearly visible after the real signal as a dumped oscillation. The same strategy adopted in [58] was thus used here to suppress the impact of cross-talk: no additional selection was applied for the efficiency measurement, as cross-talk may only occur for efficient events. In contrast, for the time resolution measurement, only the pixel with the most significant signal amplitude among the four pixels acquired was considered for each event.

Definition of the Centers

The hexagonal shapes of the pixels and the triangles used to compute the efficiency (see next Section 4.3.4) are geometrically built starting from the center of the pixels. There are several ways to find the centers, but all of them consider only those tracks that passed the quality cuts, and whose associated event of the DUT has an amplitude higher than the threshold $V_{th} = 4$ mV.

The first method defined the center coordinates as the average of each track's x and y coordinates. As explained before, this method was found to be strongly biased by the telescope's finite resolution and specific patterns created due to high levels of plane irradiation.

The second method involved defining the center coordinates that maximize the number of tracks within the DUT area. This method consisted of scanning the x and y coordinates of the center of pixel OA0 iteratively. For each center coordinate, the hexagonal area of the pixel was built, and the tracks geometrically inside the hexagon were counted. Figure 4.17 shows the number of entries inside the hexagonal area as a function of the scanning coordinate x and y of the center of pixel OA0.

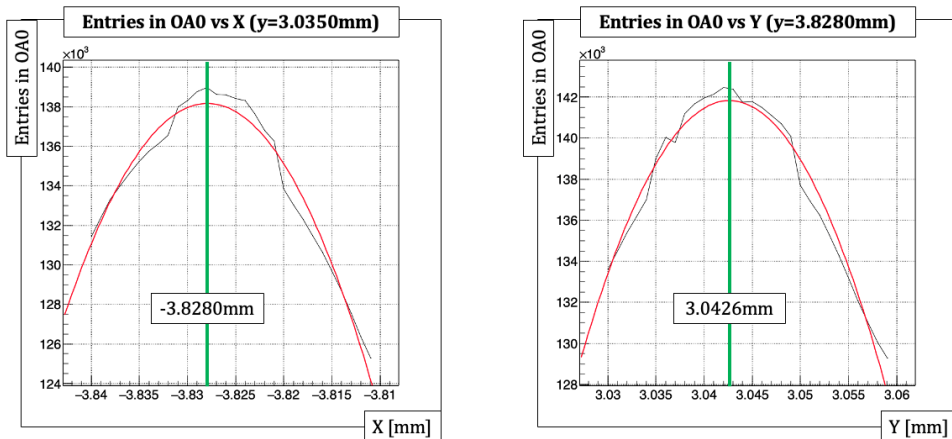


FIGURE 4.17: Number of entries inside the area of the hexagonal pixel OA0 as a function of the pixel center's x and y coordinates. The parabolic fits for each scan are superimposed.

A parabolic fit was then used to find the value of the coordinates x and y that maximizes the entries inside OA0. The two methods provided centers that differ

by about $2\text{ }\mu\text{m}$ on both coordinates. The telescope pointing resolution is known to be $\sim 10\text{ }\mu\text{m}$, so no radical changes were expected. Indeed, the efficiency was compatible with errors when using the centers with both methods. The method that maximizes the statistics was found reliable when tested on different datasets, following qualitative visual inspections. Moreover, this method appeared to suffer less from the bias introduced by the patterns on the telescope planes. Hence, it will be used in the future.

4.3.2 Data Samples

To study the response of the front-end electronics, data were acquired at three working points with amplifier bias current $I_{\text{preamp}} = 20, 50, \text{ and } 150\text{ }\mu\text{A}$. Since the front-end electronics were operated at 1.8 V , these currents correspond to a power density $P_{\text{density}} = 0.4, 0.9, \text{ and } 2.7\text{ W/cm}^2$, respectively. For the working point with $I_{\text{preamp}} = 150\text{ }\mu\text{A}$, a bias voltage scan was also performed by acquiring two additional datasets with $HV = 105$ and 115 V .

The number of triggered events in each data sample is reported in Table 4.1, which also contains the average number of telescope tracks retained for each of the four DUT pixels after the selection described in the previous section.

Working point		Total events	Selected tracks
P_{density}	HV	recorded	per DUT pixel
2.7	125	888k	67k
2.7	115	224k	16k
2.7	105	152k	12k
0.9	125	194k	25k
0.4	125	795k	64k

TABLE 4.1: Number of total recorded events and average selected telescope tracks pointing to each of the four DUT pixels for the five working points studied.

The significant reduction in the number of selected telescope tracks for recorded events is due partly to the trigger region of interest, which is much larger than the DUT area, and partly to the stringent selection criteria for telescope tracks described in the previous paragraphs.

4.3.3 Signal Amplitudes

To inspect the performance of the PicoAD proof-of-concept prototype in detail and gather information that could help improve the design of future prototypes, variations in the DUT's performance within the pixel area were investigated using the telescope's pointing capabilities. With r being the distance of the center of the hexagonal pixel from the point of intersection between the track reconstructed by the telescope and the DUT plane, the pixel area was divided into five radial regions containing a similar number of telescope tracks: $r \leq 25\text{ }\mu\text{m}$, $25\text{ }\mu\text{m} < r \leq 38\text{ }\mu\text{m}$, $38\text{ }\mu\text{m} < r \leq 48\text{ }\mu\text{m}$, $48\text{ }\mu\text{m} < r \leq 55\text{ }\mu\text{m}$ and $r > 55\text{ }\mu\text{m}$.

The variation in signal amplitude across these five radial regions was studied for pixel OA0 of the DUT at different sensor bias voltage values. For this purpose, the signal amplitude distributions in each bin in the distance r and HV value were fitted with a Landau function. Figure 4.18 shows three examples of the fitted distributions.

The most probable amplitude value obtained by the fit in each distance bin was then used to quantify the amplitude variation between the center of the pixel and its

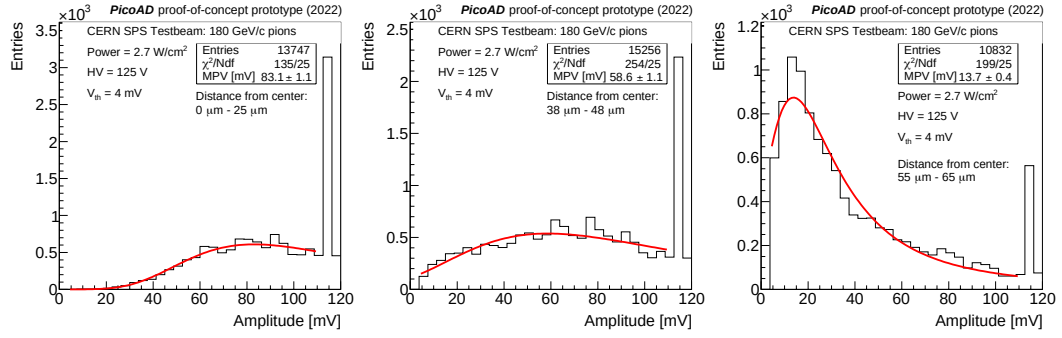


FIGURE 4.18: Amplitude distribution of signals from pixel OA0 for a power density of 2.7 W/cm^2 and $HV = 125 \text{ V}$ for telescope tracks at distances within 0-25 μm (left), 38-48 μm (center) and 55-65 μm (right) from the pixel center. The red lines show the fitted Landau probability distribution function performed in each distance range to obtain the amplitude's most probable value (MPV). The oscilloscope setting produces saturated amplitudes near 120 mV (not included in the fit).

edge. The result is shown in Figure 4.19 for the three values of the sensor bias voltage considered.

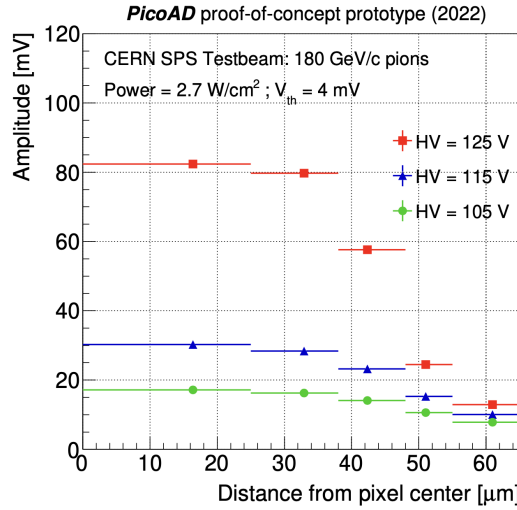


FIGURE 4.19: Most probable value of the DUT's signal amplitude recorded by pixel OA0 as a function of the distance from the pixel center. The data refer to the three values of sensor bias voltage acquired at a power density of 2.7 W/cm^2 . In each bin, the data points are positioned at the mean distance from the center of the pixel.

As expected, the signal amplitude depends on the sensor bias voltage, which is explained by the increasing gain with increasing sensor bias voltage. The significant drop in amplitude towards the edge of the pixel is due partly to charge sharing between adjacent pixels and partly to the lower electric field under the p -stop structure, which electrically separates the pixels. The pixel edge is at a distance from the pixel center between 56 μm (center of the hexagon side) and 65 μm (corner of three pixels). It should be noted that the telescope spatial resolution of approximately 10 μm implies that migrations between adjacent bins of the distance r from the pixel center are not negligible. These migrations were not unfolded in the results shown in Figure 4.19 as well as in the other results as a function of r reported in the following sections.

4.3.4 Efficiency

The detection efficiency was defined as the fraction of selected telescope tracks that resulted in a recorded signal above the discrimination threshold V_{th} in at least one of the DUT pixels. Only tracks crossing the area of the four instrumented pixels were considered, with a tolerance of $10\mu m$ beyond the pixel edges, corresponding to one standard deviation of the telescope pointing resolution.

The left panel of Figure 4.20 illustrates the efficiency measured in the whole area corresponding to the four analog pixels when the DUT was operated at $P_{density} = 2.7 W/cm^2$ (or equivalently $I_{preamp} = 150 \mu A$) and $HV = 125 V$. The right panel of Figure 4.20 shows the efficiency measured in the two triangles defined by the centers of the three left-most (OA0, OA2, OA3) and right-most (OA0, OA2, OA1) pixels.

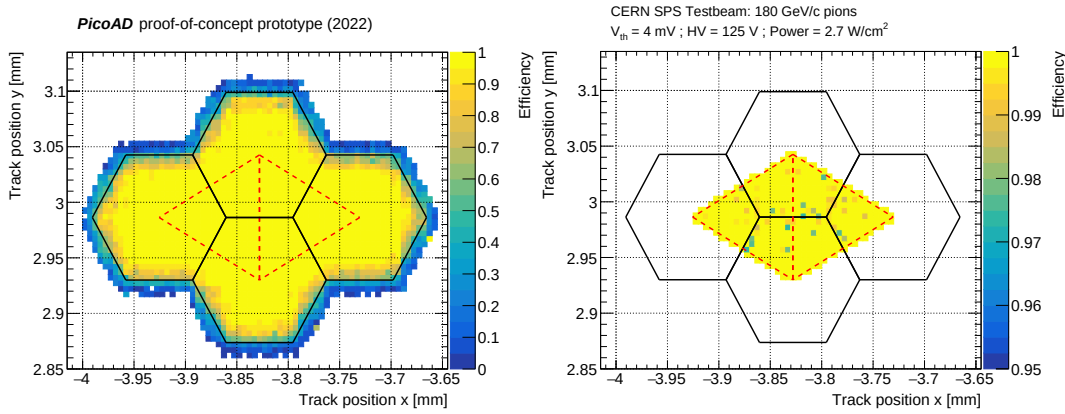


FIGURE 4.20: Efficiency measured for the DUT operated at a power density of $2.7 W/cm^2$ for a discrimination threshold $V_{th} = 4 mV$. The black lines indicate the pixel edges. The left panel shows the results for the entire surface of the four analog pixels. The apparent efficiency degradation around the external edges results from the FEI4 telescope's pointing resolution. The right panel illustrates the efficiency in the two triangular areas delimited by the centers of the three left-most and right-most pixels, which are unaffected by the telescope resolution and are therefore used throughout this study. To highlight the small inefficiencies, measured mostly in the inter-pixel regions, the color scale in the right panel starts from 95 %.

The apparent degradation in detection efficiency observed around the external edges of the four pixels can be attributed to the telescope pointing resolution of $10 \mu m$, since no efficiency degradation is visible in the five internal inter-pixel boundaries. To overcome the bias from the telescope resolution, the efficiency measurement was then restricted to the two triangular areas shown in Figure 4.20, right, defined by the centers of the three left-most (OA0, OA2, OA3) and right-most (OA0, OA2, OA1) pixels. Each triangle covers exactly one-sixth of the area of the three hexagonal pixels involved and contains the boundaries between two pixels and the three pixels' corners in the same proportion as a pixel. They provide, therefore, a good representation of the efficiency behavior across an entire pixel. With few exceptions, the bins inside the two triangles corresponding to slightly lower measured efficiencies tend to accumulate near the pixel borders. This effect can be explained by the lower signal amplitudes in these regions displayed in Figure 4.19.

The average efficiency of the two triangles provides a reliable estimation of the efficiency over the entire hexagonal pixel area, unbiased by the telescope resolution. These efficiencies are used to study different working points of bias voltage and power consumption in the detector.

Statistical Error Estimation

Two possible approaches, the Bayesian and the frequentist approaches, were used to estimate the statistical uncertainties and quote the average efficiency of the two triangles.

The standard Bayesian approach involves a Bayesian inversion, which entails computing the value of the efficiency, generating the Probability Density Function (PDF) of the efficiency using a prior distribution, and determining the values on the left and right that, when integrated, yield a specific confidence level. Figure 4.21 shows the PDF of an efficiency measured with eight efficient events over 10 total events when a uniform prior distribution is used. A confidence level of 67.3 % defines an uncertainty σ around the average efficiency ϵ .

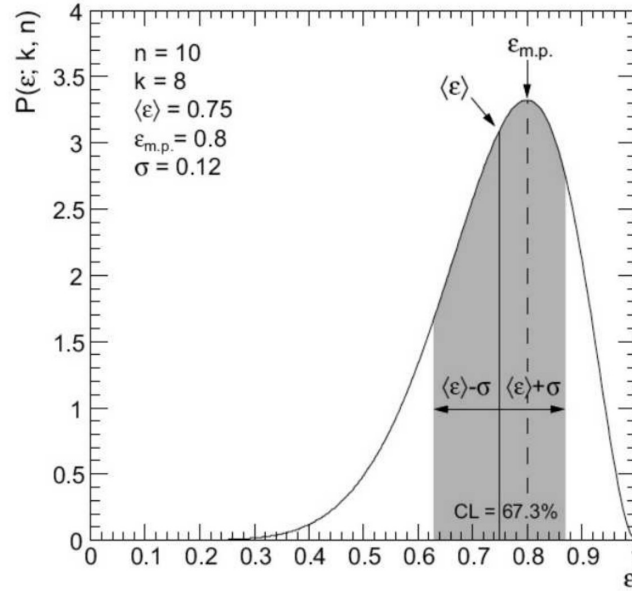


FIGURE 4.21: Probability density function of an efficiency measured with eight efficient events over 10 total events when a uniform prior distribution is used. A confidence level of 67.3 % defines an uncertainty $\sigma = 0.12$ around the average efficiency ϵ .

The Bayesian approach prescribes quoting the most probable value of the efficiency PDF as the best estimator for efficiency. For example, the data shown in Figure 4.21 would result in a quoted efficiency of $\epsilon = (0.75 \pm 0.12) \%$.

When Bayesian inversion is used, one must be careful in defining the prior probability used, as the shape of the efficiency distribution function depends strongly on it. Typically, the prior probability used is a uniform distribution between 0 and 1, knowing that the efficiency cannot exceed this interval; other prior distributions are also possible. Starting from the values of efficient tracks and total tracks, the approaches used are summarized as follows:

- a Bayesian approach with a uniform prior distribution, considering the mean as the best estimator and doing an inversion by hand;
- a Bayesian approach with a Jeffreys prior distribution [118], considering the most probable value as the best estimator;
- a Bayesian approach with a uniform prior distribution, considering the most probable value as the best estimator;

- a Clopper-Pearson approach [119], which is completely frequentist.

Figure 4.22 shows the efficiency results as a function of the preamplifier current for the left triangle and with the four different approaches quoted above. The confidence level is set to 95 % for all approaches.

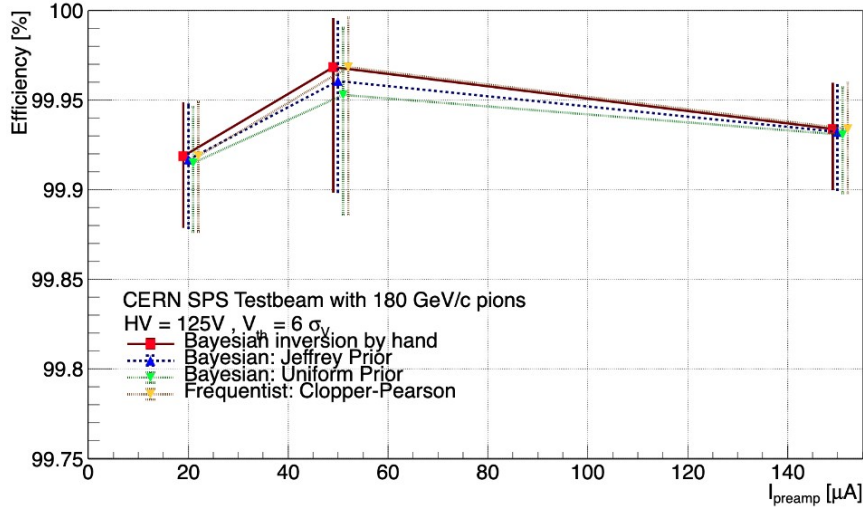


FIGURE 4.22: Comparison of the efficiency as a function of the preamplifier current for the left triangle and the four approaches quoted above. The confidence level is set to 95 % for all approaches.

The results from the four different approaches are compatible within the experimental errors. When using the Bayesian approach, one should use the mean value as the best efficiency estimator. In contrast, the frequentist approach uses the mode as the best estimator, which is simply the ratio of the number of efficient events to the total number of events. The Particle Data Group [10] recommends the frequentist approach with the Clopper-Pearson intervals when quoting the efficiency, for simplicity in estimating the efficiency mode and avoiding problems when choosing a prior distribution. Therefore, this approach was used to produce the final results.

Threshold Scan

Figure 4.23 illustrates the results of a scan over the discrimination threshold V_{th} , taken at integer multiples of the voltage noise σ_V , for $P_{density} = 2.7 \text{ W/cm}^2$ and $HV = 125 \text{ V}$.

The efficiency measured in the two triangles below 99 % only for threshold values above $13\sigma_V$, demonstrating the ASIC's capability to operate with very high detection efficiency even at relatively high thresholds.

Bias Voltage and Power Density Scans

Table 4.2 summarizes the efficiency measured in the left and right triangles shown in Figure 4.20 for the five working points at which data were acquired. For each working point, the average of the two measurements is reported in the last column of Table 4.2.

When the sensor is operated at a bias voltage $HV = 125 \text{ V}$, the efficiencies for the three power consumption values considered are observed to be compatible with 99.9 % within uncertainties. A slight decrease in detection efficiency is observed at $HV = 105 \text{ V}$, and 115 V , and the highest power consumption measured $P_{density} =$

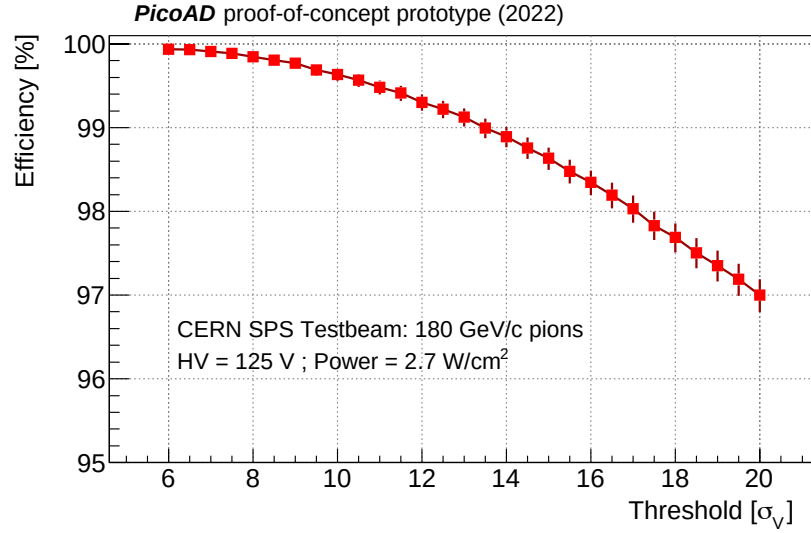


FIGURE 4.23: Detection efficiency measured within the two triangular areas shown in Figure 4.20, as a function of the discrimination threshold in multiples of the voltage noise σ_V . The data refer to $HV = 125$ V and $P_{\text{density}} = 2.7$ W/cm². The 4 mV threshold used corresponds approximately to 6.5 times the voltage noise σ_V .

I_{preamp} [μ A]	P_{density} [W/cm ²]	HV [V]	Efficiency [%]		
			<i>Left</i> triangle	<i>Right</i> triangle	Average
20	0.4	125	99.92 ^{+0.03} _{-0.04}	99.87 ^{+0.05} _{-0.04}	99.90 ^{+0.04} _{-0.05}
50	0.9	125	99.95 ^{+0.04} _{-0.09}	99.95 ^{+0.04} _{-0.09}	99.95 ^{+0.04} _{-0.09}
150	2.7	125	99.87 ^{+0.04} _{-0.05}	99.89 ^{+0.03} _{-0.05}	99.88 ^{+0.04} _{-0.05}
150	2.7	115	99.73 ^{+0.11} _{-0.15}	99.62 ^{+0.13} _{-0.17}	99.68 ^{+0.12} _{-0.16}
150	2.7	105	99.00 ^{+0.03} _{-0.03}	98.75 ^{+0.03} _{-0.04}	98.86 ^{+0.03} _{-0.03}

TABLE 4.2: Detection efficiency measured for the PicoAD proof-of-concept prototype at different I_{preamp} and HV values. The efficiencies are separately extracted in the two triangular areas connecting the pixel centers of the three *left*-most (OA0, OA2, OA13) and *right*-most (OA0, OA2, OA3) pixels shown in Figure 4.20. The average efficiencies in the two triangles are also reported. Uncertainties are statistical only.

2.7 W/cm², consistently with the expected reduction of the sensor gain at these sensor bias voltages.

Figure 4.24 shows the average efficiency for the two triangles, as reported in the last column of Table 4.2.

Efficiency in Different Regions of the Pixel

The DUT efficiencies were measured as a function of the distance from the pixel center for the five working points acquired. Figure 4.25 shows the efficiencies measured in the same five radial regions described in the previous section.

A drop in efficiency of approximately 2 % is observed at the edge of the pixel when the sensor bias decreases from SI 125 V to 105 V. The reduced electric field can explain this drop in this prototype under the *p*-stop inter-pixel structure. In the case of $HV = 125$ V, the efficiency drop at the edge of the pixel is only 0.2 % and compatible with uncertainties for the three power consumption values investigated here.

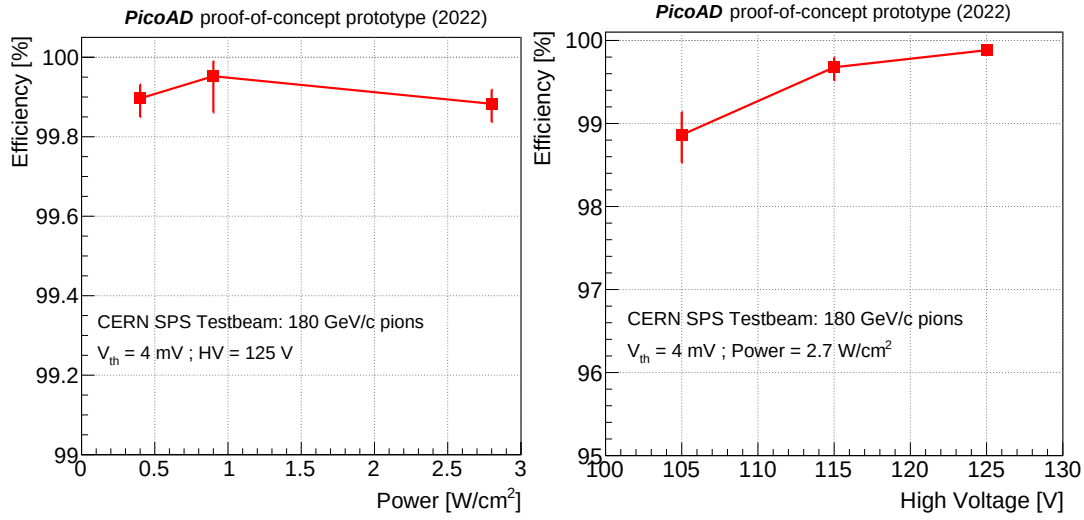


FIGURE 4.24: Detection efficiency measured within the two triangular areas shown in Figure 4.20 with $V_{th} = 4$ mV. The left panel shows the efficiency measured at $HV = 125$ V for three power density values. In comparison, the right panel shows the efficiency measured for three values of HV at a power density of 2.7 W/cm² (corresponding to $I_{preamp} = 150$ μ A).

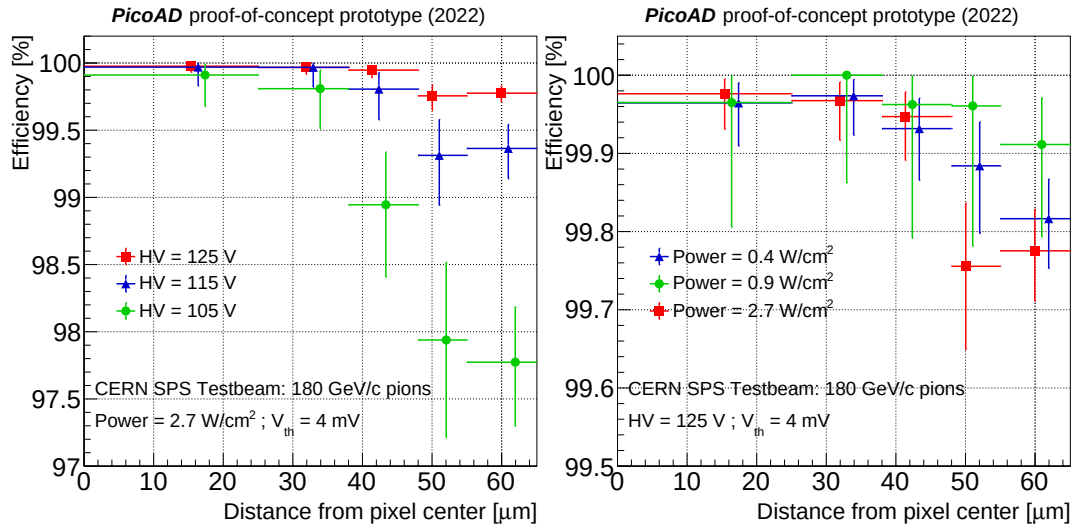


FIGURE 4.25: Detection efficiency measured within the two triangular areas shown in Figure 4.20 with $V_{th} = 4$ mV, shown as a function of the distance from the center of the pixels. The left panel shows the efficiency measured for three values of HV at a power density of 2.7 W/cm² (corresponding to $I_{preamp} = 150$ μ A). In comparison, the right panel shows the efficiency measured at $HV = 125$ V for the three power density values. To improve readability, the data are slightly shifted on the horizontal scale.

4.3.5 Time Resolution

The time resolution of the PicoAD proof-of-concept prototype was measured by constructing a coincidence system between the device under test (DUT) and the two large-area avalanche detectors (LGADs). The Time-of-Arrival (ToA) was defined as the time value at which the interpolated signal passed the threshold of 4 mV. The distribution of the difference in ToA (Δ ToA) was studied separately for the three detector pairs: DUT-LGAD0, DUT-LGAD1, and LGAD0-LGAD1, and used to

measure the corresponding standard deviation, $\sigma_{\Delta\text{ToA}}$ via a Gaussian fit. The result is a system of three equations, whose solution provides the DUT time resolution $\sigma_{t,\text{DUT}}$ and the two LGAD time resolutions $\sigma_{t,\text{LGAD0}}$ and $\sigma_{t,\text{LGAD1}}$:

$$\begin{aligned}\sigma_{\Delta\text{ToA}(\text{DUT},\text{LGAD0})}^2 &= \sigma_{t,\text{DUT}}^2 + \sigma_{t,\text{LGAD0}}^2 \\ \sigma_{\Delta\text{ToA}(\text{DUT},\text{LGAD1})}^2 &= \sigma_{t,\text{DUT}}^2 + \sigma_{t,\text{LGAD1}}^2 \\ \sigma_{\Delta\text{ToA}(\text{LGAD0},\text{LGAD1})}^2 &= \sigma_{t,\text{LGAD0}}^2 + \sigma_{t,\text{LGAD1}}^2\end{aligned}\quad (4.2)$$

The four PicoAD analog pixels were considered individually. When a signal above the threshold was recorded in more than one pixel, such as particles crossing inter-pixel regions and producing charge shared between adjacent pixels, only the signal in the pixel with the largest amplitude was considered.

The oscilloscopes showed saturation of the amplitudes at around 120 mV for pixel OA0 and 80 mV for the other three pixels. Since, in the rare case of two saturated signals, selecting the pixel with the highest amplitude was impossible, events in which two pixels had an amplitude above 75 mV were analyzed for both pixels. It was nonetheless checked that retaining or rejecting the small fraction of events with such ambiguity has no impact on the results of the time resolution measurement.

Time-Walk Correction

As Section 3.5.2 explains, signals of different amplitudes cross the discrimination thresholds with various delays. This effect, commonly referred to as time walk, results in the variation of the average of the difference ΔToA as a function of the signal amplitude and must be corrected. The plots of Figure 4.26 and Figure 4.27 show the ΔToA as a function of the signal amplitude for different detectors.

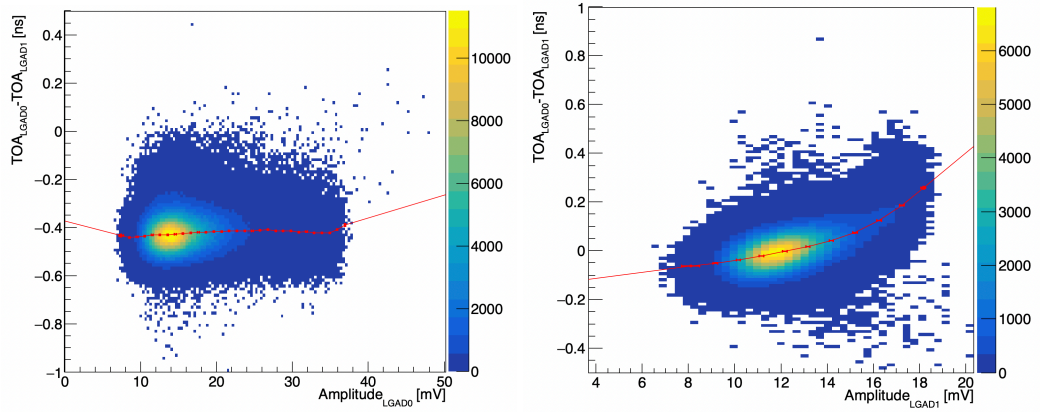


FIGURE 4.26: Distributions of the ΔToA between LGAD0 and LGAD1 as a function of the signal amplitudes measured by LGAD0 (left) and LGAD1 (right). A Gaussian fit in each bin of the horizontal axis gave the time-walk correction points (in red). The red segments show the linear interpolation between the time-walk correction points used to correct the data. The average ΔToA contains an arbitrary offset irrelevant to the time resolution measurement.

An event-by-event correction was applied to mitigate the time-walk effect on the measured time resolution. For each of the bins of the horizontal axes of the plots in Figure 4.26 and Figure 4.27, the most probable value of ΔToA was extracted via a Gaussian fit (red points in the figures), and associated to the average signal amplitude for that bin. A linear interpolation between adjacent bins (red lines in the figures) was then used based on the correction factor for each event's ΔToA value, according

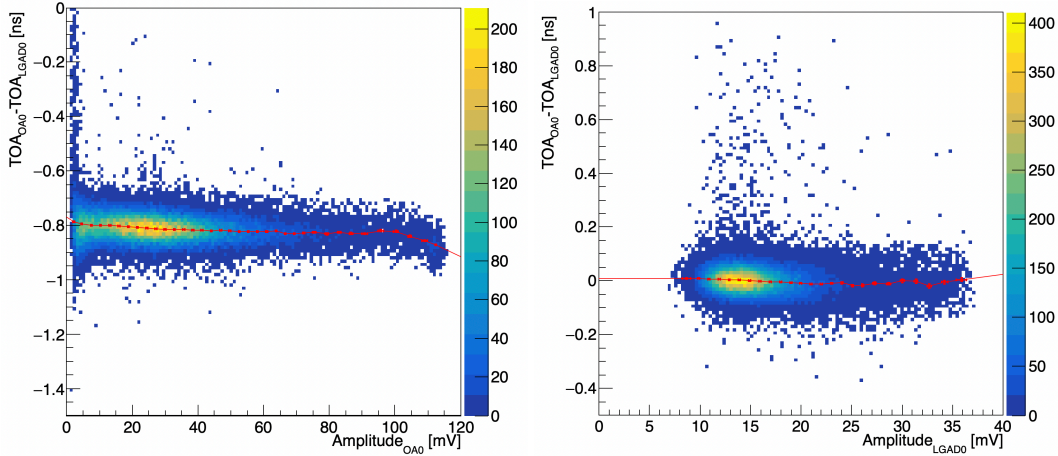


FIGURE 4.27: Distributions of the ΔToA between the "CFD-like" signals in the pixel OA0 of the DUT and the signals acquired by LGAD0 as a function of the DUT amplitude (left) and LGAD0 amplitude (right). The DUT was operated at a power density of 2.7 W/cm^2 (corresponding to $I_{\text{preamp}} = 150 \mu\text{A}$) and $HV = 125 \text{ V}$. A Gaussian fit was performed on each bin of the horizontal axis to determine the time-walk correction points (in red). The red segments show the linear interpolation between the time-walk correction points used to correct the data. The average ΔToA contains an arbitrary offset irrelevant to the time resolution measurement.

to its signal amplitude. More precisely, a first correction was determined for the first detector in the pair associated with ΔToA and then applied; subsequently, a second correction was determined for the second detector and used to obtain the final ΔToA distribution corrected for time walk. It was verified that the correction order is irrelevant, as expected for two independent detectors.

While the time-walk correction for the LGADs was always computed using their corresponding signal amplitudes, different approaches were studied for the PicoAD prototype.

- A first-time-walk correction relied solely on the DUT signal amplitudes, as adopted in [58].
- A second approach consisted of shifting by 200 ps (approximately half of the signal rise time) the waveform from the DUT pixel, subtracting it from the original waveform, and taking as ToA the time at 25 % height of the produced waveform. This method, which mimics at some level the working principle of a constant fraction discriminator (CFD) and is referred to here as "CFD-like," avoids time walk between large and small amplitude signals while at the same time reducing the small frequency-noise components that are automatically suppressed when subtracting the original and time-shifted signals to generate the CFD-like signal. The small residual time walk was then corrected using the DUT amplitude of the new waveform and then that of the LGAD, as shown in Figure 4.27.
- Finally, a hybrid approach that used the amplitude to correct for time walk signals with amplitudes smaller than 25 mV and the signal rise time between 10 mV and 20 mV for signals larger than 25 mV was also utilized. This method yielded results that were compatible with the uncertainties obtained with the CFD-like method.

The CFD-like approach was adopted to obtain the results shown below.

Estimation of the Time Jitter Between the Two Oscilloscopes

An additional time jitter affects the ΔToA measured between the two LGADs, connected to the first oscilloscope, and pixels OA1, OA2, and OA3, connected instead to the second oscilloscope. This extra jitter σ_{scopes} stems from, for instance, the different bandwidth and sampling rate of the two oscilloscopes and the longer and lower-quality cables used during the test beam measurements to connect the detectors to the second oscilloscope. Conversely, if read by the same oscilloscope as the LGADs, pixel OA0 does not exhibit this jitter. The corresponding degradation in the measured time resolution for DUT pixels OA1, OA2, and OA3 was estimated using the two SPADs [115]. The distributions of the ΔToA between LGAD0 and the SPAD0, which was read by the first oscilloscope, and between LGAD0 and the SPAD1, which was read by the second oscilloscope, were fit with a Gaussian function, as illustrated in Figure 4.28.

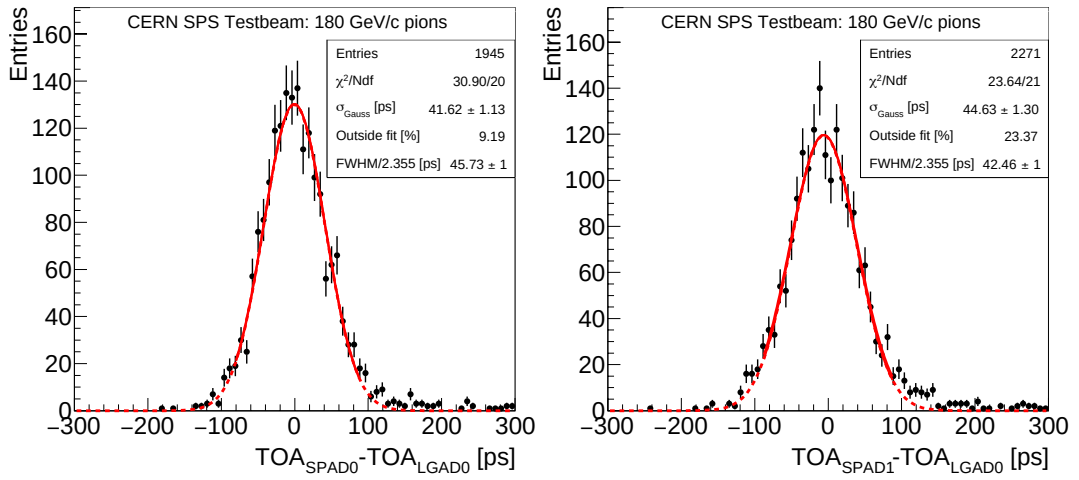


FIGURE 4.28: ΔToA between SPAD0 and LGAD0 (left), and between SPAD1 and LGAD0 (right) after time-walk correction for the LGAD0. The red lines show the results of a Gaussian fit to each distribution: the full line extends across the fit range considered, while the dashed line extrapolates the fitted function to the whole histogram.

The jitter produced by the two oscilloscopes was then estimated by subtracting, in quadrature, the two fitted σ_{Gauss} parameters, under the assumption of an identical time response between the two SPADs. This procedure provided an estimated additional contribution of $\sigma_{\text{scopes}} = (16.1 \pm 4.6)$ ps from the second oscilloscope to the measured time resolution of DUT pixels OA1, OA2, and OA3.

Methods to Measure the Time Resolution

To measure the time resolution, a Gaussian fit was performed simultaneously to all the ΔToA distributions obtained with the four DUT pixels and the two LGADs after the time-walk corrections, minimizing a single χ^2 function to extract the four DUT pixels resolutions $\sigma_{\text{DUT } 0, 1, 2, 3}$ and the LGAD resolutions σ_{LGAD0} and σ_{LGAD1} directly. The σ_{Gauss} parameters of the Gaussian functions describing the ΔToA distributions in the fit were computed using equations 4.2. Each pixel was assigned a separate time resolution parameter for the device under test (DUT).

In a second method, used as a cross-check, the time resolution of each of the four DUT pixels was estimated separately. For each DUT pixel, a Gaussian fit to the three ΔToA distributions was deployed to extract the $\sigma_{\Delta\text{ToA}(\text{DUT}, \text{LGAD0})}$, $\sigma_{\Delta\text{ToA}(\text{DUT}, \text{LGAD1})}$,

and $\sigma_{\Delta\text{ToA}(\text{LGAD0}, \text{LGAD1})}$ terms in the left-hand side of Equations 4.2. The DUT and LGAD time resolutions were then computed after solving the system. The results obtained with the two methods were consistent with statistical expectations.

In all cases, the Gaussian fits were restricted to the bulk of each ΔToA distribution by considering only bins with more entries than 25 % of the entries in the most populated bin. For all working points analyzed, the fraction of events exceeding the Gaussian fit integral was always below 5 %, indicating that non-Gaussian contributions to the timing resolution of the PicoAD proof-of-concept prototype have an overall small impact and that the resolutions quoted in the following refer to at least 95 % of the signals acquired.

As an example, the three time-walk corrected ΔToA distributions between DUT pixel OA0, LGAD0, and LGAD1 are shown in Figure 4.29 in the case the PicoAD prototype was operated at a power density of 2.7 W/cm^2 and $HV = 125 \text{ V}$.

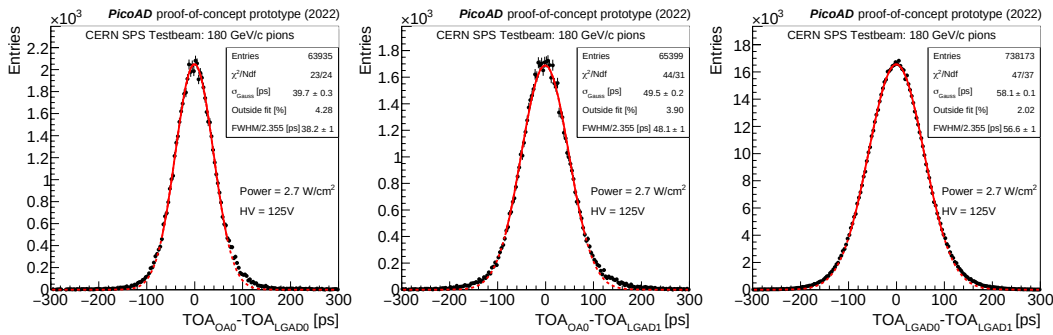


FIGURE 4.29: ΔToA difference between pixel OA0 of the DUT and LGAD0 (left), pixel OA0 of DUT and LGAD1 (center), and between LGAD0 and LGAD1 (right) after time-walk correction for the working point with $I_{\text{preamp}} = 150 \mu\text{A}$ (corresponding to a power density of 2.7 W/cm^2) and $HV = 125 \text{ V}$. The red lines show the results of a Gaussian fit to each distribution: the full line extends across the fit range considered, while the dashed line extrapolates the fitted function to the whole histogram.

The ΔToA distribution between LGAD1 and the DUT is larger than that involving LGAD0 and the DUT, indicating that LGAD0 has a better time resolution than LGAD1. In addition, the ΔToA distribution between the two LGADs is much larger than the other two ΔToA distributions that contain the DUT, which shows that the DUT has a much better resolution than the two LGAD detectors used for this measurement. The results of the Gaussian fit consist indeed of resolutions of $\sigma_{\text{DUT}} = (17.3 \pm 0.4) \text{ ps}$ for pixel OA0 of the DUT, $\sigma_{\text{LGAD0}} = (33.5 \pm 0.1) \text{ ps}$ for LGAD0 and $\sigma_{\text{LGAD1}} = (41.7 \pm 0.1) \text{ ps}$ for LGAD1.

The results for the four analog PicoAD pixels at this working point are summarized in Table 4.3 for different conditions:

- The second column of the table shows the results in the case in which the DUT was not time-walk corrected at all (while the LGADs were corrected). Even in this extreme case, the four DUT pixels can provide a time resolution of 50 ps, which is very promising for applications with, e.g., many readout channels.
- In the case in which the time walk was corrected using the signal amplitudes acquired with the oscilloscopes (third column of the table), the measured time resolutions vary between 24 ps in the case of pixel OA0 that was read by the same oscilloscope of the LGADs, and approximately 30 ps for the other three pixels that were instead read out by the second oscilloscope and therefore are subject to the σ_{scopes} jitter discussed in the next section.

- Finally, in the case of the CFD-like time-walk correction explained in the previous section, a time resolution of (17.3 ± 0.4) ps was measured for pixel OA0. The other three pixels, read instead by the second oscilloscope and therefore suffering from the additional time jitter $\sigma_{\text{scopes}} = (16.1 \pm 4.6)$ ps, have a time resolution ranging from 24.4 ps to 25.3 ps. Subtraction in quadrature of σ_{scopes} provides the time-resolution values for pixels OA1, OA2, and OA3 reported in the last column of the table; these values are compatible with the result of pixel OA0, although within the relatively-large uncertainties coming from the limited statistics that could be collected with the SPADs because of their small surface.

Time resolution σ_t [ps] for the DUT operated at $P_{\text{density}} = 2.7 \text{ W/cm}^2$ and $HV = 125 \text{ V}$					
DUT pixel	Time-walk correction			P_{density}	$\sqrt{\sigma_{t,\text{CFD}}^2 - \sigma_{\text{scopes}}^2}$
	no corr.	ampl.-based	CFD-like		
OA0	49.3 $\pm$ 1.5	24.1 $\pm$ 0.4	17.3 $\pm$ 0.4	-	17.3$\pm$0.4
OA1	49.7 $\pm$ 1.5	30.5 $\pm$ 0.3	25.3 $\pm$ 0.3	16.1 $\pm$ 4.6	19.5$\pm$3.8
OA2	47.1 $\pm$ 1.7	27.6 $\pm$ 0.3	24.4 $\pm$ 0.3	16.1 $\pm$ 4.6	18.3$\pm$4.1
OA3	47.5 $\pm$ 1.6	29.4 $\pm$ 0.3	24.8 $\pm$ 0.3	16.1 $\pm$ 4.6	18.9$\pm$3.9

TABLE 4.3: Time resolution of the four analog pixels of the DUT for the working point at a power density of 2.7 W/cm^2 and $HV = 125 \text{ V}$ for three cases: no time-walk correction; time-walk correction based on the signal amplitude; CFD-like time-walk correction. The last two columns show the time jitter σ_{scopes} between the two oscilloscopes measured using the SPADs, and the time resolution of pixels OA1, OA2, and OA3 obtained after σ_{scopes} is subtracted in quadrature from the time resolution obtained with the CFD-like time-walk correction, respectively. Only the statistical uncertainties are reported.

The time resolutions in Table 4.3 are obtained by constraining the resolution of the DUT using those of the LGADs that, for some working points, result in being worse than the DUT resolution. To verify that this method does not produce a bias, we have repeated the data analysis using the pixel OA0 of the DUT together with the LGAD0 and substituting the LGAD1, which provides worse time resolution than the LGAD0, with the SPAD0 that is read out by the same scope and therefore is not affected by σ_{scopes} . Given the much smaller active area of the SPAD compared to the LGAD active area, the data sample with the SPAD0 in coincidence with the DUT and the LGAD0 is much smaller. Therefore, the time resolution measurement suffers from a larger statistical uncertainty.

Figure 4.30 shows the distributions of the ΔToA between the DUT pixel OA0 and SPAD0 and between the LGAD0 and SPAD0.

The time resolution for the DUT pixel OA0, measured using this method, is (17.1 ± 3.5) ps, which is compatible with the statistical uncertainties of the measurement obtained with the two LGADs.

Bias Voltage and Power Density Scans

Table 4.4 summarizes the efficiency and time resolutions obtained at the highest sensor bias $HV = 125 \text{ V}$ for the three values of power density P_{density} considered during the test beam.

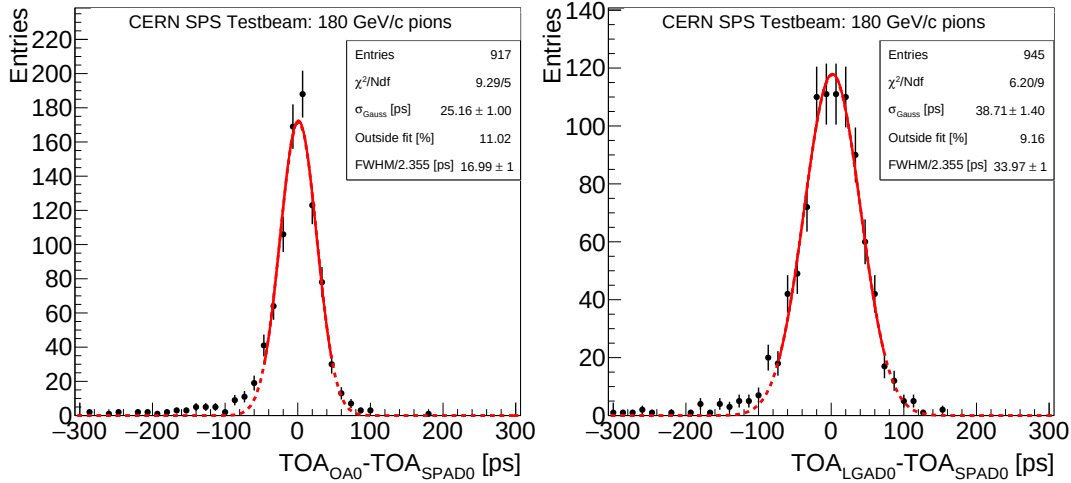


FIGURE 4.30: ΔToA between pixels OA0 of the DUT and SPAD0 (left), and between LGAD0 and SPAD0 (right), after time-walk correction for the working point with a power density of 2.7 W/cm^2 and $HV = 125 \text{ V}$. The red lines show the results of a Gaussian fit to each distribution: the full line extends across the fit range considered, while the dashed line extrapolates the fitted function to the whole histogram.

I_{preamp} [μA]	P_{density} [W/cm^2]	Efficiency [%]	Time Resolution [ps]
20	0.4	$99.90^{+0.04}_{-0.05}$	30.2 ± 0.3
50	0.9	$99.95^{+0.04}_{-0.09}$	25.1 ± 0.7
150	2.7	$99.88^{+0.04}_{-0.05}$	17.3 ± 0.4

TABLE 4.4: Power consumption per unit surface, average efficiency calculated in the two triangles of Figure 4.20 and pixel OA0 time resolution measured at $HV = 125 \text{ V}$ for the three preamplifier current values reported in the first column.

Figure 4.31 left shows the time resolution for the DUT pixel OA0 measured at $HV = 125 \text{ V}$ as a function of the amplifier power density P_{density} . The trend indicates a progressive small deterioration of the timing performance of $(17.3 \pm 0.4) \text{ ps}$ measured at $P_{\text{density}} = 2.7 \text{ W/cm}^2$ when the power consumption is reduced. However, even at 0.4 W/cm^2 , the time resolution remarkably remains 30 ps.

The right panel of Figure 4.31 reports the results of the scan in sensor bias voltage at $P_{\text{density}} = 2.7 \text{ W/cm}^2$. The PicoAD proof-of-concept prototype provides a resolution of $(28.4 \pm 0.7) \text{ ps}$ at 105 V, which demonstrates that the detector can be operated at time resolutions better than 30 ps with a 20 V plateau in sensor bias voltage.

Time Resolution in Different Regions of the Pixel

The time resolution was also studied as a function of the distance from the pixel center. Figure 4.32 left shows the results when $HV = 125 \text{ V}$ for the three values of P_{density} considered, while Figure 4.32 right shows the results obtained for $HV = 105, 115$ and 125 V when $P_{\text{density}} = 2.7 \text{ W/cm}^2$. For the working point $HV = 125 \text{ V}$ and $P_{\text{density}} = 2.7 \text{ W/cm}^2$, the data indicate a time resolution of 13 ps at the center of the pixel and 25 ps for signals associated with tracks that point to the edge of the pixel. Similar relative trends between the center and the edge of the pixel are observed for the other working points studied.

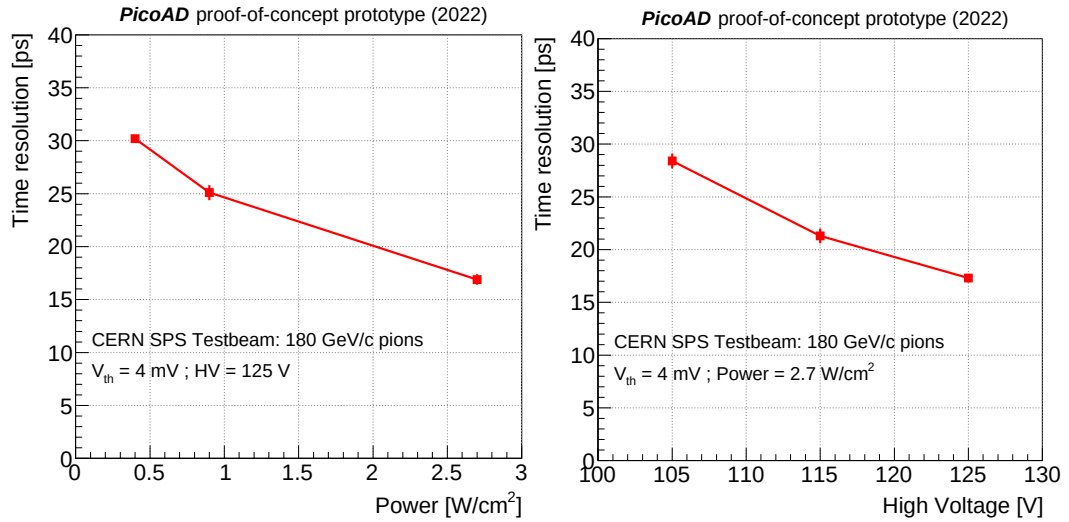


FIGURE 4.31: Time resolution for the pixel OA0 of the DUT as a function of the power density at sensor bias voltage $HV = 125$ V (left panel), and as a function of sensor bias voltage at a power density of 2.7 W/cm² (right panel).

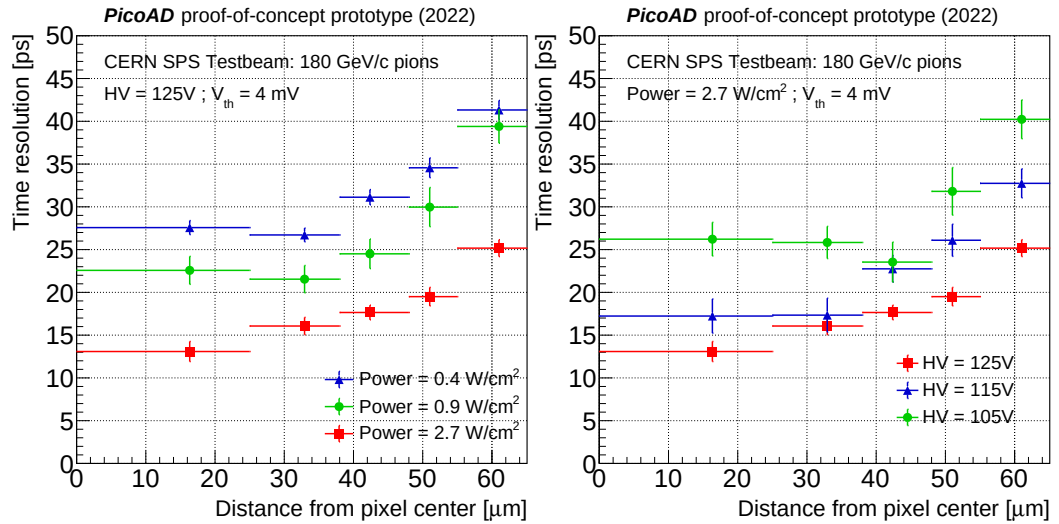


FIGURE 4.32: Time resolution for the DUT pixel OA0 as a function of the distance from the pixel center. In the left panel, the resolution is measured at $HV = 125$ V for the three values of the power density considered, while in the right panel, it is measured at a power density of 2.7 W/cm² for the three values of HV . In each bin, the data points are positioned at the mean distance from the center of the pixel.

Reduced signal amplitudes explain the lower time resolution at the pixel edge. A dedicated TCAD simulation (see Figure 4.3) shows that these lower signals result from charge sharing between adjacent pixels, partly caused by pions crossing the inter-pixel regions of the DUT, and partly from a reduced multiplication gain near the pixel edges compared to the center.

Chapter 5

MONOLITH Prototype 2 without a Gain Layer

This chapter presents a detailed characterization of the second MONOLITH prototype, manufactured in 2022, without a gain layer, and evaluates its performance using various measurement techniques.

Section 5.1 covers laboratory measurements performed before test beam and laser measurements. We describe energy-calibration runs with radioactive sources, whose spectra are analyzed to extract the preamplifier gain, the Equivalent Noise Charge (ENC), and the signal rise time of the analog front-end chain. We show some scans of these parameters as a function of the preamplifier and the feedback currents.

Section 5.2 describes the measurements obtained with a test-beam campaign. There, minimum-ionizing particles are used to probe the detector's efficiency and time resolution under realistic particle-tracking conditions. We outline the instrumentation and then define the data samples collected at different bias voltages and front-end electronics power densities. The efficiency and timing performance are benchmarks against which the latest prototype with a gain layer will be compared.

Section 5.3 turns to precision laboratory studies using a femtosecond laser to inject controlled charge pulses directly into the silicon active layer. We discuss the optical setup and alignment procedure, and then detail the characterization of the SiGe HBT front-end ASICs. By varying the laser pulse energy, we reconstruct the dependence of time resolution on deposited charge and isolate the contribution of charge-collection noise. These measurements complement the test-beam results by decoupling the electronic time jitter effects from the intrinsic limit to time resolution due to fluctuations in charge cluster creation.

Section 5.5 presents the Garfield++ simulation studies, which revealed fundamental tools to explain some of the seen behaviour of the experimental data, such as the charge-collection noise contribution to the time resolution and the small non-Gaussianity of the time of arrival distributions.

Finally, Section 5.4 quantifies the radiation tolerance of the ASICs when exposed to proton fluences up to $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$. We report on post-irradiation changes in noise, signal slope, and signal amplitude, and derive the resulting impact on signal jitter from laboratory measurements. Then, we validate these findings with a second test-beam campaign. Together, these studies establish the high radiation tolerance of these MONOLITH prototypes, which is directly derived from the high radiation tolerance of SiGe HBTs.

Some of the results presented in this chapter have been published previously. In particular, test beam results are presented in [4]. The laboratory measurements with a femtosecond laser are shown in [5]. The laboratory measurements with irradiated ASICs are presented in [6], and the test-beam measurements on the same ASICs are presented in [7].

5.1 Laboratory Measurements

This section presents the laboratory measurements done to assess the performance of the ASICs of the second MONOLITH prototype without a gain layer. The ASICs were received in January 2022. During the first months of 2022, ten ASICs were assembled on correctly printed circuit boards (PCBs; see Section 1.2.8) and intensively tested in the cleanroom laboratory at the University of Geneva.

First, we describe the experimental setup, which included the custom test board that interfaces the ASIC to precision power supplies and an oscilloscope. Special emphasis is placed on defining the radioactive sources used.

Then, we detail the core ASIC performance metrics, which are the *preamplifier gain*, the *Equivalent Noise Charge* (ENC), and the *signal rise time*. These three figures of merit were mapped in the whole parameter space of the analog front-end electronics given by the power density P_{density} and the feedback current I_{fbk} .

Finally, we present additional measurements taken immediately before the tests with the femtosecond laser and the Garfield++ simulations. These two studies require knowledge of performance metrics at specific working points and with specific ASIC samples.

5.1.1 Experimental Setup

The experimental setup consisted of the ASIC to be tested, a fast oscilloscope (Lecroy WaveMaster 820 Zi-A), typically set to 40 GSamples/s and a 4 GHz analog bandwidth, and several power supplies providing low and high bias voltages. The high bias voltage HV was always set to 200 V. The ASIC is interfaced to the fast oscilloscope and power supplies through a properly designed Printed Circuit Board (PCB), as explained in Section 1.2.8.

Laboratory measurements involved ^{55}Fe and ^{109}Cd radioactive sources. The first is a soft X-ray source, and the second is a gamma-ray source, as explained in Section 3.2.2. These two radioactive sources were used due to the known and precise value of the deposited charge within the silicon active layer. The deposited charge was known, allowing for the evaluation of the core ASIC's performance metrics.

5.1.2 ASICs Characteristics

The core ASIC performance metrics are the *preamplifier gain*, the *Equivalent Noise Charge* (ENC), and the *signal rise time*. They were measured using ^{55}Fe and ^{109}Cd radioactive sources as a function of the preamplifier current I_{preamp} and the feedback current I_{fbk} . As already mentioned in Section 3.7, the preamplifier current I_{preamp} sets the gain-bandwidth product of the HBT, while the feedback current I_{fbk} sets the working point in the gain-bandwidth curve determined by the value of I_{preamp} . The parameter space defined by these two currents, which can be controlled via software, was entirely mapped.

It is worth noting that the 10 ASICs assembled on PCBs did not always yield compatible results, with a maximum discrepancy of about 10 % to 20 % when comparing performance metrics for different ASICs under the same experimental conditions. The discrepancy can be mainly attributed to mismatches in the transistors of the front-end electronics, which resulted in a mismatch of the bias current settings. Despite that, other relevant parameters, such as signal-to-noise ratio and time jitter (a combination of the performance metrics), were compatible with experimental errors. The rule of thumb is again associated with the gain-bandwidth product, which must be constant

for a fixed power density. If a first ASIC had a higher gain than a second ASIC, it surely had a slower rise time, such that the overall time jitter remained the same.

This section presents the core ASIC performance metrics for a single ASIC that exhibits average behavior. The laboratory characterization was performed again before other relevant studies, such as test beam and femtosecond laser measurements, or Garfield++ simulations, using the same ASIC used for these studies.

The preamplifier current, I_{preamp} , is primarily responsible for the power consumption of the analog front-end electronics. The values of the preamplifier current I_{preamp} that were used were $167 \mu\text{A}$, $55 \mu\text{A}$, $20 \mu\text{A}$, and $7.5 \mu\text{A}$, which correspond to a power density P_{density} of 2.7 W/cm^2 , 0.9 W/cm^2 , 0.3 W/cm^2 , and 0.1 W/cm^2 , respectively. The power density was calculated using Equation 3.27.

Preamplifier Gain

The gain of the analog front-end electronics preamplifier A_q is defined as the ratio of the output voltage swing to the injected charge. When using the ^{55}Fe and ^{109}Cd radioactive sources, the value of injected charge is known with precisions: ^{55}Fe produces around 1640 localized electron-hole pairs, i.e., a charge of $Q_{^{55}\text{Fe}} = 0.26 \text{ fC}$, and ^{109}Cd produces around 6944 localized electron-hole pairs, i.e., a charge of $Q_{^{109}\text{Cd}} = 1.11 \text{ fC}$. Therefore, to measure the preamplifier gain, it is necessary to measure the voltage signal amplitude of the analog signal registered by the oscilloscope.

For each configuration of power density P_{density} and feedback current I_{fbk} , a set of about 10k waveforms from the differential analog output of one pixel were saved from the oscilloscope and then analyzed to find the signal amplitude for each event. The distribution of these amplitudes was fitted with a Gaussian function, yielding excellent agreement. The mean value was then saved for each configuration. The same procedure was repeated for both ^{55}Fe and ^{109}Cd radioactive sources, finding the average amplitude for the two sources: $\text{amp}^{^{55}\text{Fe}}$ and $\text{amp}^{^{109}\text{Cd}}$. The preamplifier gain A_q was then calculated as:

$$A_q = \frac{\text{amp}^{^{109}\text{Cd}} - \text{amp}^{^{55}\text{Fe}}}{Q_{^{109}\text{Cd}} - Q_{^{55}\text{Fe}}}. \quad (5.1)$$

Figure 5.1 summarizes the results of the preamplifier gain A_q as a function of the feedback current I_{fbk} for different values of power density P_{density} .

Figure 5.1 shows that the preamplifier gain A_q increases with increasing power consumption. Moreover, the preamplifier gain A_q has a specific trend, most notable for the highest power consumption $P_{\text{density}} = 2.7 \text{ W/cm}^2$, that consists in a fast rise at low feedback, then a maximum, and then a slower fall that reflects the inverse relationship between feedback strength and preamplifier gain: stronger feedback lowers the closed-loop gain and thus reduces the observed preamplifier gain. The peak is not visible for the other power consumptions as it falls at feedback currents that are too low.

The power consumption $P_{\text{density}} = 0.1 \text{ W/cm}^2$ could be measured up to a feedback current $I_{\text{fbk}} = 100 \text{ nA}$, because then the signal-to-noise ratio for the ^{55}Fe source became less than 7, meaning that the usual criteria for the signal amplitude to be higher than $7\sigma_V$ could not be applied, hence the ^{55}Fe source could not be distinguished from the noise background.

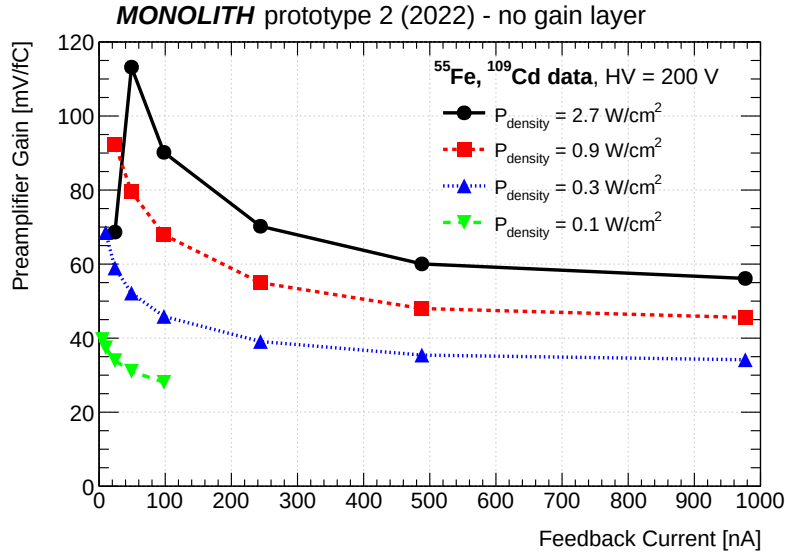


FIGURE 5.1: Preamplifier gain A_q as a function of the feedback current I_{fbk} for different values of power density P_{density} .

Equivalent Noise Charge

The Equivalent Noise Charge (ENC) is the charge in electrons that one has to feed to the preamplifier to have a voltage output whose amplitude is one standard deviation of the noise σ_V . The ENC is a standardized method for quantifying noise when comparing different amplification systems.

From the same datasets of waveforms that were used for the measurements of the preamplifier gain A_q , the standard deviation of the noise σ_V was calculated in the first half of the waveform, where no signal was present. The oscilloscope contributed a not-always-negligible contribution σ_{scope} to the background noise, hence the measured σ_V is the sum in quadrature between the pure front-end electronics contribution $\sigma_{V,\text{FE}}$ and the oscilloscope contribution σ_{scope} . The oscilloscope's contribution must be subtracted from the total to compare different amplification systems. Therefore, the formula used to calculate the corrected ENC corrected from the oscilloscope contribution was:

$$\text{ENC} = \frac{\sigma_{V,\text{FE}}}{A_q} = \frac{\sqrt{\sigma_V^2 - 2\sigma_{\text{scope}}^2}}{A_q}, \quad (5.2)$$

where σ_{scope} is the voltage noise of the oscilloscope with the open input connector, taken twice to account for the signals of the two polarities of the differential front-end electronics.

Figure 5.2 summarizes the results of the ENC corrected for the oscilloscope as a function of the feedback current I_{fbk} for different values of power density P_{density} .

Figure 5.2 illustrates that the equivalent noise charge (ENC) reaches a clear minimum before rising gently to a plateau. The lowest ENC of (89 ± 2) electrons is measured at $P_{\text{density}} = 0.9 \text{ W/cm}^2$ with a feedback current $I_{\text{fbk}} = 50 \text{ nA}$. This operating point is rarely used in practice because the ENC would be too highly sensitive to small changes in I_{fbk} . Instead, one typically chooses a larger feedback current so the ENC sits firmly on its plateau. Although the plateau ENC is somewhat higher, the boosted feedback current speeds up the front-end response, and, as a result, mitigates

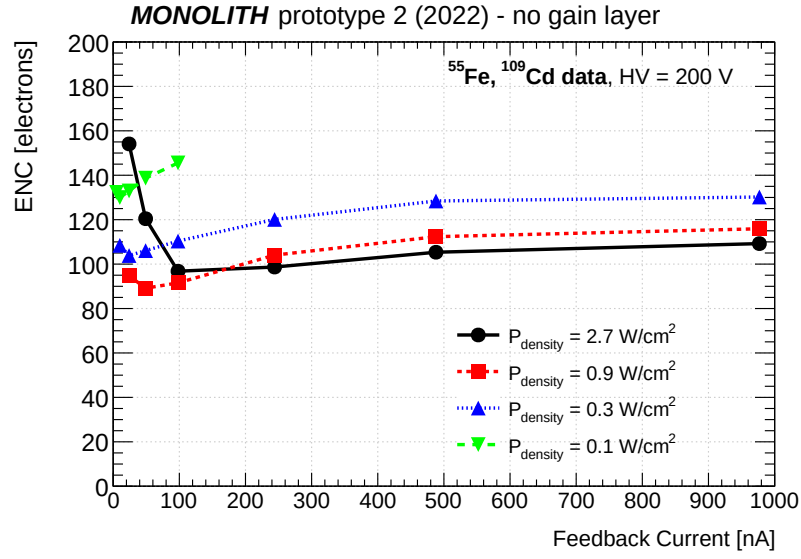


FIGURE 5.2: Equivalent Noise Charge (ENC) corrected for the oscilloscope contribution as a function of the feedback current I_{fbk} for different values of power density P_{density} .

other contributors to time resolution, such as charge-collection noise and device non-uniformities.

Rise Time

The signal's rise time is directly related to the speed of the front-end electronics, which is associated with the amplifier's bandwidth. The signal rise time is defined as the time that it takes for a signal to move from 20 % to 80 % of its amplitude. From the same datasets of waveforms that were used for the measurements of the preamplifier gain A_q and the ENC, the rise time was calculated for each waveform. The values were distributed and fitted with a Gaussian to extract the average value.

Figure 5.3 summarizes the results of the average signal rise time as a function of the feedback current I_{fbk} for different values of power density P_{density} . We present only the results obtained with the ^{55}Fe radioactive source, but the results for ^{109}Cd are analogous.

Figure 5.2 illustrates that the signal rise time enters a plateau when the feedback current is larger than about 200 nA, with a minimum value of (270 ± 13) ps when $P_{\text{density}} = 2.7 \text{ W/cm}^2$. Typically, the working point falls within this plateau of feedback current for the same reasons explained above when discussing the ENC results.

5.1.3 Characterization Before the Femtosecond Laser Measurements

Two ASICs from the second MONOLITH prototype, without a gain layer, were used for the femtosecond laser measurements (see Section 5.3). They were characterized intensively using the same technique described above. The calibrations of the preamplifier gain A_q , and the Equivalent Noise Charge (ENC) of the two MONOLITH ASICs for different power densities are reported in Table 5.1.

A feedback current of $I_{\text{fbk}} = 100 \text{ nA}$ was selected to strike a balance between preamplifier gain and speed: this modest value yields higher gain and lower ENC, at the cost

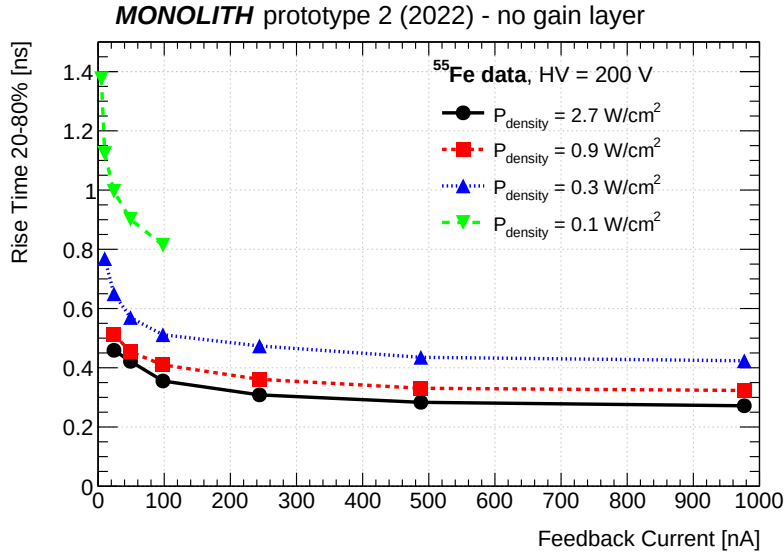


FIGURE 5.3: Signal rise time from 20 % to 80 % of the signal amplitude as a function of the preamplifier current I_{preamp} and the feedback current I_{fbk} .

ASIC	$P_{\text{density}} [\text{W/cm}^2]$	$\sigma_V [\text{mV}]$	$A_q [\text{mV/fC}]$	ENC [electrons]
DUT	2.7	1.58 ± 0.04	102.1 ± 1.2	96 ± 1
DUT	0.9	1.06 ± 0.02	67.1 ± 0.9	96 ± 1
DUT	0.36	0.77 ± 0.02	42.7 ± 0.6	108 ± 2
Reference	2.7	1.13 ± 0.03	74.7 ± 1.1	93 ± 1

TABLE 5.1: Charge gain and ENC measured with ^{55}Fe and ^{109}Cd source for DUT and reference ASICs of the femtosecond laser measurements at different values of the front-end power density. The standard deviations of the voltage and oscilloscope noise used for calculating the ENC are also reported. The feedback current was $I_{\text{fbk}} = 100 \text{ nA}$.

of a marginally slower rise time. Since the laser injection method eliminates charge-collection noise, there is no requirement for the fastest possible front-end response. Instead, the priority shifts to achieving a stable, well-calibrated conversion between output amplitude and injected charge, which benefits directly from the increased gain and reduced noise. These considerations are critical for the femtosecond-laser measurements discussed in Section 5.3.

5.1.4 Characterization Before the Garfield++ Simulations

The first objective of the Garfield++ simulations is to match the results of time resolution collected at a test beam. The best result of time resolution with the second MONOLITH prototype without a gain layer was achieved during the second test beam. In contrast, the irradiated ASICs were tested (see Section 5.4.7). The result is $(20.0 \pm 1.0) \text{ ps}$ and it was achieved with a reduced power consumption of $P_{\text{density}} = 0.9 \text{ W/cm}^2$, a high feedback current of $I_{\text{fbk}} = 2000 \text{ nA}$, and a low temperature of $T = -10^\circ\text{C}$.

In these working conditions, the core ASIC performance metrics were measured in the laboratory, giving:

- a preamplifier gain A_q equal to (56 ± 2) mV/fC;
- an ENC equal to (130 ± 5) electrons (not corrected for the oscilloscope contribution);
- a signal rise time equal to (364 ± 4) ps.

All the time resolutions reported in the following sections include the contribution from the oscilloscope. Although the digitization contribution to the time resolution is negligible, as discussed in Section 3.5.4, the influence of the oscilloscope noise may not always be insignificant. For the specific working point considered here, the oscilloscope contribution is indeed small; however, for clarity, it is essential to state that the quoted ENC of (130 ± 5) electrons includes the oscilloscope's contribution.

The second objective of the Garfield++ simulations is to disentangle the individual contributions to the time resolution arising from electronics time jitter and charge-collection noise, thereby providing essential validation of the femtosecond laser measurements. For this comparison, the reference time resolution is taken from the first test beam, which yielded a time resolution of (20.7 ± 0.3) ps. This measurement was obtained under operating conditions of high power consumption ($P_{\text{density}} = 2.7 \text{ W/cm}^2$), a feedback current of $I_{\text{fbk}} = 100 \text{ nA}$, and a room temperature of $T = 20^\circ \text{C}$. Despite the higher power consumption and elevated temperature, the lower feedback current compensated for these factors, ultimately achieving a time resolution comparable to that observed in earlier operating points.

In these working conditions, the core ASIC performance metrics were measured in the laboratory, giving:

- a preamplifier gain A_q equal to (60 ± 2) mV/fC;
- an ENC equal to (124 ± 5) electrons (not corrected for the oscilloscope contribution);
- a signal rise time equal to (375 ± 4) ps.

5.2 Test Beam

This section presents the measurements of the second MONOLITH prototype without a gain layer performed during a CERN SPS test beam with pions with momentum $120 \text{ GeV}/c$. The measurements were conducted in the experimental *North Area* at the Prevezin site of CERN during October and November 2022.

First, we present the experimental setup used for the test beam, which includes a telescope to track the incoming pions and fast detectors to provide a time reference. The telescope is the same as the one used for the test-beam campaign of the PicoAD proof-of-concept prototype, but the time references are faster (see Section 4.3.1).

Second, we present the efficiency and noise-hit ratio results, along with their dependence on the particle's position. The results show that the efficiency of this prototype is similar to that of the PicoAD proof-of-concept prototype despite not having a gain layer. Moreover, the fabrication choices for substrate doping and inter-pixel design contributed to a significantly more uniform response as a function of distance from the pixel center.

Finally, we present the results on time resolution after a time-walk correction is applied to the data. As for efficiency, the time resolution is also mapped into the available parameter space of bias voltage (HV) and power density (P_{density}).

5.2.1 Experimental Setup

A pion beam with a momentum of $120 \text{ GeV}/c$ was delivered at the CERN SPS test beam facility to characterize the MONOLITH prototype 2 ASIC and measure its detection efficiency and time resolution.

The experimental setup is illustrated in Figure 5.4. The device under test (DUT), wire-bonded on a dedicated printed-circuit board, was installed in the middle of the six detection planes of the UniGe FEI4 telescope for charged-particle tracking [114]. Two commercial MCP detectors, named MCP0 and MCP1 hereafter, with expected time resolution for minimum-ionizing particles of approximately 5 ps, were installed downstream of the telescope to provide an exact time reference.

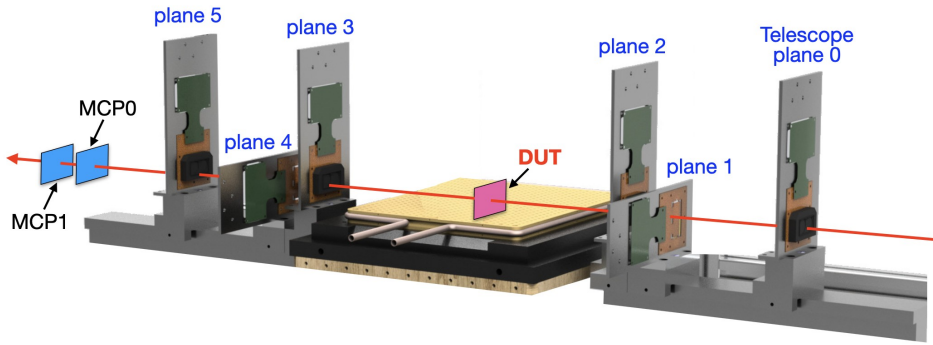


FIGURE 5.4: Schematic view of the experimental setup, showing the six planes of the UNIGE FEI4 telescope, the DUT, and the two MCPs installed downstream of the telescope in the pion beam direction.

Two oscilloscopes read analog signals from all detectors. A first oscilloscope with a sampling rate of 40 GSamples/s and an analog bandwidth set to 4 GHz was used to read the two MCP detectors and the two differential outputs from pixel OA0 of the DUT. The positive single-ended outputs of pixels OA1, OA2, and OA3 of the DUT were instead connected to a second oscilloscope with a 4 GHz analog bandwidth and a sampling rate of 20 GSamples/s. The oscilloscopes saturated the signal amplitudes to around 75 mV for all the DUT pixels.

The FEI4 telescope provided the trigger signal to the oscilloscopes. The trigger was generated by a coincidence of the region of interest of $250 \mu\text{m} \times 250 \mu\text{m}$ imposed on the first telescope plane and centered at the pixel OA0 and the last telescope plane. A window of 200 ns around the telescope trigger time was defined to record waveform data from the oscilloscopes.

Pion trajectories were reconstructed using the hits recorded by the FEI4 telescope planes, and their points of intersection with the DUT were extrapolated. Tracks were required to have hits in all six telescope planes and $\chi^2/NDF \leq 1.5$ to maximize the telescope pointing resolution. Events with more than one reconstructed track were rejected to suppress ambiguous cases. These three constraints are the same as those used for the PicoAD proof-of-concept prototype (see Section 4.3.1). The offline masking of the planes was not considered here to simplify the analysis, after verifying that it had no impact on the results. For tracks satisfying the criteria mentioned above, the telescope pointing resolution on the DUT plane was estimated in [117] to be approximately $10 \mu\text{m}$.

5.2.2 Data Samples

Large samples of data were acquired at the four sensor bias voltage values $HV = 120, 160, 200, \text{ and } 250 \text{ V}$, and the five power density values reported in Table 5.2.

The left panel of Figure 5.5 shows a typical waveform acquired for pixel OA0 of the DUT operated at power density $P_{\text{density}} = 2.7 \text{ W/cm}^2$ and bias voltage $HV = 200 \text{ V}$. The distribution of the amplitudes of the differential signal from pixel OA0 after the telescope-track selection is shown in the right panel of Figure 5.5.

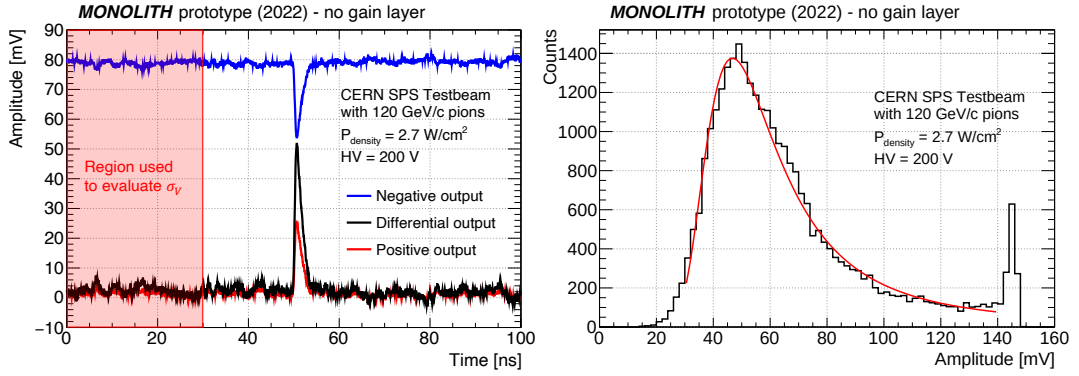


FIGURE 5.5: Characteristics of signals from pixel OA0 with the DUT operated at power density $P_{\text{density}} = 2.7 \text{ W/cm}^2$ and bias voltage $HV = 200 \text{ V}$. (Left) Example of waveforms from the single-ended outputs with the two polarities (blue and red) and the differential output (black). The region used to evaluate the noise σ_V is also shown. (Right) Distribution of the amplitudes of the differential signals; the oscilloscope's saturation generates a peak at high values. The red line results from a fit using a Landau functional form.

The analog signal amplitudes from pixel OA0 were sampled every 25 ps, and a linear interpolation was performed between consecutive samples. The voltage noise σ_V was extracted by taking the standard deviation of the differential signal samplings acquired over the time interval from 0 to 30 ns.

5.2.3 Efficiency

The pion tracks reconstructed by the FEI4 telescope and surviving the analysis selection criteria were used to measure the DUT detection efficiency. The efficiency was computed as the ratio between the number of selected tracks associated with recorded signals with amplitudes above a threshold of 7 times the voltage noise σ_V over the total number of selected tracks traversing the DUT in the area corresponding to the four analog pixels. A tolerance of $10 \mu\text{m}$ in the region outside the external boundaries of the four-pixel area was allowed to account for the finite telescope-pointing resolution.

Figure 5.6 shows the results of the efficiency measurement for the DUT operated at $P_{\text{density}} = 2.7 \text{ W/cm}^2$ and $HV = 200 \text{ V}$.

As shown in the left panel of Figure 5.6, the finite FEI4 telescope pointing resolution introduces an apparent inefficiency near the 14 external edges of the four analog pixels, making it challenging to evaluate the efficiency across the total pixel area. Since this effect is not present in the five internal edges of the four pixels, the measurement was repeated by selecting only tracks extrapolated within the two triangles connecting the OA0, OA2, and OA3 pixel centers, as well as the OA0, OA2, and OA1 pixel centers. The result is shown in the right panel of Figure 5.6, in which the color scale starts at 95 %.

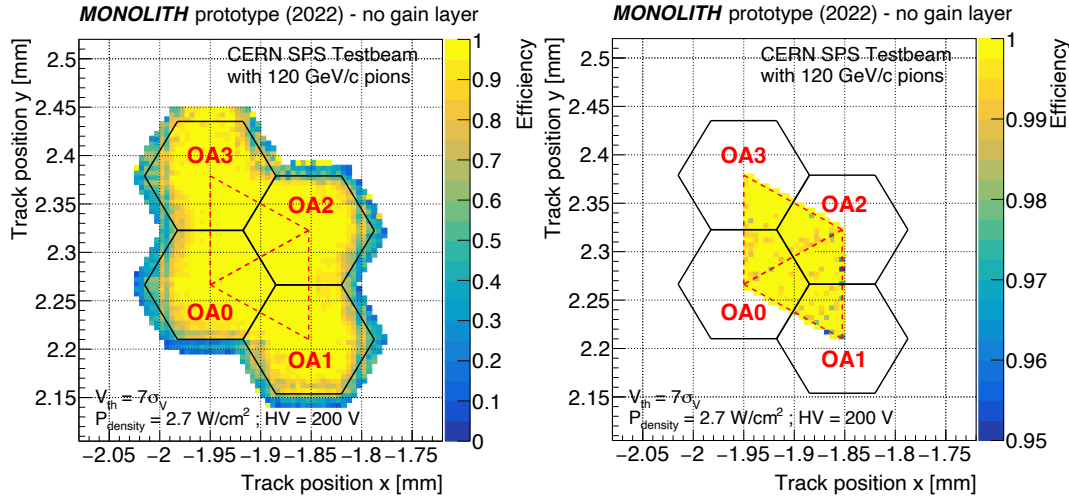


FIGURE 5.6: Detection efficiency measured for the MONOLITH prototype operated at power density $P_{\text{density}} = 2.7 \text{ W/cm}^2$ and $HV = 200 \text{ V}$, with a discrimination threshold $V_{th} = 7\sigma_V$. In the left panel, the measurement is performed in a region that extends beyond the pixel edges (black lines) by $10 \mu\text{m}$. The apparent decrease in efficiency at the external boundaries, which limits the area of the four instrumented pixels, is produced by the finite FEI4 telescope pointing resolution. The right panel shows the efficiency measured in the two triangular areas connecting the centers of pixels OA0, OA2, and OA3 and OA0, OA2, and OA1. This efficiency is unaffected by the telescope's resolution; the plot's color scale starts at 95 % efficiency.

It should be noted that the two triangles contain all the relevant pixel regions, such as the boundary between two pixels and the corner between three pixels, in the same proportion as in an entire pixel. For all working points, the efficiencies measured separately in the two triangles were found to be compatible with the statistics and were therefore averaged together. The efficiency value measured inside the two triangles for the working point of Figure 5.6 is $99.86^{+0.03}_{-0.04} \%$.

Threshold Scan

It is interesting to characterize the dependence of the detection efficiency on the discrimination threshold, as the signal-to-noise ratio of the sensor prototype influences it. The result for the DUT operated at $P_{\text{density}} = 2.7 \text{ W/cm}^2$ and $HV = 200 \text{ V}$ is shown in Figure 5.7.

Remarkably, the efficiency remains above 99 % up to threshold values of 18 times the voltage noise σ_V . A voltage threshold of $7\sigma_V$ was used throughout the data analysis, corresponding to approximately 600 electrons for this working point.

Figure 5.7 also reports the noise-hit rate measured in the laboratory, which is $4 \times 10^{-2} \text{ Hz/pixel}$ at $7\sigma_V$. Therefore, we expect fewer than one noise hit in the entire test beam data sample taken at this working point, corresponding to approximately 0.5 s of time.

The noise-hit rate, which is measured as the frequency of threshold crossing f_{NHR} , has a mathematical dependence on the applied threshold V_{th} (see, for example, [107]):

$$f_{\text{NHR}} = f_0 e^{-V_{th}^2/2\sigma_V^2}, \quad (5.3)$$

where f_0 is the frequency of zero crossing, and σ_V is the standard deviation of the voltage noise. With a calculated frequency of zero crossing of $f_0 = 1.57 \text{ GHz}$, the

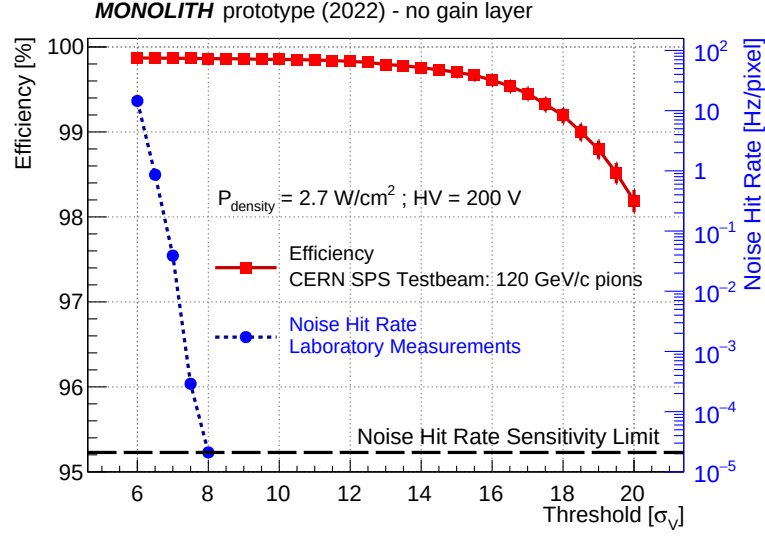


FIGURE 5.7: Detection efficiency (red squares) and noise-hit rate (blue squares) measured within the two triangular areas depicted in Figure 5.6, shown as a function of the voltage threshold in integer multiples of the voltage noise σ_V . The data refer to sensor bias $HV = 200$ V and power density $P_{\text{density}} = 2.7$ W/cm².

expected noise-hit rate frequency at $7\sigma_V$ is 3.6×10^{-2} Hz/pixel, which is compatible with the measured 4×10^{-2} Hz/pixel at $7\sigma_V$.

Bias Voltage and Power Density Scans

Table 5.2 and Figure 5.8 show the average efficiency in the two triangles for the eight available working points of bias voltage and power density at which data were acquired.

P_{density} [W/cm ²]	HV [V]	Average efficiency [%]
2.7	250	$99.80^{+0.07}_{-0.09}$
2.7	200	$99.86^{+0.03}_{-0.04}$
2.7	160	$99.71^{+0.08}_{-0.10}$
2.7	120	$99.80^{+0.07}_{-0.09}$
0.9	200	$99.83^{+0.04}_{-0.04}$
0.36	200	$99.81^{+0.04}_{-0.05}$
0.13	200	$99.77^{+0.04}_{-0.05}$
0.04	200	$99.86^{+0.05}_{-0.07}$

TABLE 5.2: Detection efficiency for the second MONOLITH prototype without a gain layer at different power density and sensor bias voltage values. A voltage threshold of $7\sigma_V$ is applied. The efficiencies are measured separately in the two triangular areas connecting the OA0, OA2, and OA3 and OA0, OA2, and OA1 pixel centers shown in Figure 5.6 and then averaged.

Within the statistical uncertainty, no significant trend is observed; all the efficiencies are compatible with 99.8 %.

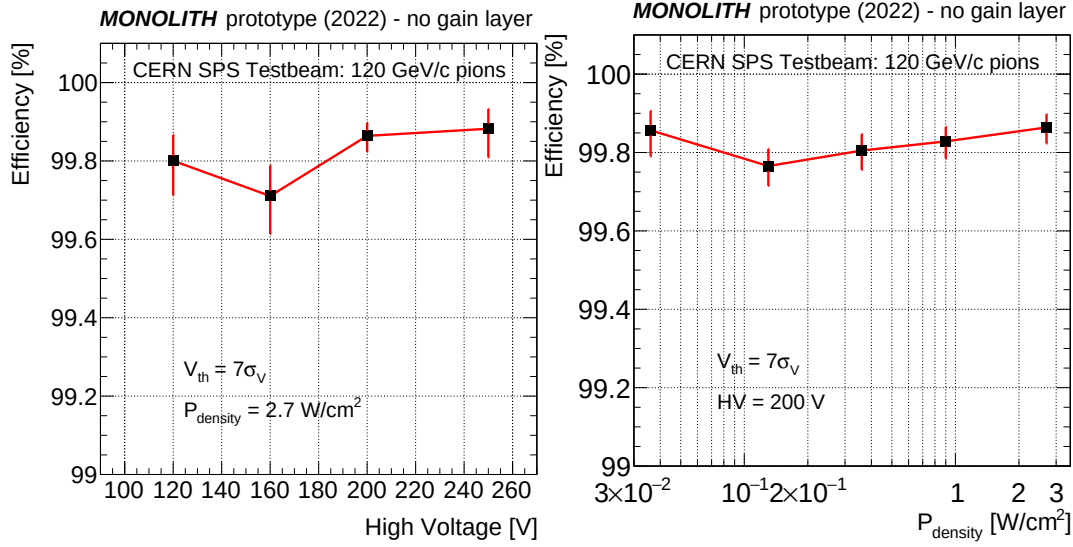


FIGURE 5.8: Detection efficiency measured within the two triangular areas depicted in Figure 5.6 with $V_{th} = 7\sigma_V$. In the left panel, the efficiency is measured at four values of sensor bias voltage for a power density of 2.7 W/cm^2 . In the right panel, the measurement is shown for five power density values at sensor bias $HV = 200 \text{ V}$.

Efficiency in Different Regions of the Pixel

Figure 5.9 shows the DUT efficiencies measured as a function of the distance from the pixel center for the eight working points acquired.

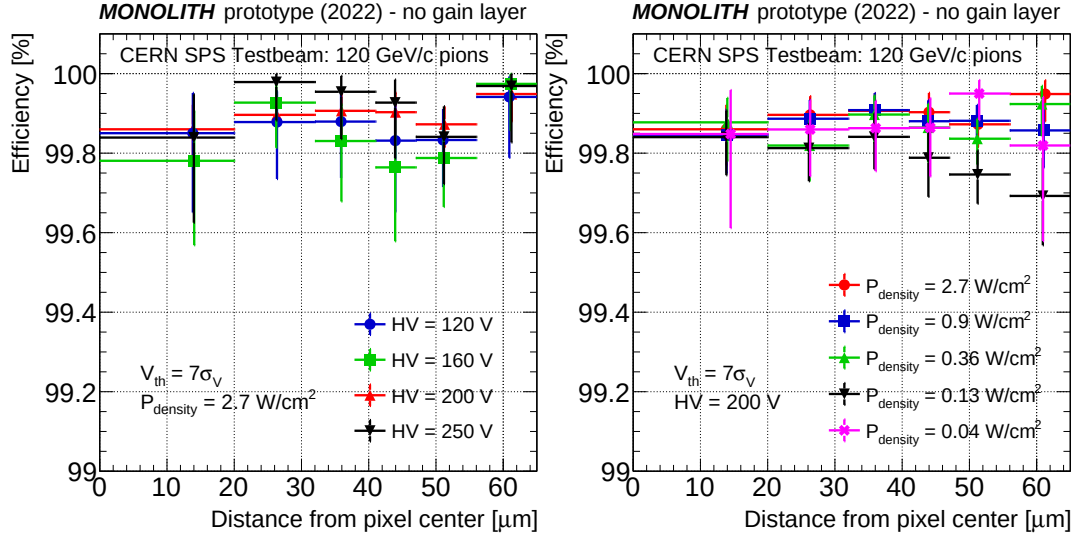


FIGURE 5.9: Detection efficiency measured within the two triangular areas depicted in Figure 5.6, in six distance intervals from the pixel center. In the left panel, the efficiency is shown for four values of HV at a power density of 2.7 W/cm^2 . The right panel shows five power density values at sensor bias $HV = 200 \text{ V}$. The data-point marker is placed at the average distance from the pixel center in each bin.

In all cases, the efficiencies remain remarkably stable at around 99.8% across the entire pixel area, indicating that no drop is measured in the inter-pixel region. This result is partly explained by the resistivity of the active layer, which was increased from $50 \Omega \text{ cm}$ of the previous prototype [58] to $350 \Omega \text{ cm}$ in the current prototype.

This modification contributes to a more uniform electric field, especially in the inter-pixel region. Moreover, the high signal-to-noise ratio of this prototype is sufficient to overcome the slight degradation in signal amplitude near the pixel edge and maintain high efficiency.

5.2.4 Time Resolution

The data samples of telescope tracks, together with the signals of the DUT and the two MCPs, were used to determine the timing performance of the MONOLITH prototype 2. The measurement was restricted to events with tracks crossing the area of pixel OA0 inside the two triangles shown in Figure 5.6 to exclude hits in the inter-pixel region that have the majority of the deposited charge in the three non-instrumented pixels adjacent to OA0, and thus should not be associated to pixel OA0.

The DUT time resolution was studied using the time of arrival ToA of the analog signals of pixel OA0, which was connected to the fastest oscilloscope. The ToA values for the DUT were computed as the time the linear interpolation between two consecutive oscilloscope samplings of the differential signal reached a $7\sigma_V$ threshold. For the MCPs, the ToA was computed as the time at which the linear interpolation between two consecutive oscilloscope samplings reached a threshold of 50 % of the total signal amplitude.

The difference in the measured ToA (called ΔToA) between the three possible detector pairs, DUT-MCP0, DUT-MCP1, and MCP0-MCP1, corresponds to the time-of-flight between the detectors. Under the assumption of a Gaussian response of the three detectors, the standard deviations extracted from the corresponding ΔToA distributions, $\sigma_{\Delta\text{ToA}}$, are interpreted as the sum in quadrature of the time resolution of the two detectors in the pair. By solving a system of three equations, similar to Equation 4.2, one for each $\sigma_{\Delta\text{ToA}}$ value, the three unknown time resolutions σ_{DUT} , σ_{MCP0} and σ_{MCP1} can therefore be retrieved.

Time-Walk Correction

The distributions of the ΔToA as a function of the amplitudes of the detectors in the pair were used to correct for time walk. As an example, Figure 5.10 shows the ΔToA distribution for the DUT and MCP0 pair as a function of the DUT (left) and MCP0 (right) signal amplitudes for the working point with a power density of 2.7 W/cm^2 and $HV = 200 \text{ V}$.

Two independent methods were deployed and cross-checked against each other to exclude potential biases to the measured timing performance introduced by the time-walk correction procedure.

- The first correction method is equivalent to the technique used for the PicoAD proof-of-concept prototype. A Gaussian fit to the ΔToA distribution was performed to determine its most probable value separately for events falling in bins of the signal amplitude. The values obtained were then linearly interpolated, resulting in a continuous parametrization of the correction as a function of the signal amplitude. Each event was then corrected twice, once for each of the detectors involved in the ΔToA .
- The second method relies on an unbinned maximum likelihood fit to extract the time-walk corrections, modeled with polynomial functions f_{DUT} , f_{MCP0} , and f_{MCP1} of the corresponding signal amplitudes, simultaneously for all detectors. A likelihood estimator $\mathcal{L}$ was constructed assuming a Gaussian time response

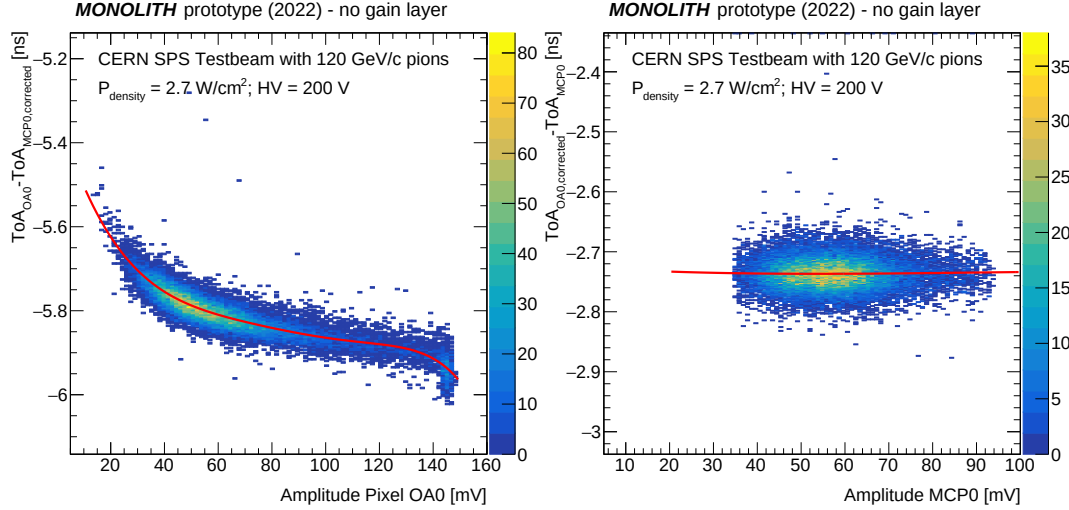


FIGURE 5.10: Distribution of the difference ΔToA between the time of arrival of DUT and MCP0 as a function of the DUT signal amplitude (left) and the MCP0 signal amplitude (right). The DUT was operated at a power density of 2.7 W/cm^2 and a sensor bias voltage of 200 V. An arbitrary offset with no influence on the measured time resolution is present in the ΔToA . The red lines show the time-walk corrections obtained with the unbinned maximum likelihood fit method described in the text. The corrections were parameterized with polynomial functions with the coefficients determined simultaneously for the three detectors.

from all sensors. The probability of observing a $\Delta\text{ToA}_{i,d}$ difference for a selected track i and the detector pair d can thus be described with a Gaussian probability density function $\mathcal{G}_{i,d}$. In the model, the standard deviation parameter $\sigma_{\Delta\text{ToA}_d}$ describes the sum in quadrature of the time resolutions of the two detectors in pair d . Additionally, being $d1$ and $d2$ the two detectors involved in the d pair associated to $\Delta\text{ToA}_{i,d}$, the total time-walk correction for track i is given by $\mu_{i,d} \equiv f_{d2}(\text{Amplitude}_{d2,i}) - f_{d1}(\text{Amplitude}_{d1,i})$. All the parameters of the model were simultaneously determined by maximizing:

$$\mathcal{L} = \log \left(\prod_i^{\text{tracks}} \prod_d^{\text{detector pairs}} \mathcal{G}_{i,d} (\Delta\text{ToA}_{i,d} | \sigma_{\Delta\text{ToA}_d}, \mu_{i,d}) \right). \quad (5.4)$$

A sixth-order polynomial function was chosen for the DUT to model the amplitude saturation effect. In contrast, third-order polynomials were used to parameterize the time-walk effect for the MCPs. This method's advantage is the absence of an arbitrary bin choice. At the same time, it is limited by its capability to precisely model the functional dependence of the time-walk correction on the signal amplitude.

These two independent methods yielded results in excellent agreement, with observed differences well below 1 ps, which confirms that systematic effects stemming from the extraction of the time-walk correction are marginal.

Methods to Measure the Time Resolution

For measuring the time resolutions, Gaussian fits of the time-walk-corrected ΔToA distributions for the three detector pairs were performed independently at each working point. To avoid possible biases of the time-resolution values coming from

the non-Gaussian tails, the range of the fit was limited to the bulk of each ΔToA distribution by including only bins populated by more than 25 % of the entries in the bin with the largest content. The σ parameters of the Gaussian fits were then used to solve the system of equations that yields the time resolutions for the DUT OA0 and the two MCPs.

As an example, the ΔToA distributions for the working point with a power density of 2.7 W/cm^2 and sensor bias voltage $HV = 200 \text{ V}$ are shown in Figure 5.11.

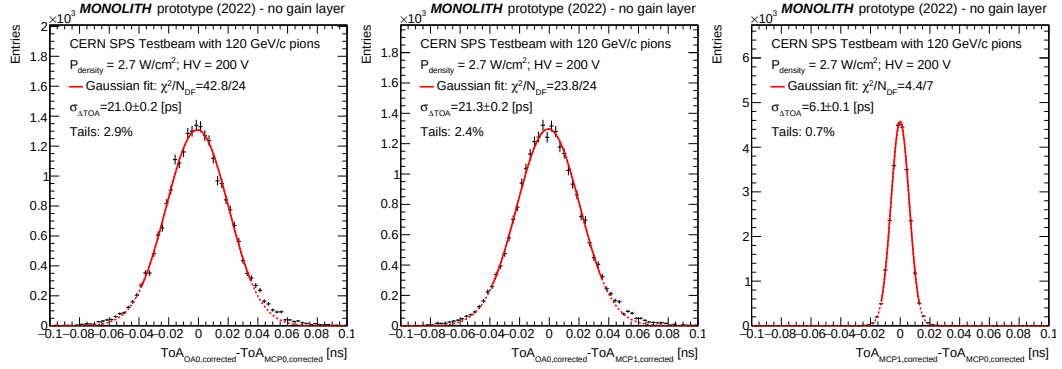


FIGURE 5.11: Distributions of the time-walk-corrected ΔToA between the DUT and MCP0 (left), DUT and MCP1 (center), and between MCP0 and MCP1 (right), for the working point with a power density of 2.7 W/cm^2 and $HV = 200 \text{ V}$. The full red lines show the results of the Gaussian fits to the bulk of each distribution, while the dashed lines extrapolate the fit to the entire histogram range.

A time resolution of $(20.7 \pm 0.3) \text{ ps}$ is measured for this working point. For the same working point, a time resolution of 48 ps was measured without correcting the time walk. The very narrow distribution of the ΔToA between the two MCPs in Figure 5.11 indicates their excellent timing performance, measured to be $(3.6 \pm 1.6) \text{ ps}$ for MCP0 and $(5.0 \pm 1.1) \text{ ps}$ for MCP1, and confirms their ability to serve as a good time reference.

The total fraction of events exceeding the Gaussian-fit integral in the ΔToA distributions involving the DUT was found to be below 3 % for all the working points, showing that the non-Gaussian component of the time response of this second MONOLITH prototype without a gain layer is overall small and that the resolutions measured with this method describe at least 97 % of the recorded signals. Section 5.5.4 presents a study on the origin of these tails, which is related to the Landau deposition of the primary charge.

Extending the time-resolution measurement to all events containing a hit in pixel OA0 (thus removing the request for the hit to be within the two triangles of Figure 5.6) does not change the time resolution values. The only effect produced is that the events with ΔToA outside the Gaussian fit increase from 3 % to 5 %; it was checked that these extra events in the tails are associated with hits in the inter-pixel region of the three sides of the OA0 pixel hexagon for which the adjacent pixels were not read out.

Bias Voltage and Power Density Scans

Table 5.3 and the left panel of Figure 5.12 report the time resolution measured at $HV = 200 \text{ V}$ for the five power density values ranging from 0.04 to 2.7 W/cm^2 . While efficiency remains at the 99.8 % level in all cases, a gradual deterioration in timing performance is visible at lower power density values. However, it remarkably remains at the level of 30 ps at a power density of 0.36 W/cm^2 and 80 ps at 0.04 W/cm^2 .

DUT operated at $HV = 200$ V and $V_{th} = 7\sigma_V$		
$P_{density}$ [W/cm ²]	Amplitude MPV [mV]	Time Resolution [ps]
2.7	48.6 ± 0.5	20.7 ± 0.3
0.9	35.8 ± 0.5	23.8 ± 0.3
0.36	22.6 ± 0.4	30.1 ± 0.4
0.13	14.2 ± 0.3	47.2 ± 0.7
0.04	16.2 ± 0.3	77.1 ± 0.9

TABLE 5.3: Time resolution at sensor bias voltage $HV = 200$ V for the five power consumption per unit surface values studied. The measurements refer to the area of pixel OA0 that is inside the two triangles of Figure 5.6. The most probable values of the amplitude of the differential signals are also reported. The uncertainties are statistical only.

The right panel of Figure 5.12 shows the time resolution as a function of the sensor bias voltage for power density $P_{density} = 2.7$ W/cm². The measurements show that the DUT can be operated in a wide HV range with a time resolution between 20 ps and 25 ps.

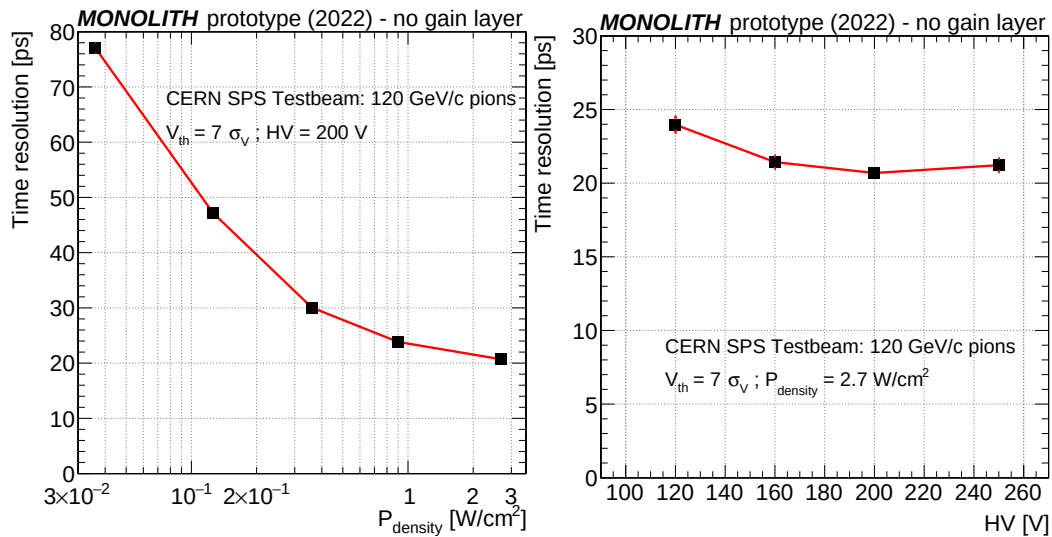


FIGURE 5.12: Time resolution measured for sensor bias voltage $HV = 200$ V as a function of the power density (left panel), and for the power density of 2.7 W/cm² as a function of sensor bias voltage (right panel).

Time Resolution in Different Regions of the Pixel

Finally, the time resolution was also measured as a function of the distance from the center of the pixel OA0. Figure 5.13 shows the results.

The left panel of Figure 5.13 shows that a mild dependence on the distance from the pixel center is measured only for $P_{density} \geq 0.36$ W/cm² when the time resolution is below ≈ 30 ps. The right panel of Figure 5.13 shows that the time resolutions measured for HV between 160 and 250 V are compatible with uncertainties. In comparison, at $HV = 120$ V, the measurements are systematically higher by approximately 2 ps. This observation may suggest that the charge drift velocity at $HV = 120$ V is not yet saturated deep within the sensor volume.

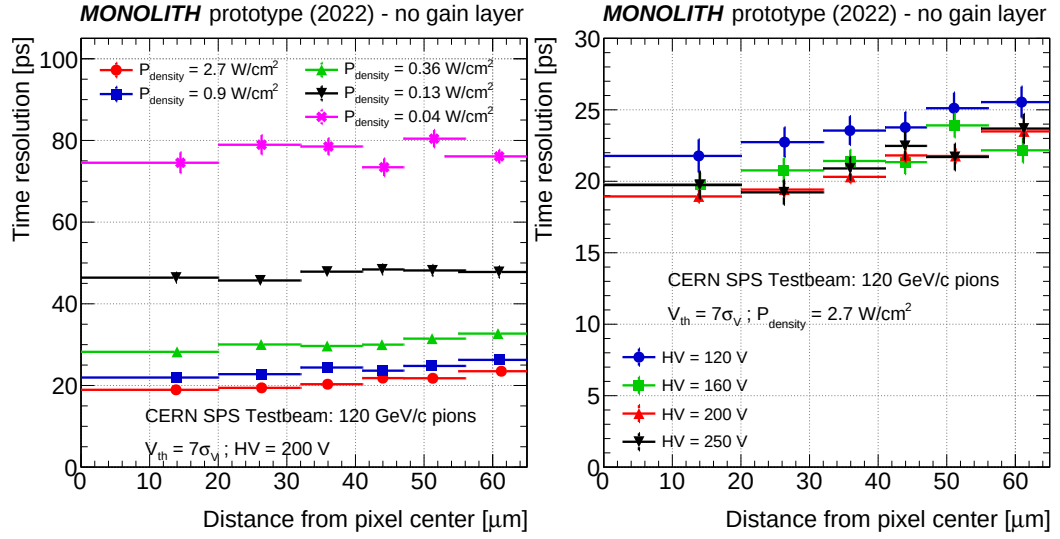


FIGURE 5.13: Time resolution as a function of the distance from the center of the pixel OA0. In the left panel, the resolution is shown for the five power density values acquired at sensor bias voltage $HV = 200$ V. The resolution in the right panel is measured at a power density of 2.7 W/cm² for the acquired four sensor bias voltage values. In each bin, the data point marker is placed at the average distance value from the center of the pixel.

Time resolution measurements are obtained by computing the time of arrival (ToA) using a simple threshold setting, which offers the significant advantage of requiring simple electronic circuitry and a straightforward signal-processing method, such as linear interpolation between consecutive oscilloscope samples. More complex signal-processing methods, such as mimicking a low-pass RC filter, a constant-fraction discriminator, or spline interpolation of the oscilloscope signal samplings, which would require more complicated electronics, marginally improve the time resolution, at most by 15 %. Therefore, they were discarded, and their results were not shown.

5.3 Femtosecond Laser

This section presents the measurements of time resolution obtained using a femtosecond laser with the second MONOLITH prototype ASIC, without a gain layer. These measurements were taken in the laboratory at the University of Geneva during the first half of 2023.

The importance of measurements using a femtosecond laser lies in the mechanism of charge creation. As mentioned in Section 3.2.1, a minimum ionizing particle creates a series of clusters with a variable number of electron-hole pairs along its path in the silicon. The charge fluctuations in each cluster are responsible for the charge-collection noise discussed in Section 3.5.1. On the contrary, an intense beam of infrared photons, as used in the femtosecond laser, creates a uniform cylinder of electron-hole pairs, which does not suffer from charge-collection noise.

Since signals produced with a femtosecond laser do not contribute to charge collection noise in the time resolution, a laser is the best method for studying the time jitter created by the analog front-end electronics. Moreover, comparing measurements from the test beam with those from a femtosecond laser allows for decoupling charge-collection noise from time resolution.

Formally, we can recall Equation 3.5, and consider some simplifications: the time walk is intrinsically corrected by taking a constant-fraction discrimination threshold and by considering very narrow amplitude regions ($\sigma_{\text{Time Walk}} = 0$), the Landau fluctuations are not present with the laser ($\sigma_{\text{Landau}} = 0$), the effect of distortions on the time resolution ($\sigma_{\text{Distortion}}$) is negligible because the laser is shoot only in the central part of the pixel, and the effect of the digitization on the time resolution (σ_{TDC}) is negligible because of the fast oscilloscope (see Section 3.5.4). Therefore, equation 3.5 becomes:

$$\sigma_t^2 \simeq \sigma_{\text{jitter}}^2 = \left(\frac{\sigma_V}{dV/dt} \right)^2, \quad (5.5)$$

where σ_V is the voltage noise of the front-end electronics, and dV/dt is the slope at the threshold of the signal.

5.3.1 Experimental Setup

The experimental setup consists of a femtosecond laser *InSight X3 Laser System*, two MONOLITH ASICs, and an oscilloscope *Lecroy WaveMaster 820Zi-B*, all connected via high-frequency cables. Of the two ASICs, one serves as a reference and the other as the device under test (DUT). Figure 5.14 shows a picture of the experimental setup (the oscilloscope is not visible).

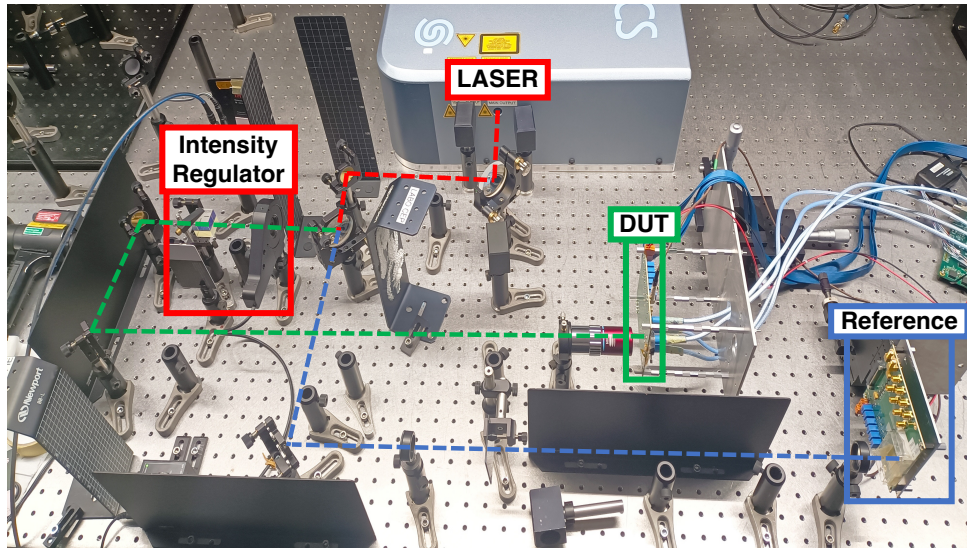


FIGURE 5.14: The experimental setup. The beam from the laser is divided into two beams by a semi-reflective lens: one goes to the DUT ASIC, and the other goes to the reference ASIC. On the DUT beamline, an intensity regulator is present.

The reference ASIC was continuously operated at a sensor bias voltage of $HV = 200$ V and a high power density of $P_{\text{density}} = 2.7 \text{ W/cm}^2$. On the contrary, the DUT ASIC was always at the same sensor bias voltage of $HV = 200$ V, but with a power density P_{density} varying between 0.36 W/cm^2 and 2.7 W/cm^2 , as summarized in Table 5.1.

The laser shoots a pulsed light beam in the infrared spectrum with a repetition rate of 80 MHz and a duration of a few hundred fs.

The reference ASIC aims to provide a very stable and precise timing reference. To achieve this, the beam arriving at the reference ASIC has a high intensity and is not well-focused, providing a very high signal and good time jitter. From the

approximation shown in the Equation 5.6, the higher the signal amplitude and the better the time resolution:

$$\sigma_t \simeq \sigma_{\text{jitter}} = \frac{\sigma_V}{dV/dt} = \frac{\sigma_V}{\text{Slew Rate}} \approx \sigma_V \frac{\text{Rise Time}}{\text{Amplitude}}. \quad (5.6)$$

The intensity of the laser beam arriving at the DUT was controlled by an intensity regulator (visible in Figure 5.14) to explore different signal amplitudes. The laser wavelength and laser focusing were fine-tuned before taking measurements. Additionally, the threshold used for the reference and the DUT ASICs was adjusted to minimize time jitter. We discuss each tuning in the following sections.

Laser Wavelength

The wavelength of the laser is tunable from 900 nm to 1100 nm and was set to maximize absorption within the active silicon layer. The second MONOLITH prototype without a gain layer can be modeled as a 300 μm thick substrate, a 50 μm thick active layer, and metal lines on top. Only the charges generated in the active layer are collected and amplified. The laser beam on the DUT was shot from the backside, crossing the substrate and then the active layer with a perpendicular incidence. As a first approximation, the intensity of the beam as a function of the distance traveled in the silicon follows the Lambert-Beer exponential function (Equation 5.7):

$$I(x) = I_0 e^{-\alpha x}, \quad (5.7)$$

where I_0 is the initial intensity, and α is the absorption coefficient of a certain wavelength crossing a medium (see Figure 3.5; a complete list of the values in silicon can be found in [103]).

Figure 5.15 shows the laser intensity as a function of the distance for different wavelengths λ , corresponding to different absorption coefficients α .

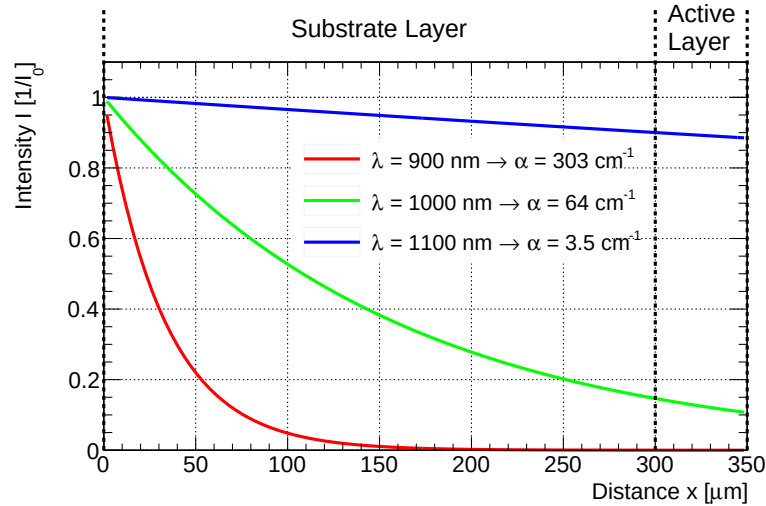


FIGURE 5.15: Intensity of the laser as a function of the distance for different wavelengths λ , corresponding to different absorption coefficients α . The laser beam passes through the 300 μm thick substrate and then the 50 μm thick active layer.

The intensity absorbed inside the active layer is maximized when the difference between the incoming and outgoing intensities in the active layer is at a maximum. Naming $x_1 = 300 \mu\text{m}$ and $x_2 = 350 \mu\text{m}$ the extreme distances of the active layer, the

formula for the intensity absorbed inside the active layer is:

$$I_{\text{active layer}} = I(x_1) - I(x_2) = I_0(e^{-\alpha(300 \text{ nm})} - e^{-\alpha(350 \text{ nm})}). \quad (5.8)$$

The maximum of absorption in the active layer can be analytically found: the maximum is when the absorption coefficient is $\alpha \simeq 30 \text{ cm}^{-1}$, which corresponds to a wavelength between 1020 nm and 1030 nm. It was measured that the wavelength that maximized the absorbed intensity in the active layer was 1042 nm.

The doping of the silicon can explain the slight discrepancy between the predicted and observed values since the expected value is for intrinsic silicon. The silicon doping shifts the absorption spectrum, as shown in [120]. The same absorption coefficient is observed for higher wavelengths with higher silicon doping, which is in the right direction to address the observed discrepancy. The wavelength of 1042 nm was used for all the data collection.

Laser Focusing

The laser beam on the DUT was centered on a single pixel and was very focused, producing most of the charge at the center of one of the four analog pixels. Figure 5.16 shows the signal from the reference ASIC and three analog pixels of the DUT.

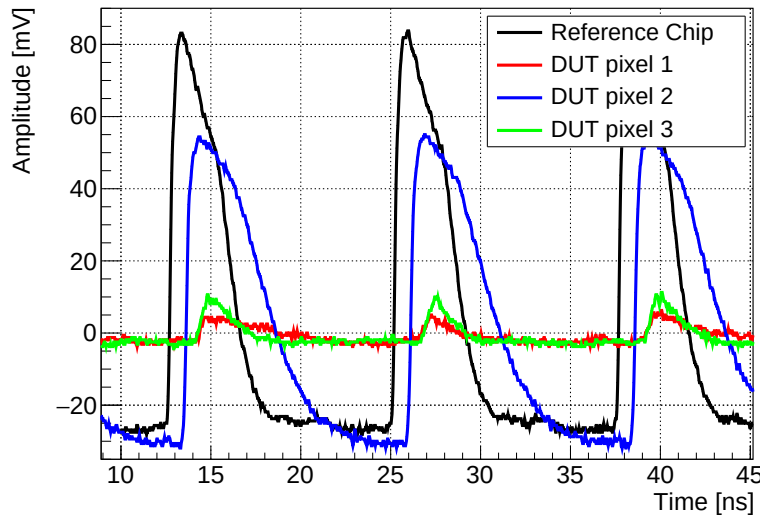


FIGURE 5.16: Signal of the laser seen by the MONOLITH ASICs (reference in black and three pixels of the DUT colored). The DUT analog output pixel OA2 (in blue) is used for timing characterization.

Figure 5.16 shows that the interested pixel (pixel OA2, in blue in the figure) has a baseline-to-peak amplitude of $\approx 85 \text{ mV}$, while the two neighboring pixels (pixel OA1 and OA3, respectively in red and green in the figure) have an amplitude of $\approx 10 \text{ mV}$. Assuming that all six neighbors of the interested pixel have the same amplitude, this means that almost 60 % of the charge is generated inside the interested pixel, and the remaining 40 % is shared between the six neighbors.

Figure 5.16 also shows that the repetition rate of the laser is 80 MHz, which corresponds to a difference in the rise of two consecutive peaks of 12.5 ns. Moreover, the DUT OA2 and the reference ASICs have a shape that is shifted by 1.2 ns, which is in agreement with the time it takes for the light to travel the 37 cm of path difference between the two beams.

Threshold Tuning

The shape of the signals from the MONOLITH prototypes was studied to determine the optimal threshold, V_{th} , for computing the arrival time in timing measurements. As mentioned in Section 3.5.3, the threshold must be set at the maximum slope to achieve the lowest possible time jitter σ_{jitter} , and, hence, the best possible time resolution. The signal has a characteristic quick rise with a subsequent slow fall, as visible in Figure 5.17.

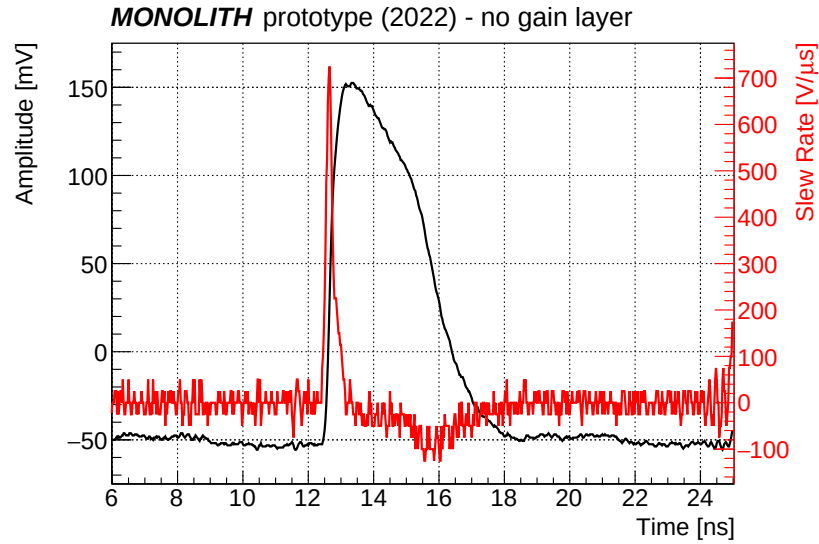


FIGURE 5.17: Example of a waveform (in black) of the second MONOLITH prototype without a gain layer used as a reference. On the same plot, the slew rate is calculated point by point using the waveform data points (in red).

Figure 5.17 also shows the slew rate as a function of time. The slew rate was calculated as the numerical derivative between two subsequent waveform points. The slew rate is not constant during the signal's rise but has a sharp peak. Hence, a point during the signal's rise maximizes the slew rate.

Figure 5.18 shows the signal's slew rate as a function of the signal amplitude at which the slew rate was calculated. A set of 1000 waveforms was considered for this plot.

Figure 5.18 helps to see better where to fix the threshold to minimize the time resolution for the reference ASIC. The average amplitude for the reference ASIC is 200 mV, which is the difference between the highest amplitude registered and the baseline, calculated as the mean of the noise in the 5 ns before the signal. On average, the maximum slew rate of $\simeq 700 \text{ V}/\mu\text{s}$ happens when the signal reaches an amplitude of $\simeq 20 \text{ mV}$, which corresponds to $\simeq 35\%$ of the maximum signal amplitude. Therefore, the threshold used to compute the time of arrival ToA is the constant fraction discriminator at 35 % of the maximum amplitude.

Figure 5.18 shows that the slew rate has no big variation in the central part of the signal. Amplitudes ranging between -20 mV and 60 mV have slew rates higher than $\simeq 600 \text{ V}/\mu\text{s}$. This interval corresponds to a constant fraction between 15 % and 55 % of the maximum signal amplitude. Despite the time resolution measurements were conducted with a constant fraction threshold of 35 %, compatible results were also found using constant fraction thresholds in the range between 15 % and 55 %, showing the solidity of this MONOLITH prototype to threshold variation.

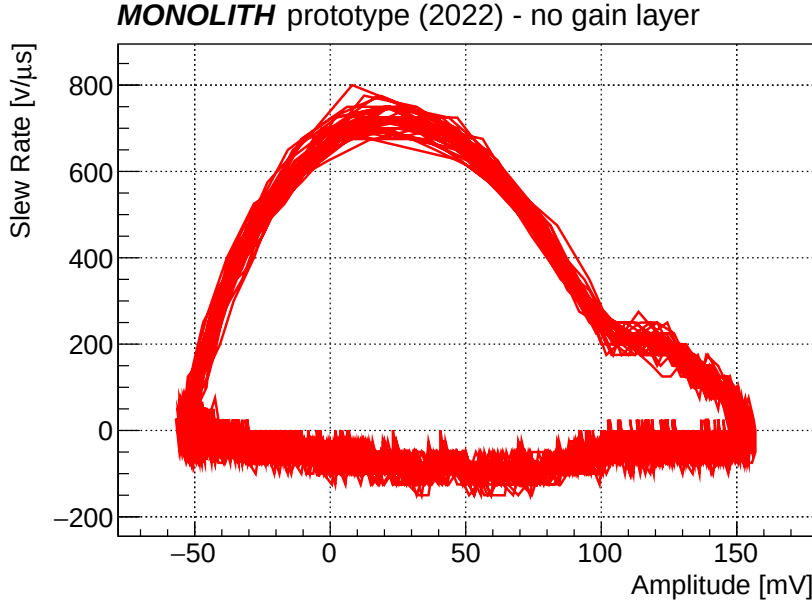


FIGURE 5.18: Slew rate calculated from the waveforms of the MONOLITH ASIC used as a reference as a function of the signal's amplitude. A set of slew rates from 1000 waveforms has been superimposed.

Finally, Figure 5.18 also shows that signals with high amplitudes (bigger than $\simeq 150$ mV) suffer from a saturation of the analog front-end electronics. The constant slew rate corresponding to amplitudes from 100 to 150 mV is not expected. This effect is due to saturation in the preamplifier or the driver stage. The saturation does not affect the time resolution, since the threshold used to calculate the arrival time is in the initial part of the signal, well before saturation occurs.

5.3.2 Time Resolution of the Reference ASIC

The time resolution of the DUT is measured in combination with the reference ASIC, as explained in the next Section 5.3.3. Therefore, the time resolution of the reference ASIC must be known with extreme precision. However, in the experimental setup used for the femtosecond laser measurements, there was no possibility of installing a third ASIC to measure the time resolution starting from a system of three objects, as done for the test beam (see, for example, Equation 4.2). Hence, an alternative method was employed to measure the time resolution of the reference ASIC.

According to the femtosecond laser specifications, the repetition rate of the laser is exact (less than 100 fs); hence, we assumed that the laser made a negligible contribution to the time resolution measured.

To measure the time resolution of the reference ASIC, we took the difference in time of arrival (ToA) between consecutive signals. Figure 5.19 shows the histogram of the difference between the ToA of two consecutive signals measured by the reference ASIC.

The standard deviation of the Gaussian fit of the distribution of ToA difference is $\sigma_{\text{Gauss}} = (3.81 \pm 0.03)$ ps. This value is the convolution of the time resolution of the reference chip taken two times; hence, the time resolution of the reference chip is $\sigma_{t,\text{ref}} = (2.69 \pm 0.04)$ ps.

In summary, the MONOLITH ASIC used as a reference has a time resolution of $\sigma_{t,\text{ref}} = (2.69 \pm 0.04)$ ps for the measured signal amplitude of (230 ± 3) mV. The

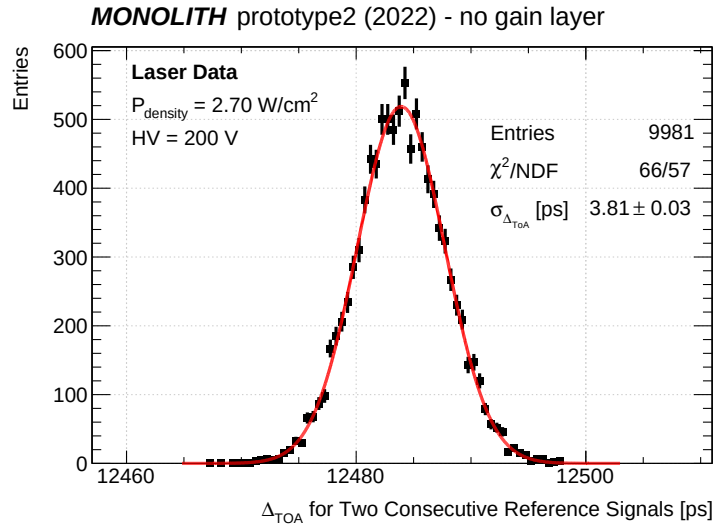


FIGURE 5.19: Difference between the ToA of two consecutive signals measured by the reference chip. The red line superposed to the data shows the result of a fit using a Gaussian functional form.

signal amplitude remained constant for all measurements, allowing for a precise and reliable time reference to extract the time resolution of the DUT ASIC. Despite this time resolution being obtained with a high charge, which is calculated to be (17.2 ± 0.2) k electrons according to the gain reported in Table 5.1, it is impressive that the second MONOLITH prototype without a gain layer can deliver such a precise time resolution.

5.3.3 Time Resolution of the DUT

The reference ASIC has a known time resolution of $\sigma_{t,\text{ref}} = (2.69 \pm 0.04)$ ps, as determined in the previous section. The laser intensity was varied on the DUT to scan different signal amplitudes. The time of arrival ToA for DUT and reference ASICs was computed as the constant fraction threshold of 35 % with a linear interpolation between consecutive oscilloscope samples. As explained in the previous section, the 35 % corresponds to a region where the slope is maximized. Then, a histogram of the ToA difference was built for each amplitude.

For each amplitude, around 10k waveforms were saved from the oscilloscope. Each waveform is a buffer of 500 ns and contains 38 good laser impulses, each with an associated signal on the DUT and the reference ASIC, for a total of about 38k ToA difference for each amplitude. It was found that the amplitude could be fixed during data taking only with a precision of a few millivolts (mV). Since we wanted an exact amplitude, we applied an offline selection, considering only signals whose amplitude was within the desired range of ± 0.15 mV, thus limiting the statistics for each amplitude. Figure 5.20 shows two examples of histograms of the difference of ToA between the DUT and the reference ASIC.

The histograms in Figure 5.20 were fitted with a Gaussian curve up to a height corresponding to 10 % of the maximum. The standard deviation σ_{Gauss} of the Gaussian fit is given by the time resolution of the DUT ($\sigma_{t,\text{DUT}}^2$) and the reference ASIC ($\sigma_{t,\text{ref}}^2$), according to: $\sigma_{\text{Gauss}} = \sqrt{\sigma_{t,\text{DUT}}^2 + \sigma_{t,\text{ref}}^2}$.

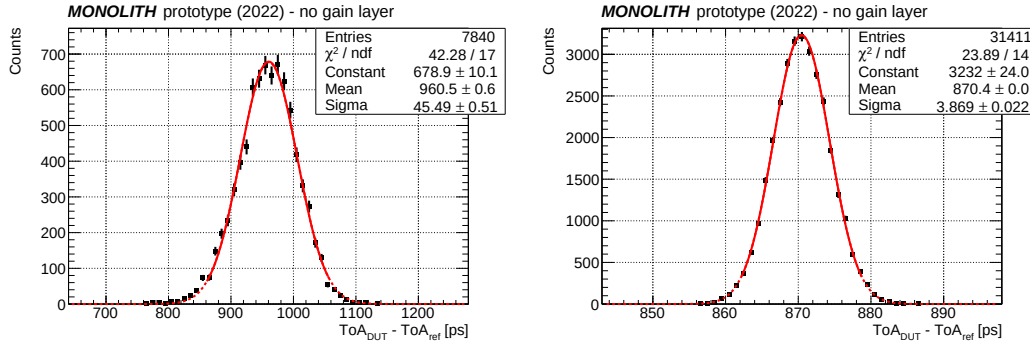


FIGURE 5.20: (Left) Difference between the ToA given by the DUT and the ToA given by the reference when the amplitude of the DUT is (20.00 ± 0.15) mV. (Right) Difference between the ToA given by the DUT and the ToA given by the reference when the amplitude of the DUT is (180.00 ± 0.15) mV. The horizontal scale is different.

Time Resolution of the DUT vs. Signal Amplitude

The time resolution of the reference ASIC is known ($\sigma_{t,\text{ref}} = (2.69 \pm 0.04)$ ps), so it is possible to calculate the time resolution of the DUT for each amplitude ($\sigma_{t,\text{DUT}}$). Figure 5.21 shows the time resolution of the DUT $\sigma_{t,\text{DUT}}$ as a function of the DUT signal amplitude. Table 5.1 shows the three power density consumptions.

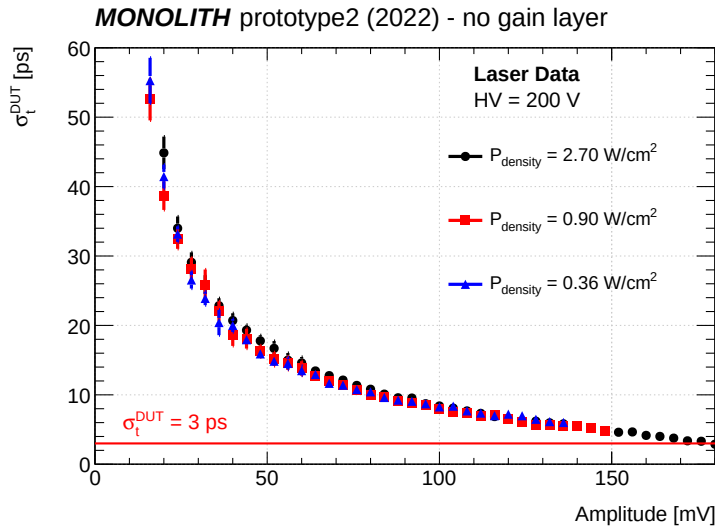


FIGURE 5.21: Time resolution of the DUT as a function of the amplitude for the three power consumption levels of Table 5.1. The red line corresponds to a time resolution of 3 ps.

Since we are dealing with precise time resolutions, it should be pointed out that, despite the ToA was taken as a constant fraction discrimination, some residuals time-walk contributions of the order of few ps remained, because the approximation of a fixed peaking time is not exactly valid for this prototype. Limiting the measured amplitudes to a specific value ± 0.15 mV also limited the time-walk contribution to less than 0.1 ps. Therefore, the electronics time jitter is the only contribution to the measured time resolution $\sigma_{t,\text{DUT}}$.

Figure 5.21 shows that, as expected from Equation 5.6, the time resolution measured with the femtosecond laser improves as the signal amplitude of the DUT increases.

At the highest power density studied $P_{\text{density}} = 2.7 \text{ W/cm}^2$, shown by the black dots in Figure 5.21, a time resolution $\sigma_{t,\text{DUT}} = 45 \text{ ps}$ is measured for signal amplitudes of 20 mV, while $\sigma_{t,\text{DUT}} = 3 \text{ ps}$ is reached at 180 mV. The gain values reported in Table 5.1 can be used to convert these amplitudes into the number of electrons; the result is approximately 1200 and 11k electrons, respectively. Data at larger signal amplitudes, which were used in the case of the reference ASIC to produce the reference timestamp, show effects of saturation of the amplifier and, therefore, were not considered in the case of the DUT.

The time resolutions as a function of the signal amplitude shown in Figure 5.21 are independent of the front-end electronics power density within uncertainties. This observation, combined with 5.6, reveals an important aspect: the product $\sigma_V * \text{Rise Time}$ is constant for a fixed amplitude, even for different power consumption.

Moreover, in the case of signals generated by a laser beam, the uniform charge distribution inside the sensor volume does not produce additional jitter at large charge values due to charge-collection noise (CCN, also known as Landau noise). Therefore, the time resolution depends only on the time jitter of the front-end electronics. In contrast, in the case of a MIP, the non-uniform charge deposition in a sensor would create a CCN that depends on the total deposited charge; the consequence would be that, for example, a given amplitude obtained with larger power density would imply a more minor electric charge and therefore smaller CCN, and provide better resolution, and the three curves would not overlap.

Time Resolution of the DUT vs. Injected Charge

The power density influences the gain of the preamplifier: higher power density leads to a higher gain. The amplitude is converted into deposited charge using the preamplifier gain from Table 5.1. Figure 5.22 shows the result of time resolution as a function of the deposited charge for the DUT.

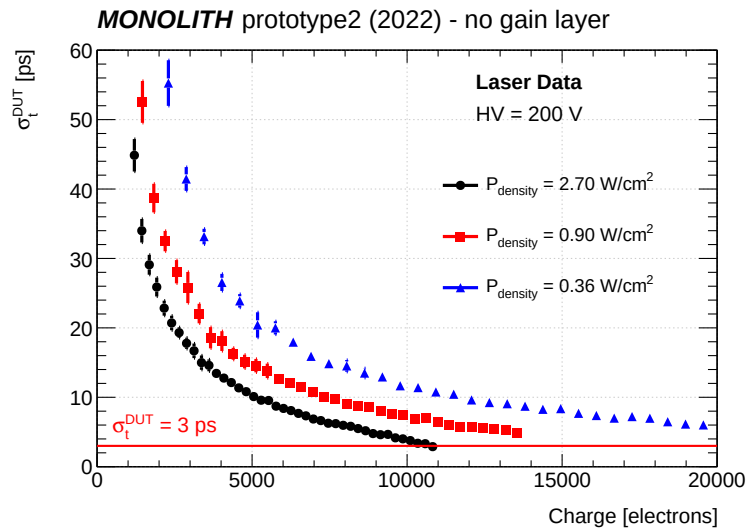


FIGURE 5.22: Time resolution of the DUT as a function of the deposited charge for the three power consumption levels of Table 5.1.

Figure 5.22 shows another aspect of Equation 5.6. Working out the substitution Amplitude = $q_{in} A_q$, where q_{in} is deposited charge, one gets:

$$\sigma_t \approx \frac{\sigma_V * \text{Rise Time}}{q_{in} A_q} \propto \frac{1}{q_{in}}. \quad (5.9)$$

The time resolution is inversely proportional to the deposited charge, and the preamplifier parametrizes the gain dependency as A_q . This result can be interpreted as follows: for a fixed injected charge, and therefore for a specific shape of the induced current, higher power consumption leads to a higher signal amplitude, resulting in intrinsically better time jitter.

5.3.4 Comparison Between Laser and Test-beam Time Resolutions

A comparison of time resolution measurements obtained with MIPs and with a laser can isolate the charge-collection noise introduced by fluctuations in the collected charge in the case of a MIP. This comparison is done in Figure 5.23, which reports the difference in arrival time between the DUT and the reference ASIC produced by the femtosecond laser, shown as a function of the DUT signal amplitude. The figure also reports the test beam data of this MONOLITH prototype as presented in the previous Section 5.2.4, taken with the same experimental conditions.

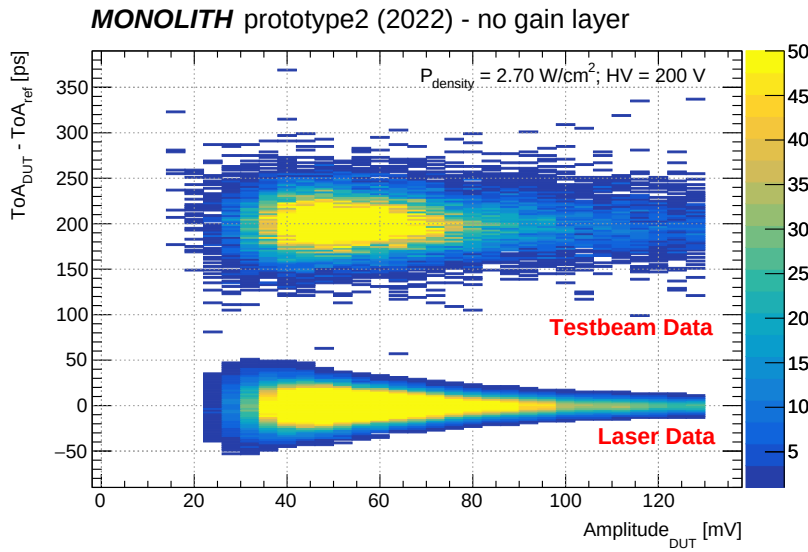


FIGURE 5.23: Difference in arrival time between the reference timestamp and the DUT as a function of the DUT signal amplitude. Both test beam (top of the plot) and laser measurements (bottom of the plot) were performed at a power density of 2.7 W/cm^2 and a sensor bias voltage of 200 V. For this comparison, the laser data were weighted according to a Landau distribution, and the Δ_{TOA} of the test beam data were shifted by 200 ps to avoid overlap with the laser data.

It's essential to notice that a different ASIC was used for the test beam and the femtosecond laser measurements, whose distributions are shown in Figure 5.23. The ASIC used during the test beam could not be used for the laser measurements because the PCB did not have a hole below the ASIC, so the laser could not be shot from the backside. However, previous laboratory measurements have shown that the two ASICs exhibit similar behavior and yield compatible results. The data shown in Figure 5.23 were taken in the same experimental conditions, namely a sensor bias voltage of 200 V, a front-end electronics power density of 2.7 W/cm^2 , and the

same settings of the oscilloscope. Moreover, the laser beam profile had a radius of approximately $30\text{ }\mu\text{m}$, measured using an array of photodiodes. Therefore, the test beam data were restricted to tracks reconstructed by the test beam-telescope crossing the sensor within $30\text{ }\mu\text{m}$ from the pixel center. In the case of the test beam data, the reference timestamp was provided by two MCP detectors, each with a time resolution of 5 ps , and the data were corrected for signal time walk.

The width of the ΔToA distribution of the laser data gets smaller for higher amplitudes, as already found in Figure 5.21. In contrast, the width of the distribution of the test beam data remains rather large at large amplitudes. This difference could be attributed to the charge-collection noise contribution to the time resolution, which is present only for the MIP data.

To quantify this difference, Figure 5.24 shows the time resolution as a function of the signal amplitude obtained with the test beam data (black squares) and with the laser (blue dots). In the case of the test beam data, the signal amplitudes were limited to 140 mV by the oscilloscope readout. In both cases, for each bin of signal amplitude, the time resolution of the DUT was obtained by subtracting in quadrature the timestamp of the reference detector from the $\sigma_{\Delta\text{TOA}}$.

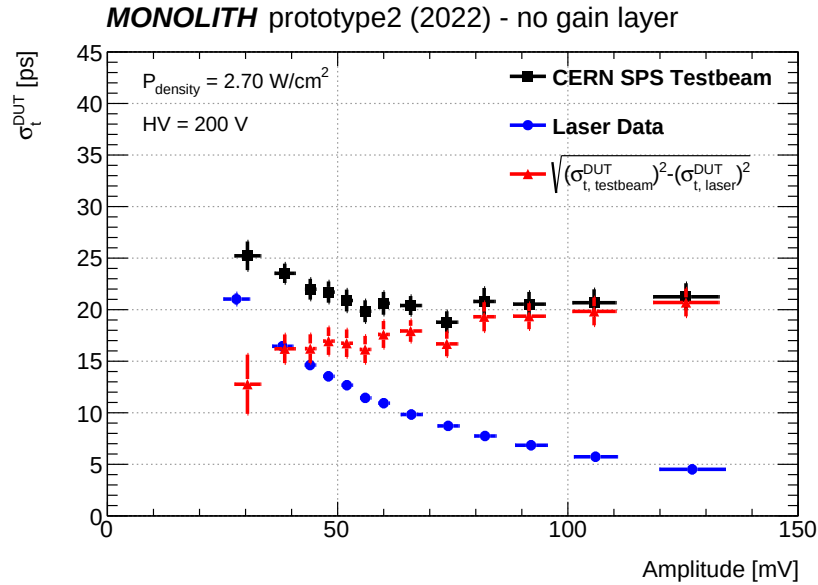


FIGURE 5.24: Comparison of the time resolutions as a function of the signal amplitude obtained with the femtosecond laser (blue dots) and with MIPs (black squares) at the CERN SPS test beam. The data were taken at a power density of $2.7\text{ W}/\text{cm}^2$ and a sensor bias voltage of 200 V . The red triangles show the difference in quadrature between the time resolutions for each bin in amplitude.

The red triangles in Figure 5.24 show the difference in quadrature between the time resolutions measured at the test beam and with the laser and provide a measurement of the time jitter produced in the test beam data by the CCN as a function of the signal amplitude. This result shows that the electronics jitter dominates the time resolution at low signal amplitudes, while the CCN dominates the time resolution at high signal amplitudes. It also shows that the CCN contribution is not constant but increases with signal amplitude.

Extraction of the Charge-Collection Noise Contribution

To gather more information on the time jitter produced by the CCN, it is instructive to compare the laser and the test beam data for the three power densities listed in Table 5.1 as a function of the electric charge instead of the amplitude. This comparison is done in Figure 5.25, which presents the laser and test beam resolutions and their difference in quadrature. The injected charge in the front-end electronics was obtained by dividing the signal amplitude by the gain for each power density reported in Table 5.1.

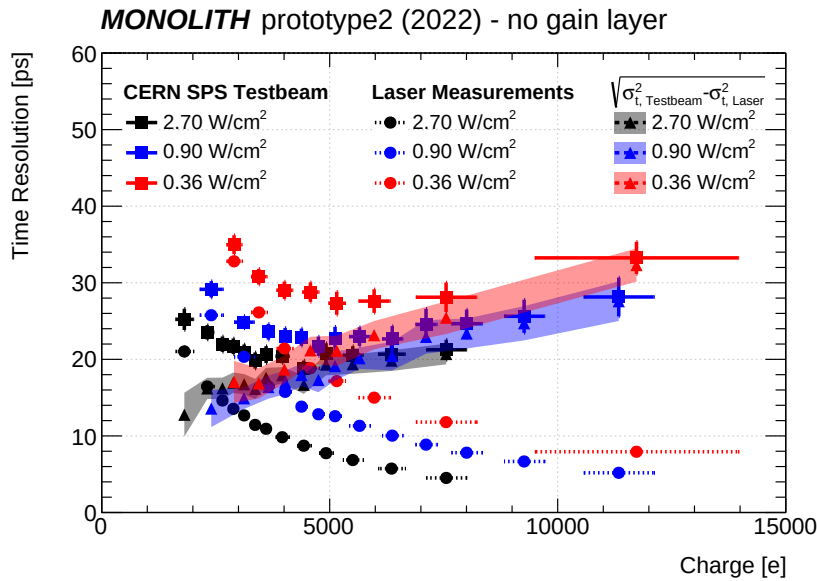


FIGURE 5.25: Time resolutions as a function of the charge injected in the front-end electronics, obtained with MIPs at the CERN SPS test beam and with the femtosecond laser. The data were taken at a sensor bias voltage of 200 V and the three power densities listed in Table 5.1. The bands show the difference in quadrature between the time resolutions for each charge bin, the width corresponding to one standard deviation.

Only signals with amplitudes less than or equal to 140 mV were considered for this study since test beam signals with amplitudes beyond that value were saturated by the oscilloscope setting and, therefore, could not be time-walk corrected with sufficient accuracy. For this reason, the data displayed as a function of the injected charge has different ranges, and the data taken at a higher power density of the front-end electronics are limited to smaller charge values.

Both laser and test beam data in figure 5.25 show better time resolutions when the power density of the front-end electronics is higher. On the other hand, the two data sets behave very differently as a function of the charge. While the time resolutions of the laser data improve significantly at large charge values, the test beam data show time resolutions that worsen at large charge values. The presence of time jitter related to the CCN, which is calculated as the difference in quadrature and is represented by the colored bands, can explain the worsening at high charges. For the three power density values, the CCN time jitter shows an apparent increase with the charge that produces the signals. It dominates the time resolution at the most significant charge values measured. The red band showing the CCN contribution at 0.36 W/cm² seems to rise more quickly than the other two bands. However, this observation is marginal since the band limits represent one standard deviation of the measurement uncertainty.

5.4 Radiation Tolerance

As demonstrated in the previous sections, SiGe HBTs can be used to produce low-noise, low-power, and high-speed front-end electronics that can be integrated into a monolithic ASIC, achieving a fully efficient and thin detector with excellent time resolution. These properties reflect the fundamental and crucial requests for future collider experiments. Still, another fundamental request must be satisfied: the silicon detectors for future collider experiments must be radiation tolerant, in principle up to fluences of $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$, because of the high luminosity involved in these colliders, which will lead to a high number of particle traversing the detectors.

Previous studies on SiGe HBT radiation tolerance can be found in [121, 122]; however, no exhaustive studies have been conducted on SiGe HBT electronics integrated in a monolithic ASIC. Hence, to study the radiation tolerance, eight samples of the second MONOLITH prototype without a gain layer were irradiated with 70 MeV protons up to $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$.

In this section, we present the results of laboratory measurements performed using radioactive sources to evaluate the timing performance of the SiGe HBT amplifier implemented in the second MONOLITH prototype, which does not include a gain layer. The measurements focus on characterizing the single-ended output of the four analog pixels, which consists of a fast charge amplifier made from a SiGe HBT and a two-stage analog driver. This setup allows for the direct measurement of the analog pulse using a fast oscilloscope. Then, we present the test-beam measurements performed on the same ASICs to more accurately assess the efficiency and timing performance using 120 GeV/c pions delivered by the CERN SPS.

5.4.1 Proton Irradiation of the SiGe HBT ASICs

Eight prototype ASICs of the second MONOLITH prototype, without a gain layer, were wire-bonded onto readout boards and characterized using a ^{55}Fe radioactive source in the clean rooms of the University of Geneva in terms of gain and Equivalent Noise Charge (ENC) during October and November 2022.

The eight boards were then shipped to the Cyclotron and Radioisotope Center (CYRIC) proton irradiation facility at Tohoku University in Japan, which has an azimuthal-varying field cyclotron and a beamline for the radiation damage test of semiconductors [123]. The beamline supplies a high-intensity proton beam with a kinetic energy of 70 MeV and a beam current of $1.5 \mu\text{A}$. The eight ASICs were irradiated with proton fluences varying from $2 \times 10^{13} \text{ n}_{\text{eq}}/\text{cm}^2$ to $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$. The electronics on the ASICs were powered during irradiation. A 15 mm-thick aluminum mask was positioned in front of the boards to shield the active components surrounding the ASICs. By the end of January 2023, we had the ASICs on the boards back, and we started testing them at the clean room laboratory at the University of Geneva during the first half of 2023.

Figure 5.26 shows the readout board hosting one of the two ASICs irradiated at the maximal fluence of $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$.

Since not all the protons were stopped by the shielding, the radiation damaged some of the Low-Dropout Regulators (LDOs) used to generate low-voltage biases to the ASIC. A clear correlation was found between the number of damaged LDOs, their position on the board, and the radiation fluence. Copper wires were soldered to the ASIC to directly supply the correct low voltage via a power supply, thereby bypassing damaged LDOs.

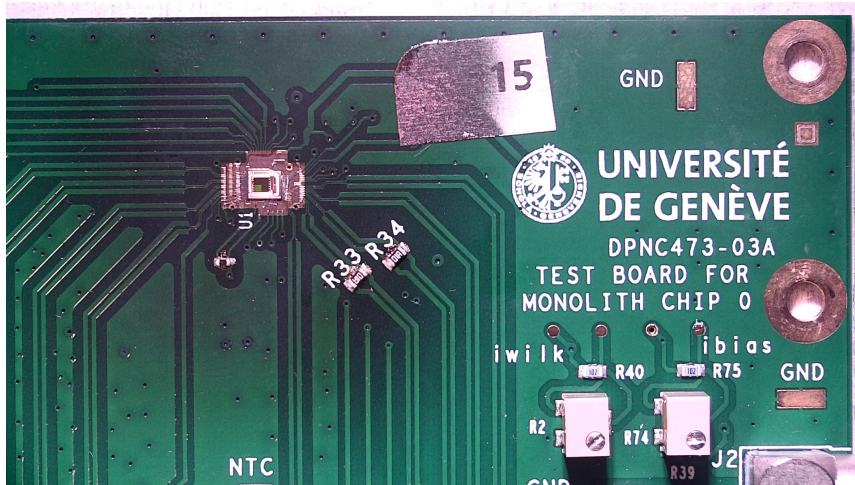


FIGURE 5.26: Photograph of a board irradiated with a proton fluence of $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$. The area around the ASIC was damaged by radiation, and some of the electronic components on the board were severely damaged.

After irradiation, the ASICs were retested in the clean rooms at the University of Geneva. One of the two ASICs irradiated at $3 \times 10^{15} \text{ n}_{\text{eq}}/\text{cm}^2$ developed a short circuit on the digital power supply, and it became unusable after irradiation. This damage might be related to radiation on the ASIC itself or other causes associated with the transport, storage, or manipulation of the ASIC. One likely cause could be the failure of the LDOs on the board during irradiation, which set the voltage on the ASIC to 3.0 V, well above the 1.2 V specified for the LVMOS in SG13G2. This damaged ASIC was excluded from the following analysis.

As a result, seven irradiated ASICs were used in this study. In addition, three non-irradiated ASICs were characterized (one of them is the same studied in the test-beam campaign in Section 5.2). They were used to have a reference for the behavior of the ASIC in the absence of radiation. Table 5.4 lists the five proton fluences considered, the number of boards, and the number of characterized analog pixels for each irradiation point, considering four analog outputs per ASIC.

Fluence [1 MeV $\text{n}_{\text{eq}}/\text{cm}^2$]	Boards prepared	Boards utilised	Analog pixels characterised
0	3	3	12
2×10^{13}	1	1	4
9×10^{13}	1	1	4
6×10^{14}	2	2	8
3×10^{15}	2	1	4
1×10^{16}	2	2	8

TABLE 5.4: Summary of the boards prepared for this study and the proton fluence at which they were exposed. The last two columns report the total number of boards and analog pixels that were tested.

During transportation, storage in the clean rooms, and measurement with radioactive sources, the temperature of the ASICs was maintained at -35°C to prevent any unwanted annealing.

Radiation Damage to the Current Mirrors

It was observed that radiation damage limited the range of the preamplifier current that could be set. The problem is related to damage to the current mirrors responsible for powering the MOSFET that sets the preamplifier current to the amplification stage (see Figure 3.16 for an explanation of the amplification stage). The change in characteristics of the bias generation circuit can be easily verified by checking that the measured current absorption of the analog power supply (*true preamplifier current*) is equal to the current set by software to the preamplifier (*set preamplifier current*), as done in Figure 5.27.

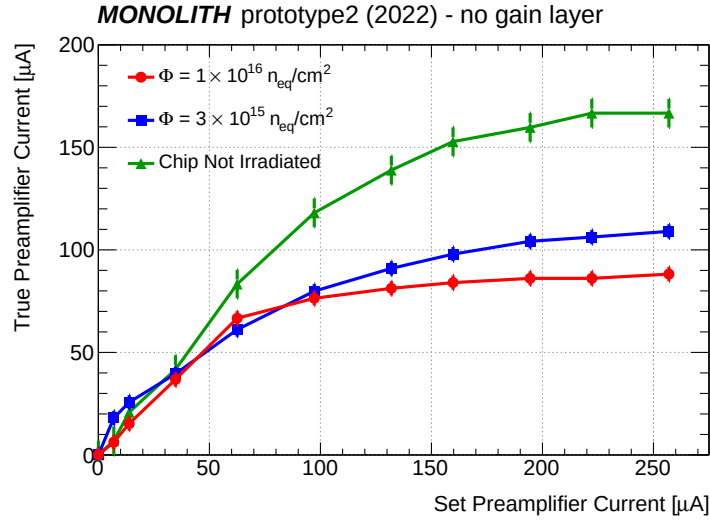


FIGURE 5.27: True preamplifier current vs. set preamplifier current for three proton fluences. The saturation level of the measured preamplifier current shows that the maximum current available to polarize the transistor diminishes with increasing fluence. A set current of 50 μA , which corresponds to $P_{\text{density}} = 0.9 \text{ W}/\text{cm}^2$, was chosen for all the measurements of this study.

As shown in Figure 5.27, it was found that the curve of measured current vs. set current saturates rapidly at high proton fluence. Consequently, it was decided to operate all ASICs at $I_{\text{preamp}} = 50 \mu\text{A}$ to compare them at the same power consumption, despite being subject to different proton fluences. This current corresponds to a power density $P_{\text{density}} = 0.9 \text{ W}/\text{cm}^2$ at which all ASICs could be operated, in contrast to the maximum value $P_{\text{density}} = 2.7 \text{ W}/\text{cm}^2$ that can be achieved before irradiation.

Since radiation damage causes an increase in the base leakage current of the HBT, the working point of the front-end electronics needs to be adjusted. The best performance is achieved by increasing the feedback current I_{fbk} supplied to the HBT base up to the point in which the bias generation circuit is not able to provide the extra feedback current anymore, as observed for I_{preamp} . It's essential to note that the limit to the maximum deliverable current from the biasing circuit is related to the circuit design, rather than a technological limitation.

To choose a working point suitable for all the ASICs, scans were conducted as a function of I_{fbk} before data taking with the radioactive source. The value $I_{\text{fbk}} = 2.0 \mu\text{A}$ was chosen, which allowed satisfactory operation of all the irradiated ASICs. This feedback current value was also used to operate the three not irradiated ASICs, in contrast to the value $I_{\text{fbk}} = 0.1 \mu\text{A}$ used for one of the not irradiated ASICs in the previous characterizations (in Section 5.2 and Section 5.3).

5.4.2 Characterization with a ^{90}Sr Radioactive Source

^{90}Sr is an almost pure β^- emitter, with two main decays with peak energies of 0.55 MeV and 2.28 MeV. Most of this source's electrons behave similarly to minimum-ionizing particles in a silicon detector, such as the second MONOLITH prototype without a gain layer.

Waveforms and Signal Amplitudes

The three top panels in Figure 5.28 show the average of 10 signals acquired by a 1 GHz bandwidth oscilloscope from one of the positive single-ended analog outputs in the case of an ASIC not irradiated and one of the ASICs irradiated with fluences of 3×10^{15} and $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$. The 10 signals were selected with amplitude within $\pm 5 \text{ mV}$ from the mode of the Landau distributions.

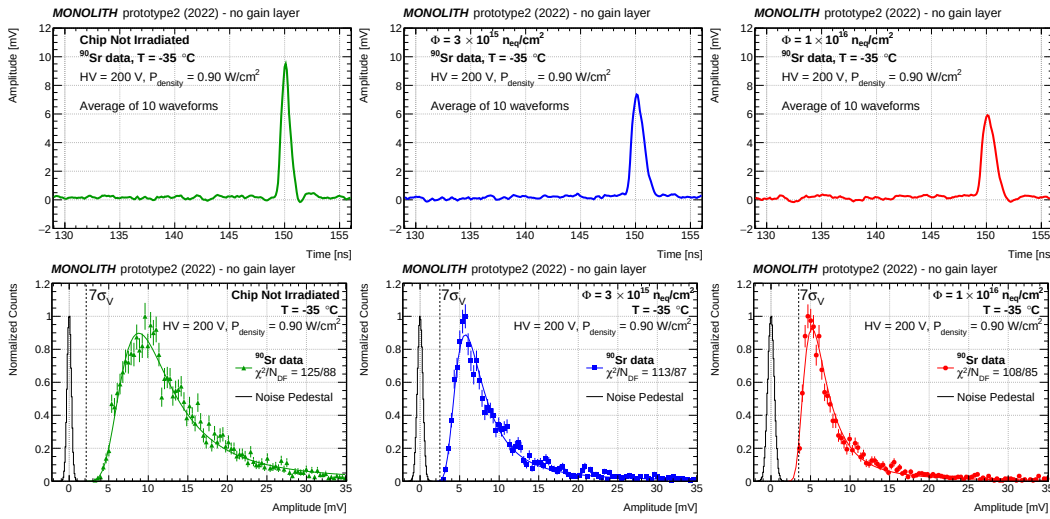


FIGURE 5.28: Waveforms (top) and amplitude distributions (bottom) acquired using a ^{90}Sr radioactive source at a preamplifier power consumption of $0.9 \text{ W}/\text{cm}^2$, for a not irradiated (left), irradiated at $3 \times 10^{15} \text{ n}_{\text{eq}}/\text{cm}^2$ (center) and $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$ (right) ASIC. The waveforms are the average of 10 signals, selected to have an amplitude within $\pm 5 \text{ mV}$ from the mode of the Landau distributions. The colored lines superimposed to the amplitude distributions are the results of fits using a Landau functional form, which were used to obtain the values of the mode of the distributions; the bottom panels also display the noise pedestals (black histograms) as well as the amplitude that corresponds to the $7\sigma_V$ threshold cut (dashed vertical lines).

The most probable signal amplitude from the ASIC not irradiated is around 9 mV , and it decreases to approximately 6 mV at $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$.

The signal amplitude distributions for the same three pixels are shown in the bottom panels of Figure 5.28, where the results of the fits using a Landau functional form are superimposed. The figures also display the noise pedestals used to determine the voltage noise σ_V , used to set the $7\sigma_V$ threshold for the analysis. The dashed vertical lines display the position of the threshold in each case, indicating that only in the case of the ASIC irradiated to $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$ the voltage threshold cut into the amplitude distribution, creating an inefficiency at the level of a few percent.

Main Characteristics vs. Proton Fluence

Figure 5.29 presents the main characteristics of the signals produced by the electrons emitted by the ^{90}Sr radioactive source, namely the signal amplitude (which depends on the amplifier gain), the voltage noise σ_V , the slope dV/dt at the $7\sigma_V$ threshold and the signal-to-noise ratio. All quantities were measured at a temperature of -35°C and $P_{\text{density}} = 0.9\text{ W/cm}^2$ for each available analog pixel and then averaged between the pixels that were subject to the same fluence (see Table 5.4). The data displayed as black squares were collected at $HV = 200\text{ V}$ and $V_{\text{CCA}} = 1.8\text{ V}$. The green lines represent the average values measured using the 12 pixels of three ASICs that were not irradiated.

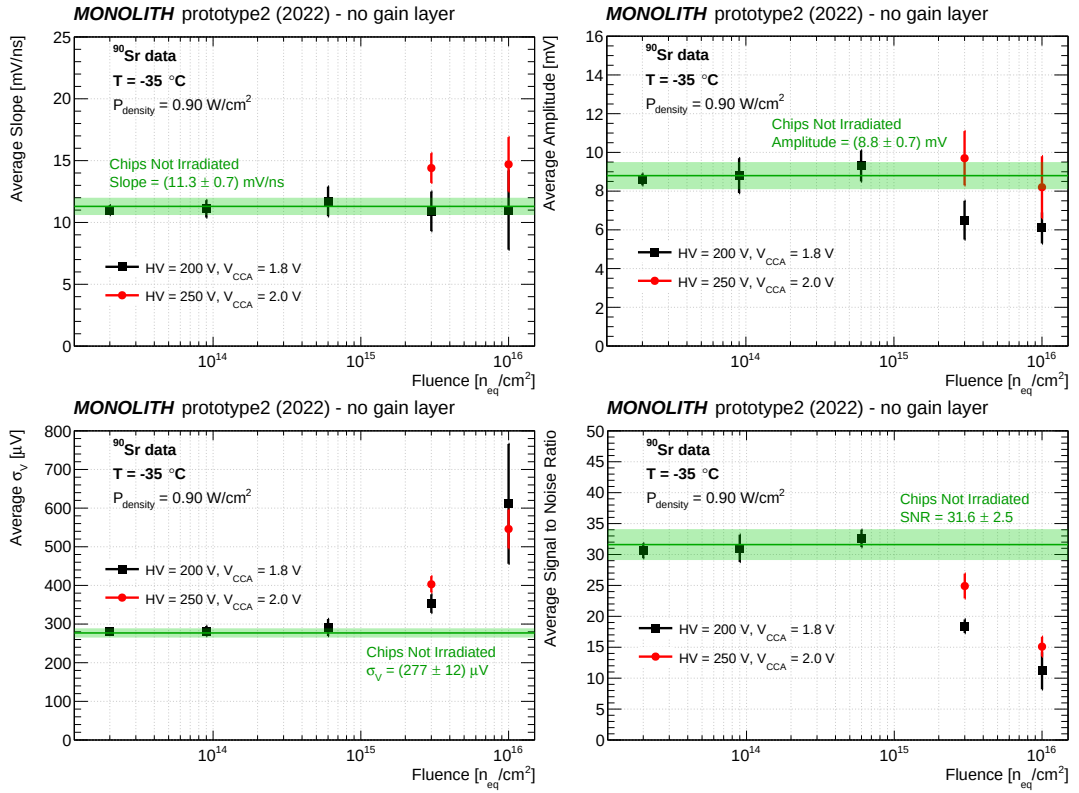


FIGURE 5.29: Average signal slope at the $7\sigma_V$ threshold value (top-left), average mode of the signal amplitude (top-right), average σ_V (bottom-left), and average signal-to-noise ratio (bottom-right), all measured with a ^{90}Sr radioactive source as a function of the fluence. Data were collected at -35°C and $P_{\text{density}} = 0.9\text{ W/cm}^2$, and with two different working points: $HV = 200\text{ V}$ and $V_{\text{CCA}} = 1.8\text{ V}$ (black squares), or $HV = 250\text{ V}$ and $V_{\text{CCA}} = 2.0\text{ V}$ (red dots). The error bars displayed represent the standard deviation of the average values. The green lines represent the average values measured with the 12 analog pixels of the three not irradiated ASICs, operated under the same conditions as the black squares. The green bands represent the standard deviation of these average values.

Figure 5.29 shows that the signals are unaffected by radiation up to fluences of $6 \times 10^{14}\text{ n}_{\text{eq}}/\text{cm}^2$. For larger fluence values, the average amplitude degrades, and the average voltage noise σ_V increases, thereby reducing the average signal-to-noise ratio by up to a factor of three. The signal slope appears to be less affected by radiation than the other signal characteristics.

For one of the two ASICs irradiated at $1 \times 10^{16}\text{ n}_{\text{eq}}/\text{cm}^2$, with a V_{CCA} value supplied of 1.8 V , the signal-to-noise ratio of two of the four analog pixels was not

large enough for a smooth operation. This variation can probably be attributed to process mismatch. These two channels were recovered by increasing the V_{CCA} of the ASIC to 2.0 V and the sensor bias to 250 V, which increased the signal amplitude and led to a large enough signal-to-noise ratio to operate them well. For the two highest fluence values, data were also taken at $HV = 250$ V and $V_{CCA} = 2.0$ V (red dots). The improved performance at high proton fluence of this second working point shows that V_{CCA} and the sensor bias HV represent two other parameters, in addition to I_{fbk} , that allow smooth and effective operation of SiGe HBT front-end electronics even at a proton fluence of 1×10^{16} n_{eq}/cm^2 .

5.4.3 Estimation of the Time Jitter of the Signal

Characterizing a timing detector using a radioactive source has some differences and challenges compared to a test-beam measurement. First, there is no control over the track selection that an external beam telescope would provide, which results in tracks at all angles being included in the analyzed sample. Second, fast detectors like MPCs do not give an external time reference.

Although measurements with a radioactive source cannot replace test-beam measurements, they offer flexibility. They can provide valuable information early in the qualifying process for a timing detector. For these reasons, data taken with a ^{90}Sr source were used to assess the radiation tolerance of the second MONOLITH prototype without a gain layer.

The time jitter of the signal was evaluated using the method described below, which utilizes the time jitter computed at the most probable value (mode) of the signal amplitude distribution. For each signal, the time jitter measured with the ^{90}Sr radioactive source was obtained with the usual formula for the time jitter, e.g., Equation 3.19. The signal slope dV/dt was calculated by linear interpolation of the oscilloscope signal samplings between thresholds of $6\sigma_V$ and $7\sigma_V$, at which the signal time of arrival from the test beam was taken. The time jitter at the most-probable signal amplitude, called $\sigma_t^{90\text{Sr}}$, was computed for each pixel in the following way.

1. The values of the time jitter of all signals in the pixel were computed using formula 3.19, by subtracting in quadrature the oscilloscope contribution to σ_V , which was typically 150 μV . Then, all the time jitter values were plotted in bins of the signal amplitude, and the average values of the time jitter and the amplitude for each bin were retained.
2. In parallel, the signal amplitude distribution was fitted with a Landau functional form, as shown in the bottom panels of Figure 5.28, to obtain the mode of the Landau distribution.
3. Finally, the value of the time jitter for the most-probable value $\sigma_t^{90\text{Sr}}$ for the pixel was obtained by linear interpolation of the data points of the corresponding time jitter vs. amplitude plot, computed at the mode value of its amplitude distribution.

Figure 5.30 shows the average time jitter values as a function of the average signal amplitudes in bins of amplitude, for a pixel of an ASIC not irradiated, one irradiated at 3×10^{15} n_{eq}/cm^2 and one at 1×10^{16} n_{eq}/cm^2 .

The time jitter vs. signal amplitude distributions of the three ASICs in Figure 5.30 show the same trend. Still, the time jitter worsens systematically with increasing proton fluence, indicating a degradation of time resolution with exposure to higher

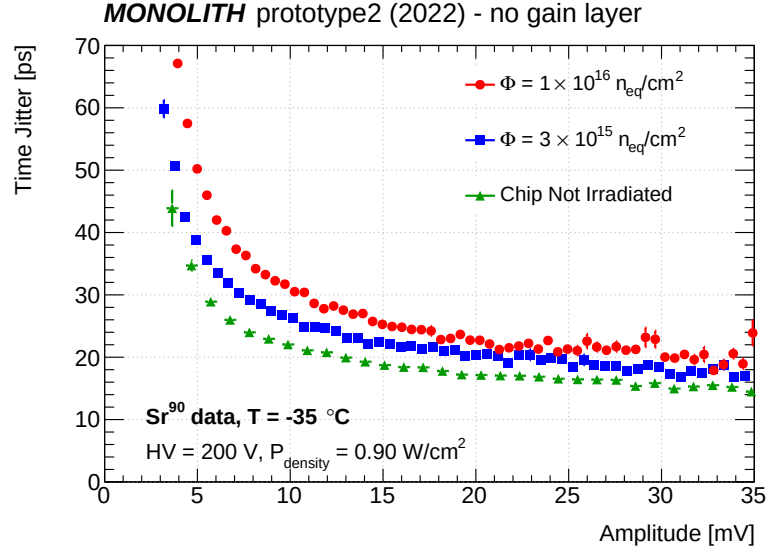


FIGURE 5.30: Average value of the time jitter calculated according to equation 3.19 in each bin of the signal amplitude, measured with a ^{90}Sr radioactive source for three different fluences. The data refers to one of the four analog pixels in each ASIC. For the sake of clarity, the results obtained with the four ASICs irradiated up to $6 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$ are not shown in the figure since they overlap with the distributions obtained for the ASICs not irradiated.

fluence. As said before, the slope at a threshold of $7\sigma_V$ of the signal is less affected by the proton fluence. Therefore, the degradation of the time jitter must be attributed mainly to the increase in the voltage noise σ_V , the other ingredient to calculate the time jitter according to Equation 3.19.

The accuracy of this method for non-irradiated ASICs was established by re-analyzing the test beam data of the previous Section 5.2.4. It was found that it provides a signal time jitter of $(22.7 \pm 0.3) \text{ ps}$ to be compared with the time resolution of $(23.8 \pm 0.3) \text{ ps}$ (see, for example, Table 5.3 when $P_{\text{density}} = 0.9 \text{ W}/\text{cm}^2$). Therefore, we can conclude that the results obtained with the method utilized here to estimate the time resolution using the electrons from a ^{90}Sr source are accurate at the 10 % level for the non-irradiated boards. This method leads to such a slight deviation between time jitter and time resolution because, as confirmed from Section 5.5, the contribution of the charge-collection noise to the time resolution in this MONOLITH prototype is low compared to the electronics time jitter.

Time Jitter vs. Fluence

Plots analogous to those in Figure 5.30 were produced for each of the 40 analog pixels of the ten ASICs used for this study and were utilized to estimate the time jitter with the method described above.

Table 5.5 reports the resulting time jitter of the 40 pixels, measured at the same working point. The last column shows the average time jitter obtained for all pixels exposed to the same proton fluence. The error quoted represents the standard deviation of the pixels used for the average.

The time jitter of the three non-irradiated ASICs is $\sigma_t^{90\text{Sr}} = (21.0 \pm 0.4) \text{ ps}$. This result is compatible with errors with the time jitter obtained with the ASICs irradiated up to $6 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$. For larger proton fluence, the time jitter at the same sensor

Fluence [n _{eq} /cm ²]	$\sigma_t^{90\text{Sr}}$ [ps]				Average $\sigma_t^{90\text{Sr}}$ [ps]
	OA0	OA1	OA2	OA3	
$HV = 200\text{ V}, V_{\text{CCA}} = 1.8\text{ V}$					
0	22.1	20.5	18.8	19.9	21.0 ± 1.4
	22.1	22.7	19.5	19.6	
	22.9	22.2	21.1	20.7	
2×10^{13}	21.4	22.2	21.2	22.4	21.8 ± 0.6
9×10^{13}	21.4	22.5	21.0	21.8	21.7 ± 0.6
6×10^{14}	21.5	22.4	20.2	20.9	21.5 ± 0.8
	20.7	22.3	22.6	21.3	
3×10^{15}	32.7	33.2	31.4	32.8	32.5 ± 0.8
1×10^{16}	43.3	50.9	44.0	47.5	56.6 ± 16.3
	84.9	79.6	48.9	53.7	
$HV = 250\text{ V}, V_{\text{CCA}} = 2.0\text{ V}$					
3×10^{15}	28.7	29.0	28.5	29.5	28.9 ± 0.4
1×10^{16}	33.2	36.2	35.5	33.6	39.6 ± 6.6
	51.4	46.9	38.4	41.7	

TABLE 5.5: Time jitter of the 40 analog pixels used for this study. A row reports the proton fluence at which the ASIC was exposed and the time jitter σ_t^{90Sr} measured for the four analog pixels in the ASIC. The last column shows the time jitter obtained by averaging all the pixels exposed to the same proton fluence; the error is the standard deviation of the values. The data were collected using the single-ended signals produced by the electrons emitted by a ^{90}Sr source. The last three rows report the time-jitter values obtained at $HV = 250 \text{ V}$ and $V_{CCA} = 2.0 \text{ V}$, which provide a better and more uniform time jitter for the various pixels.

bias voltage of $HV = 200 \text{ V}$ and $V_{CCA} = 1.8 \text{ V}$ becomes $\sigma_t^{90Sr} = (32.5 \pm 0.8) \text{ ps}$ at $3 \times 10^{15} \text{ n}_{eq}/cm^2$ and $\sigma_t^{90Sr} = (56.6 \pm 16.3) \text{ ps}$ at $1 \times 10^{16} \text{ n}_{eq}/cm^2$.

Table 5.5 shows that two pixels of one of the ASICs irradiated to $1 \times 10^{16} \text{ n}_{eq}/cm^2$ have time jitter of approximately 80 ps and are responsible for the significant standard deviation at this fluence. As anticipated in the previous section, to recover the performance of these analog outputs, the V_{CCA} was increased to 2.0 V, and the sensor bias to $HV = 250 \text{ V}$. At this modified working point, the average time jitters, reported in the bottom part of Table 5.5, become $\sigma_t^{90Sr} = (28.9 \pm 0.4) \text{ ps}$ at $3 \times 10^{15} \text{ n}_{eq}/cm^2$ and $\sigma_t^{90Sr} = (39.6 \pm 6.6) \text{ ps}$ at $1 \times 10^{16} \text{ n}_{eq}/cm^2$.

The average time jitters are also shown in Figure 5.31 by black squares or red dots and for the not irradiated ASICs by the horizontal green bar. Since the signal slope does not change significantly between different levels of fluence (as shown by the top-left panel of Figure 5.29), the worsening of the time jitter can be mainly attributed to the increase of the voltage noise σ_V with increasing fluence reported in Figure 5.29 bottom-left.

Average Amplitude vs. Sensor Bias

A high level of radiation crossing a silicon sensor damages the substrate, generating a variation of the resistivity and an increase in the concentration of charge traps ([124], [125]). These might lead to a visible change in the charge produced in the sensor and collected by the front-end electronics after irradiation. Indeed:

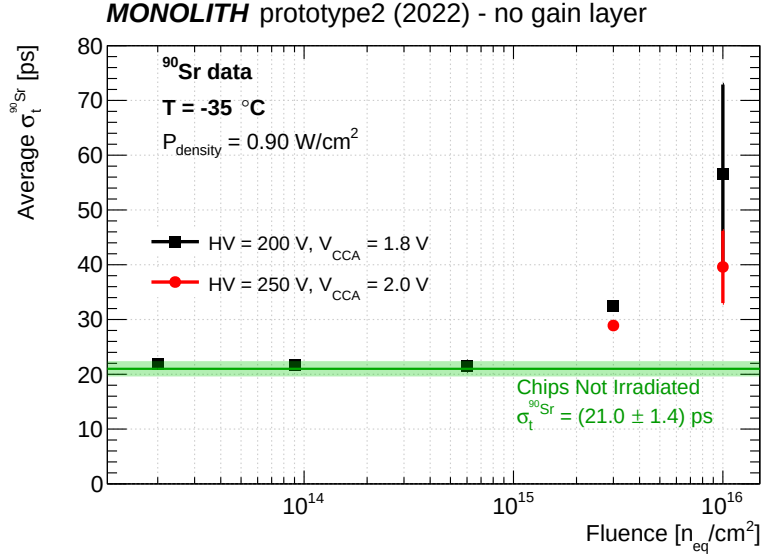


FIGURE 5.31: Average time jitter σ_t^{Sr} measured with the ^{90}Sr radioactive source as a function of the fluence. Data were collected at $-35\text{ }^{\circ}\text{C}$ and $P_{\text{density}} = 0.9\text{ W/cm}^2$, and with two different working points: $HV = 200\text{ V}$ and $V_{\text{CCA}} = 1.8\text{ V}$ (black squares), or $HV = 250\text{ V}$ and $V_{\text{CCA}} = 2.0\text{ V}$ (red dots). For each proton fluence, the time jitter was calculated independently for each analog pixel and then averaged across all pixels. The error bars displayed represent the standard deviations of the average values. The green line represents the average value measured using the 12 analog pixels of the three non-irradiated ASICs, operated under the same conditions as the black squares; the green band represents the standard deviation of the average value.

- if the resistivity of the substrate changes, the voltage needed to fully deplete a sensor before irradiation may not be sufficient to deplete the sensor after irradiation fully, and the primary charge produced is less.
- The traps formed in the silicon substrate by radiation may capture part of the drifting charge so that the total charge collected by the preamplifier is less than the total charge collected by a sensor not irradiated.

Since both effects could be mitigated by increasing the sensor bias voltage, the consequence of irradiation on the sensor substrate was studied using data collected with the ^{90}Sr source at different sensor bias voltage values. To single out variations of performance that could be attributed solely to the sensor and avoid contributions from the electronics, the parameters of operation of the electronics were kept at the standard working-point values of $I_{\text{preamp}} = 50\text{ }\mu\text{A}$, $I_{\text{fbk}} = 2.0\text{ }\mu\text{A}$ and $V_{\text{CCA}} = 1.8\text{ V}$ during the sensor bias voltage scan.

Figure 5.32 shows the dependence of the mode of the distribution of the signal amplitudes on the sensor bias voltage for the ASIC not irradiated of the first row of Table 5.5 and the ASIC irradiated to $1 \times 10^{16}\text{ n}_{\text{eq}}/\text{cm}^2$ of the ninth row of Table 5.5.

In the case of the ASIC not irradiated, the amplitude mode does not vary with the sensor bias voltage, demonstrating that 200 V is a sufficient bias to fully deplete the sensor. On the other hand, in the case of the ASIC irradiated to $1 \times 10^{16}\text{ n}_{\text{eq}}/\text{cm}^2$, the mode of the amplitude is found to increase with increasing sensor bias.

The ^{90}Sr source is an almost pure emitter of β^- particles that deposit charge all along their path in the substrate; thus, the amount of charge produced is expected to increase for larger depleted volumes. Additionally, an increase in the depleted volume with increasing sensor bias voltage also decreases the sensor's capacity, resulting in

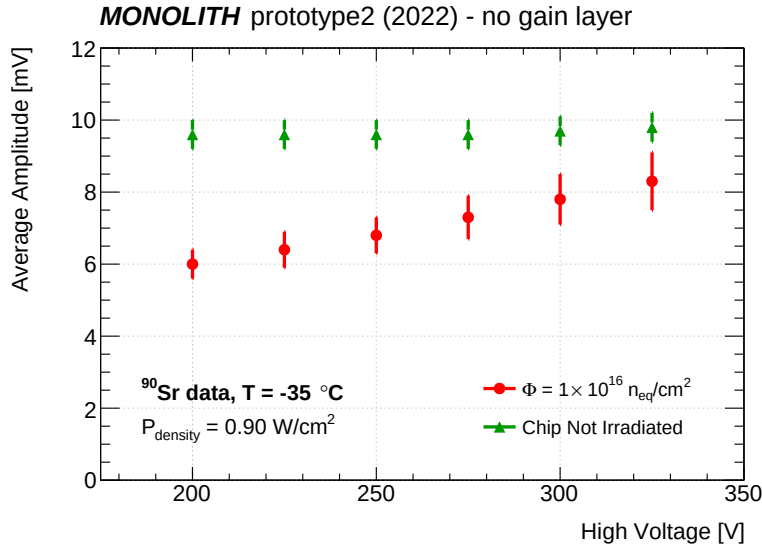


FIGURE 5.32: Average mode of the signal amplitude measured with ^{90}Sr as a function of the sensor bias voltage for one of the three ASICs not irradiated and one of the two ASICs irradiated to $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$. The values plotted represent the mode of the amplitude distributions obtained by fits using a Landau functional form. The amplitude was calculated independently for each of the four analog pixels in an ASIC and then averaged. The error bars displayed represent the standard deviation of the average values.

a higher signal amplitude for the same amount of charge. The trend observed in Figure 5.32 suggests that increasing the sensor bias voltage above 200 V increases the depletion volume for the irradiated ASIC, thus producing more charge and reducing the capacitance. A concurrent reduction of charge trapping with increasing bias voltage might also contribute to the measured trend.

Time Jitter vs. Sensor Bias

To quantify the effect of this change in signal amplitude on the timing performance, Figure 5.33 shows the average time jitter of the most probable signal amplitude $\sigma_t^{90\text{Sr}}$ as a function of the sensor bias voltage, measured for the same two ASICs of Figure 5.32.

When the sensor bias voltage is varied between 200 and 325 V, the value of $\sigma_t^{90\text{Sr}}$ remains flat at approximately 20 ps for the ASIC not irradiated. On the contrary, it gradually improves from $\sigma_t^{90\text{Sr}} = (46.4 \pm 3.5) \text{ ps}$ to $(38.0 \pm 1.9) \text{ ps}$ for the ASIC irradiated to $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$.

5.4.4 Test Beam Experimental Setup

A test-beam campaign was conducted in the summer of 2023 to test the irradiated samples that had already been characterized in the laboratory. The primary motivations were to study the performance degradation after proton irradiation in terms of efficiency and time resolution.

The experimental setup of the test beam was the same as the one used for the test beam measurements of the second MONOLITH prototype without a gain layer (see Figure 5.4 in Section 5.2.1). In particular, the UNIGE FE-I4 telescope [114] was used to provide the tracking information. The Device Under Test (DUT) was mounted at the center of the telescope on a cooling plate and covered by an insulating box to

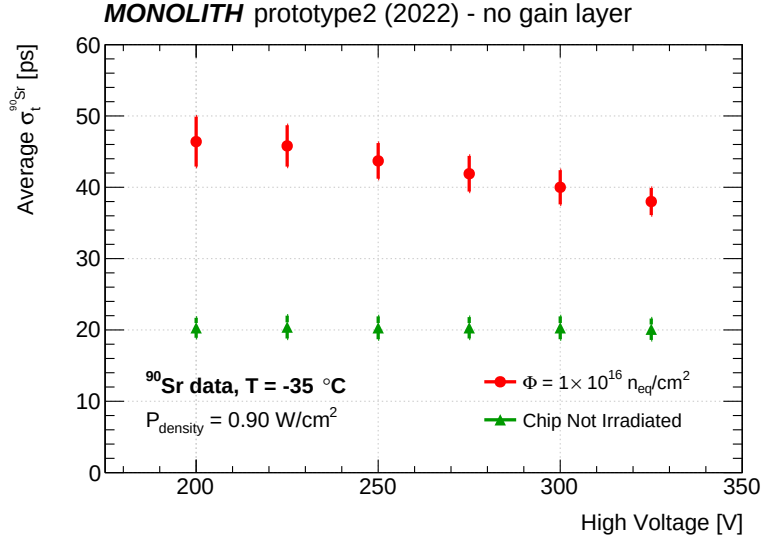


FIGURE 5.33: Average time jitter $\sigma_t^{90\text{Sr}}$ measured with the ^{90}Sr radioactive source as a function of the sensor bias voltage for one of the two ASICs irradiated to $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$ (red dots) and for one of the non-irradiated ASICs (green triangles). Data were collected at -35°C . The front-end electronics were operated at the standard working point with $I_{\text{preamp}} = 50 \mu\text{A}$ (corresponding to $P_{\text{density}} = 0.9 \text{ W}/\text{cm}^2$), $I_{\text{fbk}} = 2.0 \mu\text{A}$, and $V_{\text{CCA}} = 1.8 \text{ V}$. The time jitter was calculated independently for the four analog pixels and then averaged. The error bars displayed represent the standard deviation of the average values.

maintain a constant temperature of -10°C for the ASIC. Two MicroChannel Plate detectors (MCPs) were positioned downstream of the telescope and used as a time reference.

Two fast oscilloscopes read the analog outputs of the two MCP detectors and the DUT. The connections and configurations of the oscilloscope are slightly different from those of the previous test beam. The first oscilloscope has a sampling rate of 20 GSamples/s and an analog bandwidth set to 4 GHz for the two channels connected to the MCPs and 1 GHz for the two channels connected to the pair of single-ended analog outputs of the pixel used to measure the DUT time resolution. Two other pairs of single-ended outputs of adjacent pixels were attached to the second oscilloscope with a sampling rate of 10 GSamples/s and an analog bandwidth set to 1 GHz.

The waveforms acquired for the DUT's positive and negative single-ended signals were subtracted offline, point by point, to construct the differential signals. Only the differential signals were used throughout the analysis since they provide a better signal-to-noise ratio.

The usual quality cuts on the tracks, as already described in Section 5.2.1, were also used here, i.e., only one track crossing all six telescope planes and whose χ^2/NDF was less than 1.5.

5.4.5 Test Beam Data Samples

Given the limited time availability during the test beam experiment, data were taken only with four irradiated boards out of the seven characterized in the laboratory in the previous Section 5.4.2. In addition, data were also retaken with the ASIC not irradiated, and the results are presented in Section 5.2.

Table 5.6 summarizes the power density of the front-end electronics and the sensor bias voltage of the 18 datasets taken at the test beam. All boards were operated at a

front-end feedback current of $2.0 \mu\text{A}$ and, to avoid any unwanted annealing, stored and operated at a temperature of -10°C .

Proton Fluence [$1 \text{ MeV } n_{\text{eq}}/\text{cm}^2$]	P_{density} [W/cm^2]			High Voltage [V]			
0	0.13	0.54	0.9	200			
9×10^{13}	0.13	0.54	0.9				
6×10^{14}	0.13	0.54	0.9				
3×10^{15}	0.13	0.54	0.9				
1×10^{16}	0.13		0.54	250			
			0.9	150	200	250	300

TABLE 5.6: Power density and sensor bias voltage of the 18 data samples taken at the CERN SPS test beam with the four irradiated boards and the board not irradiated. The proton-fluence values are reported in the first column in $1 \text{ MeV } n_{\text{eq}}/\text{cm}^2$ values.

5.4.6 Efficiency

Since the number of channels available in the two oscilloscopes allowed recording of the signals from only two out of six neighboring analog pixels, the calculation of the efficiency was restricted to the telescope tracks whose extrapolation on the pixel surface was inside the region defined by the triangle connecting the center of the three pixels (the pixel under study and the two adjacent pixels that were read out). This triangle represents all the areas of a pixel in the correct proportions. It provides an unbiased result based on the telescope's pointing resolution, which has been observed to exhibit artificial inefficiency at the edge of pixels if the adjacent pixel is not read out.

Threshold Scan

Figure 5.34 shows the efficiency as a function of the voltage noise σ_V , in the case of the board not irradiated operated at a sensor bias voltage of 200 V and of the board irradiated at $1 \times 10^{16} n_{\text{eq}}/\text{cm}^2$ operated at 300 V.

The increase in HV for the irradiated board was necessary to fully deplete the sensor and operate it at very high efficiency after the change in resistivity due to radiation exposure. The voltage noise σ_V was measured event-by-event using the waveform samplings recorded in a time interval of 200 ns preceding the signal.

At the nominal threshold value $V_{\text{th}} = 7\sigma_V$, the ASIC not irradiated provides a detection efficiency of $99.96^{+0.01}_{-0.02}\%$ and, given the large signal-to-noise ratio, it maintains an efficiency at the level of 99.8 % even with a threshold value of $14\sigma_V$. On the contrary, in the case of the ASIC that was exposed to $1 \times 10^{16} n_{\text{eq}}/\text{cm}^2$, since the signal-to-noise ratio diminishes, the efficiency was found to depend more on the threshold. At $HV = 300 \text{ V}$ varies from 99.7 % at a threshold $V_{\text{th}} = 7\sigma_V$ to approximately 97.0 % already at $V_{\text{th}} = 10\sigma_V$.

Figure 5.34 also reports the noise-hit rate as a function of the threshold. The noise-hit rate shows the expected exponential drop. It increases by a factor of approximately five after the ASIC has received a proton fluence of $1 \times 10^{16} n_{\text{eq}}/\text{cm}^2$.

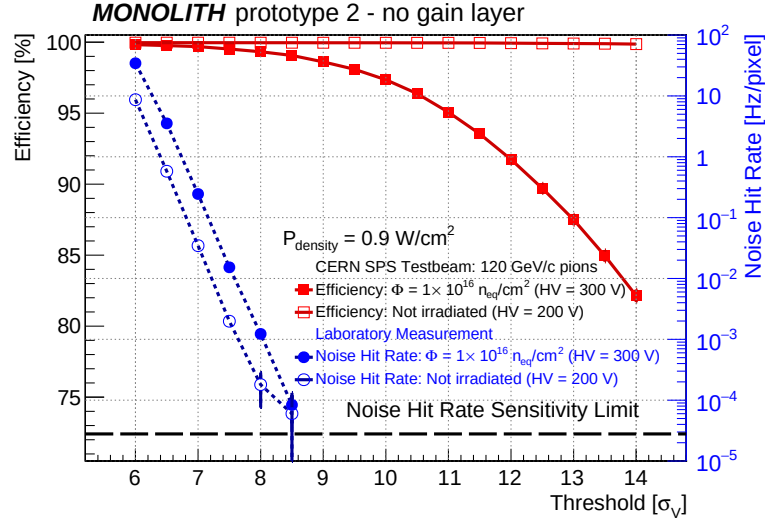


FIGURE 5.34: Detection efficiency measured at a power density $P_{\text{density}} = 0.9 \text{ W/cm}^2$ as a function of the voltage threshold. The data refer to the ASIC not irradiated operated at $HV = 200 \text{ V}$ (open red squares), and to the ASIC irradiated at $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$ operated at $HV = 300 \text{ V}$ (full red squares). The voltage threshold is given in units of the voltage noise σ_V , which is $300 \mu\text{V}$ before irradiation and raises to $600 \mu\text{V}$ after a fluence of $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$. The plot also reports the noise-hit rate (values on the right-hand axis scale) measured with the two ASICs.

Bias Voltage and Power Density Scans

Figure 5.35 shows the detection efficiency for different fluence values: on the left panel as a function of the sensor bias voltage, on the right panel as a function of the power density at which the front-end electronics were operated.

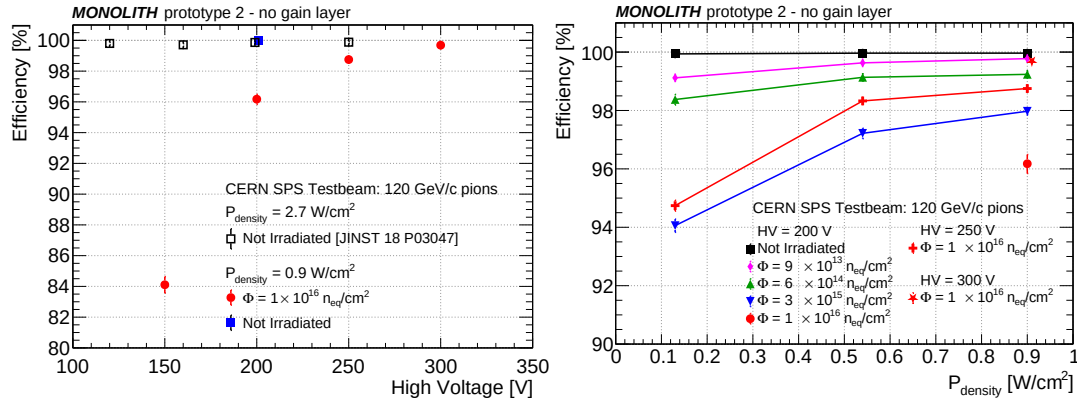


FIGURE 5.35: Detection efficiency measured at a threshold value $V_{\text{th}} = 7\sigma_V$ for different fluences: (left) as a function of the sensor bias voltage, (right) as a function of the front-end electronics power density. The black open squares on the left plot are results obtained with the ASIC not irradiated at $P_{\text{density}} = 2.7 \text{ W/cm}^2$ from the previous test beam (see Section 5.2.3). The five ASICs proton fluence between zero and $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$ were operated at sensor bias voltage $HV = 200 \text{ V}$. The sample irradiated at $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$ was also operated at a sensor bias voltage of 250 V and 300 V .

Only one data set at $HV = 200 \text{ V}$ was acquired during this test beam with the ASIC not irradiated. On the left panel, the data taken with the ASIC not irradiated in the previous test beam (see Section 5.2.3) are superimposed as open black squares for

comparison. It should be noted that the ASIC that was not irradiated was operated in different conditions between this test beam and the previous one (see Section 5.2). In addition to the various power density, the data during the last test beam were acquired at feedback current $I_{fbk} = 0.1 \mu\text{A}$ and a temperature of 20°C in contrast with the $2.0 \mu\text{A}$ and -10°C of the present data. This change in operating conditions increases the signal-to-noise ratio, resulting in slightly better efficiency and time resolution.

The left panel of Figure 5.35 shows that after $1 \times 10^{16} \text{ n}_{eq}/\text{cm}^2$ a bias voltage of 300 V is just enough to reach the detection efficiency plateau. This behavior is entirely different from that of the non-irradiated ASIC, for which the efficiency plateau is reached already at $HV = 120 \text{ V}$. This measurement, as well as the drop in detection efficiency shown in figure 5.34, could in large part be attributed to the change in resistivity of the silicon bulk, which we estimated to vary from the initial $350 \Omega \text{ cm}$ to approximately $50 \Omega \text{ cm}$ extrapolating the data of [125, 126]. The $50 \mu\text{m}$ thick sensor would not be fully depleted at such resistivity. Concurrent radiation damage of the front-end electronics cannot be excluded and should be measured by irradiating the SiGe HBT alone.

The right panel of Figure 5.35 shows that, at a sensor bias voltage $HV = 200 \text{ V}$, a steady increase of the efficiency is measured as P_{density} increases. This effect is more pronounced at the highest proton fluences. In the case of the ASIC irradiated to $1 \times 10^{16} \text{ n}_{eq}/\text{cm}^2$, at $HV = 200 \text{ V}$ data were taken only at $P_{\text{density}} = 0.9 \text{ W}/\text{cm}^2$, which result in a detection efficiency of $(96.2 \pm 0.3) \%$ (red circle in the figure). For this ASIC, the scan in P_{density} was performed at $HV = 250 \text{ V}$, and the red crosses display the results. At $HV = 300 \text{ V}$, only a dataset at $P_{\text{density}} = 0.9 \text{ W}/\text{cm}^2$ was acquired (red star in the figure), which corresponds a detection efficiency of $99.69^{+0.06}_{-0.07} \%$.

Efficiency in Different Regions of the Pixel

Finally, the detection efficiency is shown as a function of the distance from the center of the pixel in Figure 5.36. All the data points shown refer to a sensor bias voltage of 200 V, except the dataset at $1 \times 10^{16} \text{ n}_{eq}/\text{cm}^2$, which was obtained at 300 V.

At no or minor irradiation fluences, the efficiency does not depend on the position within the pixel area. However, starting from a fluence of $6 \times 10^{14} \text{ n}_{eq}/\text{cm}^2$, the efficiency is measured to drop in the inter-pixel area. However, increasing the sensor bias voltage allows for the recovery of efficiency far from the pixel center, as shown by the data at $HV = 300 \text{ V}$ (red stars).

5.4.7 Time Resolution

The time resolution of the DUT was measured using the timestamps of two precise MCP detectors located outside the FE-I4 telescope, downstream of the pion beam, as a reference. The sample of telescope tracks used for measuring the time resolution was the same as that used for detection efficiency, with the only exception that all tracks within a pixel were considered here, rather than just the tracks within the triangle defined by the three analog pixel centers mentioned above.

For this test beam, the time of arrival (ToA) was taken when the signals reached 50 % of the maximum value of the signal amplitude. For both the DUT and the MCPs, an offline linear interpolation between the oscilloscope samplings gave this ToA. Taking the ToA at a fixed percentage of the signal amplitude has the advantage of correcting the time walk between signals with different amplitudes, assuming that signals of all amplitudes have consistent rise times. Figure 5.37 shows the

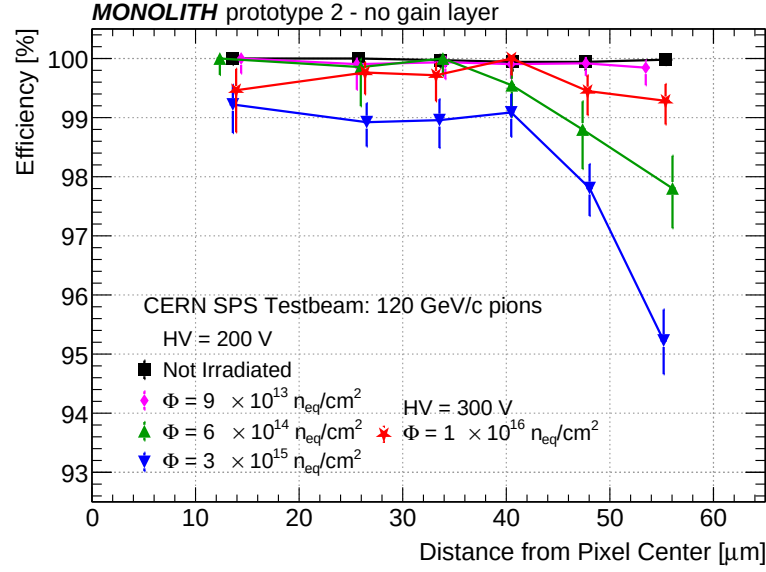


FIGURE 5.36: Detection efficiency as a function of the distance from the pixel center. The data refer to power density $P_{\text{density}} = 0.9 \text{ W/cm}^2$, voltage threshold $V_{\text{th}} = 7\sigma_V$ and sensor bias voltage $HV = 200 \text{ V}$, with the exception of the data taken at $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$, for which the sensor bias voltage was 300 V . The efficiency was computed in the following bins of distance from the pixel center in microns: $[0 - 21]$, $[21 - 30]$, $[30 - 37]$, $[37 - 44]$, $[44 - 51]$, $[51 - 65]$. In each bin, the data points are plotted at the mean track distance from the center of the pixel.

distributions of $\text{ToA}_{\text{DUT}} - \text{ToA}_{\text{MCP}}$ for the not irradiated ASIC (left panel), and the ASIC irradiated at $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$ (right panel).

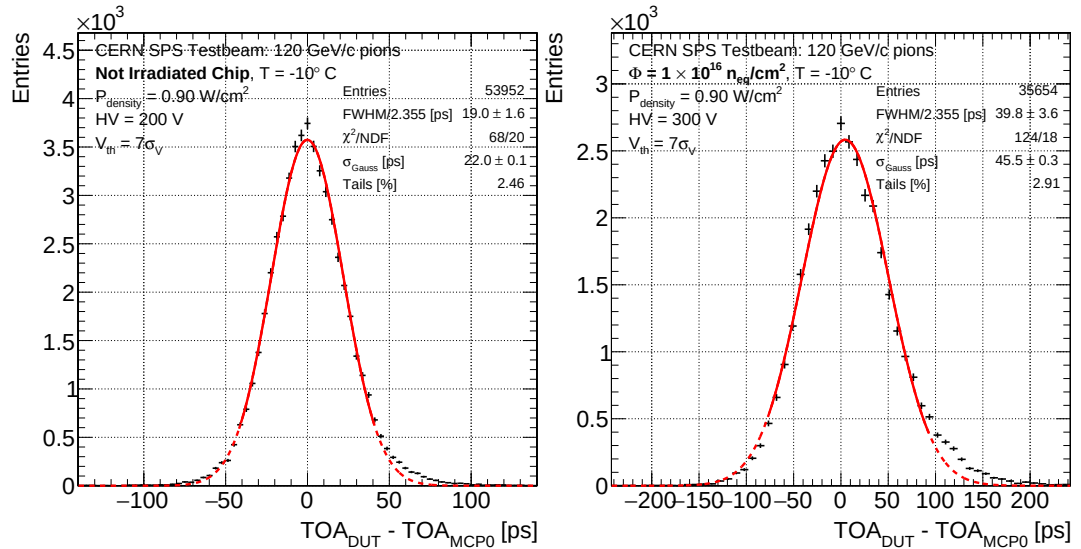


FIGURE 5.37: Difference in ToA between the DUT and one of the MCP detectors used to provide a precise reference time. The left-hand panel shows the results for the board not irradiated, while the right-hand panel shows those for the board irradiated to $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$. A Gaussian fit extending to the first bin containing 20 % of the maximum value of the distribution is superimposed on the distributions.

It was observed this method carries the drawback of making more symmetric the distributions of the ToA difference, by amplifying the typically minor distortion which deviates slightly from a Gaussian distribution and generates a tail at positive

values of $\text{ToA}_{\text{DUT}} - \text{ToA}_{\text{MCP}}$. The causes of slight non-Gaussianity were discussed in Section 5.5. Their effect is visible in Figure 5.37.

The distributions of ToA difference in Figure 5.37, were fitted with Gaussian functions up to the bins containing more 20 % of the entries of the bin with the highest number of entries to extract the σ_{Gauss} . These values for the three pairs of detectors were then used to compute the time resolution of the DUT, as done in the previous test beam.

Bias Voltage and Power Density Scans

Figure 5.38 shows the time resolution for different values of fluence: on the left panel, as a function of the sensor bias voltage, on the right panel, as a function of the power density at which the front-end electronics were operated.

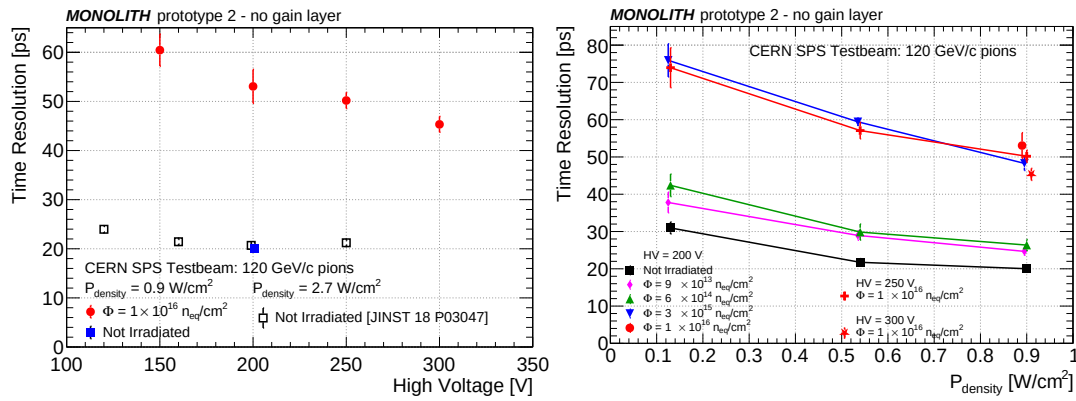


FIGURE 5.38: Time resolution measured at a constant threshold of 50 % for different fluences: (left) as a function of the sensor bias voltage, (right) as a function of the front-end electronics power density. The black open squares on the left plot are the results obtained with the ASIC not irradiated at $P_{\text{density}} = 2.7 \text{ W/cm}^2$ from the previous test beam (see Section 5.2.4). The five ASICs proton fluence between zero and $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$ were operated at sensor bias voltage $HV = 200 \text{ V}$. The sample irradiated at $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$ was also operated at a sensor bias voltage of 250 V and 300 V.

For comparison, the left panel of Figure 5.38 also reports the time resolution measured with the ASIC not irradiated at $P_{\text{density}} = 2.7 \text{ W/cm}^2$ in the previous test beam (see Section 5.2.4).

The left panel of Figure 5.38 shows that, while the ASIC not irradiated reaches the plateau of time resolution already at $HV = 160 \text{ V}$, the time resolution measured with the ASIC irradiated at $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$ does not seem to have yet reached a plateau value at 300 V.

The right panel of Figure 5.38 shows that the time resolution steadily improves with increasing P_{density} and degrades with increasing proton fluence. In the P_{density} range studied, it varies between 30 and 20 ps in the case of the ASIC not irradiated and between 75 and 50 ps in the case of the ASIC irradiated to $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$.

Time Resolution in Different Regions of the Pixel

The time resolution is shown as a function of the distance from the center of the pixel in Figure 5.39. All the data refer to a sensor bias voltage of 200 V, with the exception of the dataset at $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$, which was acquired at 300 V.

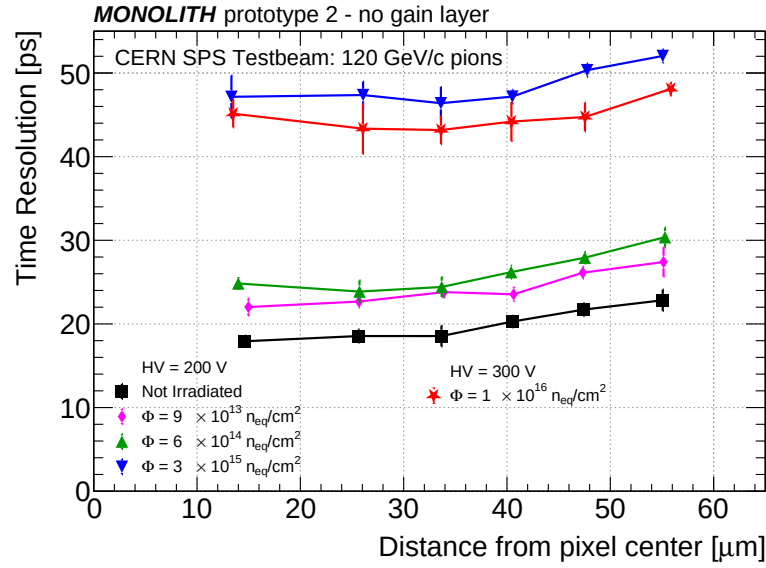


FIGURE 5.39: Time resolution as a function of the distance from the pixel center. The data refer to $P_{\text{density}} = 0.9 \text{ W/cm}^2$, $V_{\text{th}} = 7\sigma_V$ and $HV = 200 \text{ V}$, with the exception of the data taken at $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$, for which the sensor bias voltage was 300 V. The time resolution was computed in the same bins as figure 5.36.

Figure 5.39 shows that, for all the proton fluences, the time resolution is better in the central region of the pixel than in the inter-pixel region, where it increases by approximately 5 ps.

Time Resolution vs. Proton Fluence

To give more emphasis to the effects of the proton fluence on this second MONOLITH prototype without a gain layer, the results of time resolution already presented were simplified in a single plot as a function of the proton fluence, which is shown in Figure 5.40 at $P_{\text{density}} = 0.9 \text{ W/cm}^2$.

The results obtained with the datasets taken at sensor bias voltage $HV = 200 \text{ V}$ vary between 20 ps for the board not irradiated and 53 ps for the board irradiated to $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$. For the last ASIC, increasing the sensor bias voltage to 300 V provides a time resolution of 45 ps, thus improving it by approximately 10 %.

It should be noted that, for each of the irradiation values studied, the ratio of the time resolutions at $P_{\text{density}} = 0.13$ and 0.9 W/cm^2 is constant, showing a relative improvement of approximately 35 % independently of the fluence, as reported in Table 5.7.

Table 5.7 indicates that the dependence of the transition frequency from the collector current, which is observed in [122] not to change up to $1 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$, might not be affected by radiation damage even at $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$.

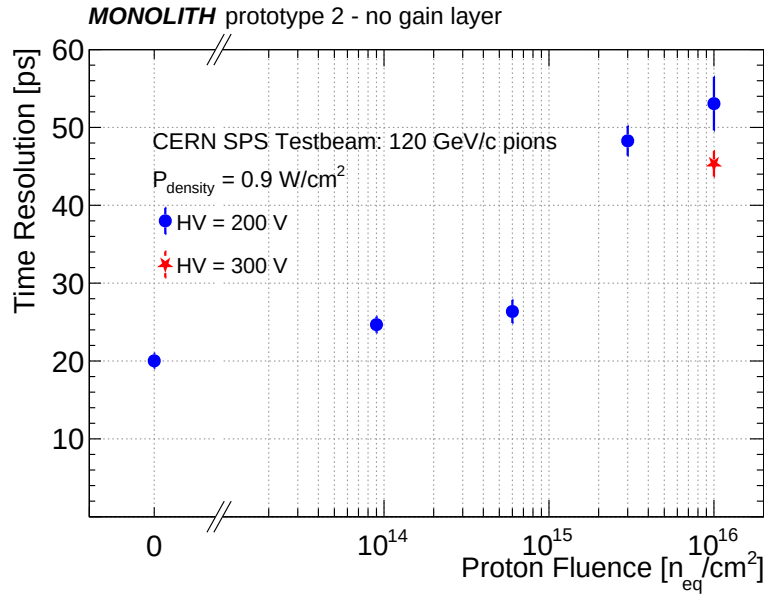


FIGURE 5.40: Time resolution as a function of the proton fluence for $P_{\text{density}} = 0.9 \text{ W/cm}^2$. The blue dots show the results obtained at a sensor bias voltage $HV = 200 \text{ V}$, while the red star shows the result obtained at $HV = 300 \text{ V}$ with the board irradiated to $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$.

Fluence [1 MeV $\text{n}_{\text{eq}}/\text{cm}^2$]	Time Resolution [ps]		Ratio
	$P_{\text{den.}} = 0.13 \text{ W/cm}^2$	$P_{\text{den.}} = 0.9 \text{ W/cm}^2$	
0	31.0 ± 1.6	20.0 ± 1.0	1.55 ± 0.15
9×10^{13}	37.8 ± 2.8	24.7 ± 1.0	1.53 ± 0.21
6×10^{14}	42.4 ± 3.0	26.4 ± 1.5	1.61 ± 0.21
3×10^{15}	75.9 ± 4.5	48.3 ± 1.9	1.57 ± 0.17
1×10^{16}	74.0 ± 5.4	50.2 ± 1.6	1.57 ± 0.19

TABLE 5.7: Time resolution measured with the five ASICs at $P_{\text{density}} = 0.13$ and 0.9 W/cm^2 and their ratio. The data were taken at sensor bias voltage $HV = 200 \text{ V}$ for all boards except the board irradiated to $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$, for which the sensor bias voltage was 250 V .

5.5 Garfield++ Simulations

Accurate modeling of the physical processes governing signal formation in silicon detectors is essential for understanding the ultimate performance limits of timing systems. To this end, Garfield++ simulations were employed to complement and verify the experimental measurements of this second MONOLITH prototype without a gain layer.

First, we provide a brief description of the simulation framework and methodology. Using electric field maps derived from TCAD simulations and front-end electronics delta response derived from Cadence simulations, the Garfield++ framework was configured to replicate realistic operating conditions.

Then, we address a critical validation step: comparing the simulated time resolution with the measured one from a test beam. Discrepancies between simulation and experiment were carefully analyzed to identify possible sources of mismatch, such as simplifications in the preamplifier modeling. Several techniques were implemented

to confirm the amount of discrepancy and correct the simulation with appropriate compensating factors.

Then, we focus on estimating the charge-collection noise, a key factor in determining the time resolution. Simulating numerous charge depositions and analyzing the resulting signal fluctuations enabled the quantification of the impact of carrier statistics and spatial charge distribution on the overall time jitter. A comparison with the results of the femtosecond laser measurements confirmed the measured behavior of the charge-collection noise contribution to the time resolution at varying charge levels.

Finally, we investigate a problem observed in all test beam measurements: a deviation from the ideal Gaussian distribution of arrival times. These tails in the timing distributions were reproduced in much simpler simulation conditions, making it possible to understand that they are a mathematical feature of the amplifier shaping of the induced current signals.

5.5.1 The Simulation Framework

Garfield++ is a C++ toolkit designed explicitly for simulating ionization, charge transport, and signal formation processes in particle detectors [97]. Initially intended for gaseous detectors, it has since been extended to handle solid-state detectors, such as silicon. It is a versatile platform for studying timing, charge collection dynamics, and signal induction in modern fast-timing devices. In the context of the MONOLITH project, Garfield++ was utilized to simulate the microscopic processes underlying signal generation and time resolution in the monolithic silicon prototypes.

The simulation chain begins by importing a detailed electric-field map derived from TCAD simulations [82]. This field governs the drift and diffusion of electron-hole pairs created inside the silicon active layer. Garfield++ models the motion of individual carriers under the influence of the electric field. The drift trajectories are computed step by step, and the motion of each charge is used to evaluate the induced current on the electrodes using the Shockley–Ramo theorem.

The instantaneous current signals calculated by Garfield++ are convolved with a delta response function derived from Cadence simulations [108] that characterizes the shaping behavior of the preamplifier. The data sample of the obtained output signals is analyzed using the same software as the experimental data to extract the arrival time and calculate the time resolution.

5.5.2 Test Beam Time Resolution Matching

The first objective of the Garfield++ simulations is to match the results of time resolution collected at a test beam. We chose one of the best result, which is (20.0 ± 1.0) ps and it was achieved with a reduced power consumption of $P_{\text{density}} = 0.9 \text{ W/cm}^2$, a high feedback current of $I_{\text{fbk}} = 2000 \text{ nA}$, a low temperature of $T = -10^\circ\text{C}$, and an high voltage $HV = 200 \text{ V}$ (see Section 5.4.7). In these working conditions, the core ASIC performance metrics were measured in the laboratory (see Section 5.1.4), giving a preamplifier gain A_q equal to $(56 \pm 2) \text{ mV/fC}$, an ENC equal to (130 ± 5) electrons (not corrected for the oscilloscope contribution), and a signal rise time equal to $(364 \pm 4) \text{ ps}$.

Profile of the Electric Field from TCAD Simulations

The one-dimensional electric-field map along the depth of the silicon active layer was derived from TCAD simulations [82]. Figure 5.41 shows the electric field and the drift velocity for electrons and holes as a function of the depth of the detector for a bias voltage of $HV = 200$ V. The drift velocities were calculated starting from the electric field using Equation 2.14.

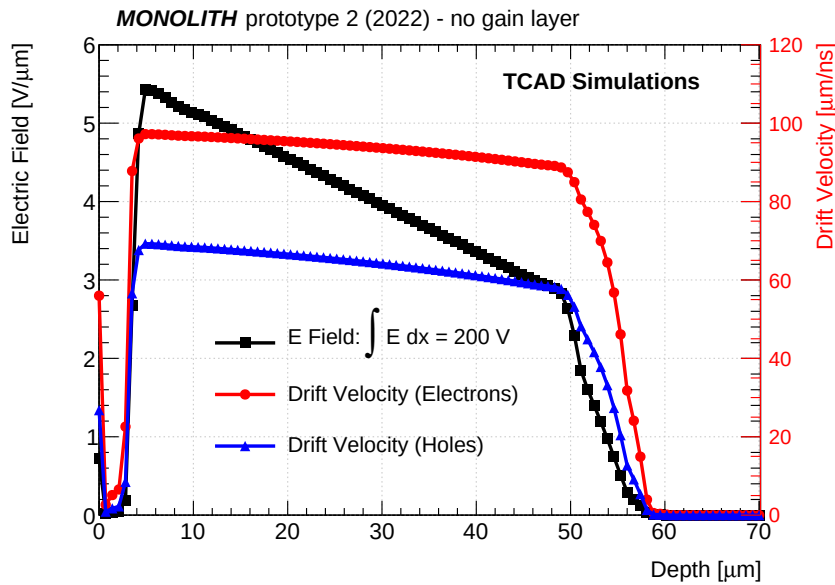


FIGURE 5.41: Electric field and drift velocity for electrons and holes as a function of the depth of the detector for a bias voltage of $HV = 200$ V. The electric field profile was simulated with TCAD, and the drift velocities were calculated starting from the electric field using Equation 2.14.

The thickness of the active layer of this second MONOLITH prototype without a gain layer is nominally $50\text{ }\mu\text{m}$ of epitaxial layer. However, the simulations show that it is slightly lower. The collection electrode, towards which the electrons drift, at a depth of $0\text{ }\mu\text{m}$, has a very high doping necessary to host the electronics and creates a few μm of low electric field and, by consequence, low drift velocity. Charge carriers created in this region do not induce current on the collection electrode; therefore, this region must not be considered active.

The region between $50\text{ }\mu\text{m}$ and $70\text{ }\mu\text{m}$ in Figure 5.41 represents a small portion of the $300\text{ }\mu\text{m}$ -thick wafer substrate, which is highly p-doped. The TCAD simulations accounted for the effects of dopant diffusion during thermal annealing in the fabrication process. Consequently, an additional $20\text{ }\mu\text{m}$ of substrate was retained to account for these effects. As shown in Figure 5.41, the electric field in this region gradually decreases to zero between $50\text{ }\mu\text{m}$ and $60\text{ }\mu\text{m}$, resulting in a slow component in the induced current from charges generated there.

Current Induction

The electric field profile visible in Figure 5.41 was imported into the Garfield++ framework. An external program called *High Energy Electrons and Delta-rays* (HEED) was used within the Garfield++ framework to model the microscopic ionization processes [98]. It gave the distribution of primary clusters of electron-hole pairs for a

Minimum Ionizing Particle (MIP). Then, Garfield++ simulated the drift under the given electric field and the diffusion of this cluster's distribution. The movement of charge induced an instantaneous induced current according to the Shockley-Ramo theorem [105, 106], as explained in Section 3.3.

Figure 5.42 shows a set of induced currents and their average signals, obtained with the TCAD electric field of Figure 5.41 and the charge deposition generated by HEED for MIPs shot perpendicularly to the detector surface.

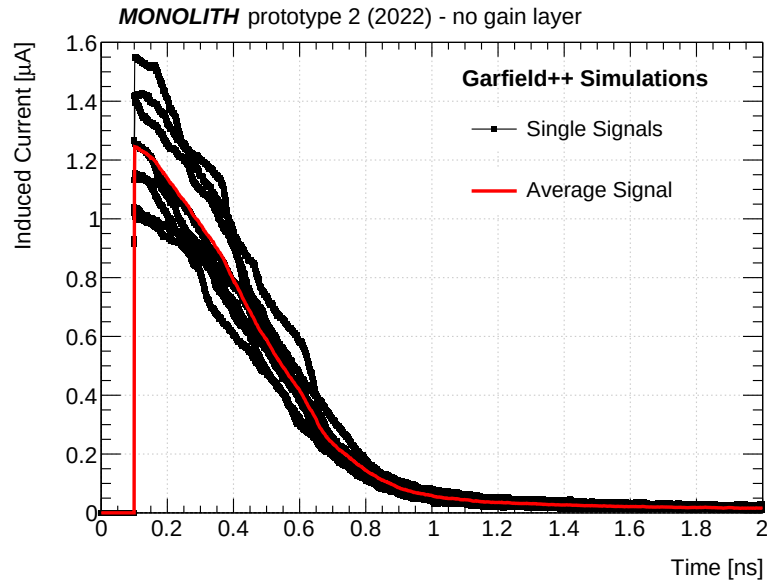


FIGURE 5.42: Examples of induced current signals and their average, as calculated from the Garfield++ simulations using the TCAD electric field of Figure 5.41 and the charge deposition generated by HEED.

Figure 5.42 shows that the single signals have some fluctuations in their shape. With the collection of a cluster of charges by the collection electrode, the cluster is absorbed and suddenly stops contributing to the induction, causing a sudden drop in the current. These fluctuations are referred to as *charge-collection noise* and contribute to the detector's total time resolution.

The time interval between consecutive simulation points was set to 5 ps, a parameter adjustable via software; however, reducing this interval increases computational demands. Notably, the onset of signal induction occurs within a timeframe shorter than this interval. Specifically, the signal's leading edge manifests at a sub-picosecond scale, as electrons can accelerate to velocities of 1×10^7 cm/s in approximately 0.14 ps, and a minimum ionizing particle (MIP) traverses a 50 μm thick silicon sensor in about 0.16 ps. While theoretically, setting a discrimination threshold at this initial signal rise could yield exceptional time resolution, achieving this would necessitate amplification electronics with infinite bandwidth, which is a practical impossibility due to the resultant infinite noise.

Front-end Electronics Noise and Amplification

As a final step, the instantaneous current signals calculated by Garfield++ were convoluted with the analog front-end electronics. The electronics noise was added as a Gaussian smearing of each point of the induced current with a standard deviation equal to the measured ENC of (130 ± 5) electrons. The signals obtained were then

convolved with the amplifier's delta response, as given by Cadence simulations [108], following Equation 3.9.

The data sample of the obtained output signals is analyzed using the same software as that for the experimental data and compared with the test beam measurements. It is worth noticing that only a subsample of test beam tracks was selected: the one falling in a circular area of radius $30\text{ }\mu\text{m}$ around the pixel center. This selection was done to avoid the effects of charge sharing, the weighting field, and electric field non-uniformities in the inter-pixel region. Moreover, the test beam tracks were estimated to have an angle of at most 0.7° , which is reasonably straight, similar to the MIP used during the HEEP simulation.

The initial comparison between the simulated events and the test beam data revealed a discrepancy: the rise time of the simulated signals was approximately $(220 \pm 4)\text{ ps}$, significantly shorter than the measured $(364 \pm 4)\text{ ps}$. A detailed analysis of the simulations identified the source of the mismatch: the Cadence simulations produced an amplifier delta response that was unrealistically fast. Several independent methods were employed to evaluate the discrepancy between the simulated and actual delta responses, all yielding compatible correction factors. These corrections were subsequently applied to the simulations, resulting in good agreement with the measured data.

The three methods focused on determining the correction factor k to be applied to the simulated amplifier delta response $h^{\text{SIM}}(t)$ to match the experimental response $h^{\text{EXP}}(t)$. Since the simulated delta response was observed to be faster than the measured one, it was evident that the correction needed to stretch the time axis of the simulation. Accordingly, the relationship was defined as $h^{\text{EXP}}(t) = h^{\text{SIM}}(t/k)$, $k > 1$ acts as a stretching factor, slowing the simulated delta response. The three methods used were:

- the comparison of the Fourier transforms;
- the comparison of the frequency of zero crossing;
- the comparison of the autocorrelation functions.

Comparison of the Noise Fourier Transforms

The first method involved comparing the Fourier transforms of the noise. Considering events without signals from MIPs $n^{\text{EXP}}(t)$, the Fourier transforms were calculated using the Fast-Fourier transform algorithm: $\mathcal{F}[n^{\text{EXP}}(t)] = N^{\text{EXP}}(\omega)$. On the other hand, events containing only noise were directly simulated with the Cadence tool using the same amplifier delta response $h^{\text{SIM}}(t)$ as the one used for the Garfield++ simulations, and giving simulated noise waveforms $n^{\text{SIM}}(t)$. The Fourier transforms were calculated as well using the Fast-Fourier transform algorithm: $\mathcal{F}[n^{\text{SIM}}(t)] = N^{\text{SIM}}(\omega)$.

When the amplifier delta response $h(t)$ is stretched by a factor $k > 1$, i.e., $h(t) \rightarrow h(t/k)$, the noise signal is also stretched by the same factor, i.e., $n(t) \rightarrow n(t/k)$. For the Fourier transforms, it holds:

$$\mathcal{F}[n(t)] = N(\omega) \rightarrow \mathcal{F}[n(t/k)] = k N(k\omega), \quad (5.10)$$

where ω is the angular frequency.

Therefore, the stretching factor $k > 1$, which slows down the simulated delta response, has the opposite effect of compressing the frequency axis of the Fourier

transform, while introducing a multiplicative factor on the amplitude. The experimental Fourier transform $N^{\text{EXP}}(\omega)$ matched the one for the stretched simulated delta response $k N^{\text{SIM}}(k\omega)$ when the stretching factor was found to be (2.66 ± 0.08) .

Comparison of the Frequency of Zero Crossing

The second method involved comparing the frequency of zero crossings. Again, we considered only noise signals $n^{\text{SIM}}(t)$ simulated with Cadence, and measured experimentally $n^{\text{EXP}}(t)$. The frequency of zero crossing f_{ZC} is the frequency with which the voltage signal crosses the signal baseline threshold, and it can be measured directly by counting the zero crossings in a certain amount of time, or calculated using the formula:

$$f_{\text{ZC}} = 2 \sqrt{\frac{\int_0^\infty f^2 w(2\pi f) df}{\int_0^\infty w(2\pi f) df}}, \quad \text{with } w(2\pi f) = \lim_{T \rightarrow \infty} \frac{2}{T} |N(\omega)|^2. \quad (5.11)$$

where $w(2\pi f)$ is the noise power spectrum, calculate the noise power spectrum starting from the Fourier transform $N(\omega)$ of the noise signal $n(t)$.

It can be analytically shown that a stretching factor $k > 1$ that slows down the simulated delta response $h(t) \rightarrow h(t/k)$ also affects the frequency of zero crossing in the same way: $f_{\text{ZC}} \rightarrow f_{\text{ZC}}/k$. The test beam experimental frequency of zero crossing $f_{\text{ZC}}^{\text{EXP}}$ matched the one for the stretched simulated delta response $f_{\text{ZC}}^{\text{SIM}}/k$ when the stretching factor was (2.67 ± 0.04) . This value is extremely close to the one obtained by comparing the Fourier transforms of the noise, confirming that both methods are consistent and lead to the same result.

Comparison of the Noise Autocorrelation Function

The third and final method involved comparing the autocorrelation functions of the noise. The autocorrelation function $R(\tau)$ of the noise $n(t)$ is defined as:

$$\mathcal{R}[n(t)] = R(\tau) = \int_{-\infty}^{\infty} n(t + \tau) n(t) dt. \quad (5.12)$$

As said before, when the amplifier delta response $h(t)$ is stretched by a factor $k > 1$, i.e., $h(t) \rightarrow h(t/k)$, the noise signal is also stretched by the same factor, i.e., $n(t) \rightarrow n(t/k)$. For the autocorrelation functions of the noise, it holds:

$$\mathcal{R}[n(t)] = R(\tau) \rightarrow \mathcal{R}[n(t/k)] = k R(\tau/k), \quad (5.13)$$

where ω is the angular frequency.

Therefore, the stretching factor $k > 1$, which slows down the simulated delta response, has an analogous effect on the autocorrelation function of the noise, while introducing a multiplicative factor to the amplitude. The experimental autocorrelation function $R^{\text{EXP}}(\tau)$ matched the one for the stretched simulated delta response $k R^{\text{SIM}}(\tau/k)$ when the stretching factor was found to be (2.5 ± 0.5) . This method was less accurate, but it still yielded a result compatible with the other two.

Comparison of the Signal Rise Time and the Amplitude

After all three methods confirmed that the amplifier delta response was too fast, and they all agreed on a correction factor of 2.66, the amplifier delta response was

effectively stretched by this factor by multiplying the time coordinate. Subsequently, the convolution step with the induced current signals was repeated with Garfield++.

The comparison between test beam data and the new simulated dataset was smooth. Figure 5.43 compares signal rise time (left) and amplitude (right).

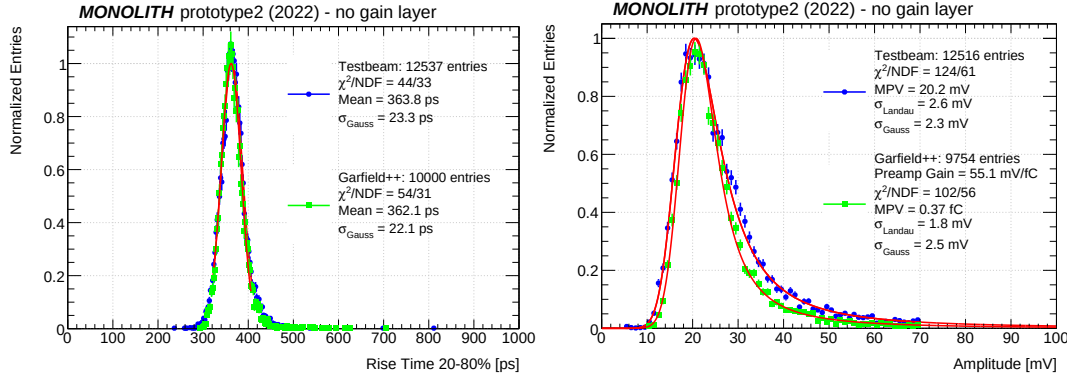


FIGURE 5.43: Comparison between test beam and simulated data of signal rise time (left) and signal amplitude (right).

The left panel of Figure 5.43 shows that the rise time now matches smoothly, both in the mean value and in the width of the distribution, after the appropriate correction factor of 2.66.

The signal amplitudes from the Garfield++ simulations are in units of charge, i.e., fC, whereas the test beam data are in units of voltage, i.e., V. Therefore, both signal amplitude distributions were fitted with a convolution of a Landau and a Gaussian function. The Most Probable Values (MPVs) from the fits were matched, as shown in the right panel of Figure 5.43, by multiplying the simulated MPV by the ratio of the two. This ratio corresponds to the preamplifier gain A_q : the value found is (55.1 ± 1.9) mV/fC, which is compatible with the value measured using radioactive sources, (56 ± 2) mV/fC.

In the right panel of Figure 5.43, the value of σ_{Gauss} from the amplitude fits indicates the effect of the voltage noise on the distribution width: the similar values between experimental and simulated data are a straightforward confirmation that the correct voltage noise was simulated. On the other hand, the value of σ_{Landau} is strictly related to the physics of the interaction. The slight discrepancy between the experimental and simulated values may suggest that some minor effects were not fully considered. Initially, the simulated σ_{Landau} was even smaller, but the simulation was refined by inputting the correct beam composition and the momentum spread measured at the test beam. The nominally 120 GeV/c pion beam contained only about 60 % pions, with the remainder being protons, electrons, positrons, and a small percentage of mixed hadrons, and exhibited a momentum spread up to 1 %.

Comparison of the Time Resolution

As explained in Section 5.4.7, the time of arrival from this test beam sample was taken at a constant fraction of 50 % of the signal height, and no time-walk correction was applied. The distribution of the arrival time difference between the DUT and the time reference (a fast MCP) is visible in the left panel of Figure 5.44. The time of arrival was taken at a constant fraction of 50 % also for the simulated data, and the distribution of time of arrival is visible in the right panel of Figure 5.44.

The time resolution of the DUT from the experimental data ($\sigma_{t,\text{DUT}} = (20.1 \pm 0.9)$ ps in the left panel of in Figure 5.44) must be compared with the Gaussian standard

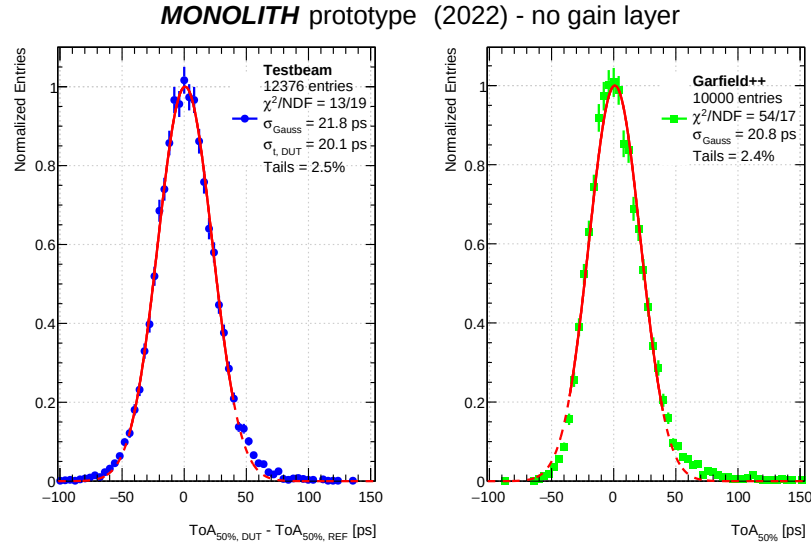


FIGURE 5.44: Comparison between the difference in arrival time between the DUT and the time reference from test beam data (left) and time of arrival distribution from a Garfield++ simulation.

deviation of the distribution of the simulated data ($\sigma_{t,\text{Gauss}} = (20.8 \pm 0.9) \text{ ps}$ in the right panel of in Figure 5.44): the two results are in perfect agreement.

The comparisons in Figure 5.43 and Figure 5.44 confirm that the Garfield++ simulation is correctly set up and can deliver the correct results for the second MONOLITH prototype without a gain layer.

The front-end electronic noise can be turned off by setting the ENC in the Garfield++ simulation to zero, leaving only the purely charge-collection noise contribution to the time resolution. The same analysis on the simulated waveforms resulted in a time resolution of $(11.2 \pm 0.6) \text{ ps}$, indicating that: first, the expected time resolution from the charge-collection noise is achieved (a $\sigma_{\text{Landau}} \simeq 10 \text{ ps}$ was calculated in Section 3.5.1), and, second, the electronics noise time jitter is the most relevant contribution to the time resolution of this second MONOLITH prototype without a gain layer.

5.5.3 Charge-Collection Noise Estimation

The second objective of the Garfield++ simulations is to disentangle the individual contributions to the time resolution arising from electronics time jitter and charge-collection noise, thereby providing essential validation of the femtosecond laser measurements. We chose the same experimental conditions that yielded $(20.7 \pm 0.3) \text{ ps}$, as the femtosecond laser measurements were taken under the same conditions. These conditions are: a high power consumption ($P_{\text{density}} = 2.7 \text{ W/cm}^2$), a feedback current of $I_{\text{fbk}} = 100 \text{ nA}$, and a room temperature of $T = 20^\circ \text{C}$ (see Section 5.2.4). The core ASIC performance metrics were measured in the laboratory, giving a preamplifier gain A_q equal to $(60 \pm 2) \text{ mV/fC}$; an ENC equal to $(124 \pm 5) \text{ electrons}$ (not corrected for the oscilloscope contribution); a signal rise time equal to $(375 \pm 4) \text{ ps}$.

We tuned the amplifier delta response of the Garfield++ simulations as described in the section above to match the slightly lower ENC and the marginally slower rise time. We subsequently verified that the simulation yielded the same results as the test beam data.

Three Garfield++ simulations were performed using the same tuned amplifier delta response.

- The first simulation, also used to calibrate the working point, included the correct Landau-distributed charge deposition for a MIP and the nominal ENC of (124 ± 5) electrons. This simulation is designed to reproduce the test beam data for MIP-like charge deposition, incorporating the contributions of charge collection noise and amplifier time jitter to the time resolution.
- The second simulation used uniform charge deposition and had a nominal ENC of (124 ± 5) electrons. The uniform charge deposition was achieved by redistributing the charge clusters uniformly; thus, the total charge per event still follows a Landau distribution, but the spatial distribution of the clusters within an event is uniform. This simulation aims to reproduce the femtosecond laser data for a variable injected charge, isolating only the amplifier time jitter contribution to the time resolution, with no contribution from charge-collection noise.
- The third simulation included the correct Landau charge deposition for a MIP, but with the ENC set to zero electrons. This case is designed to reproduce the difference in quadrature between test beam and femtosecond laser measurements, assuming no correlations are present. It includes only the charge-collection noise contribution to the time resolution, with no amplifier time jitter.

Figure 5.45 compares the results between test beam data, femtosecond laser measurements, and the three different simulations.

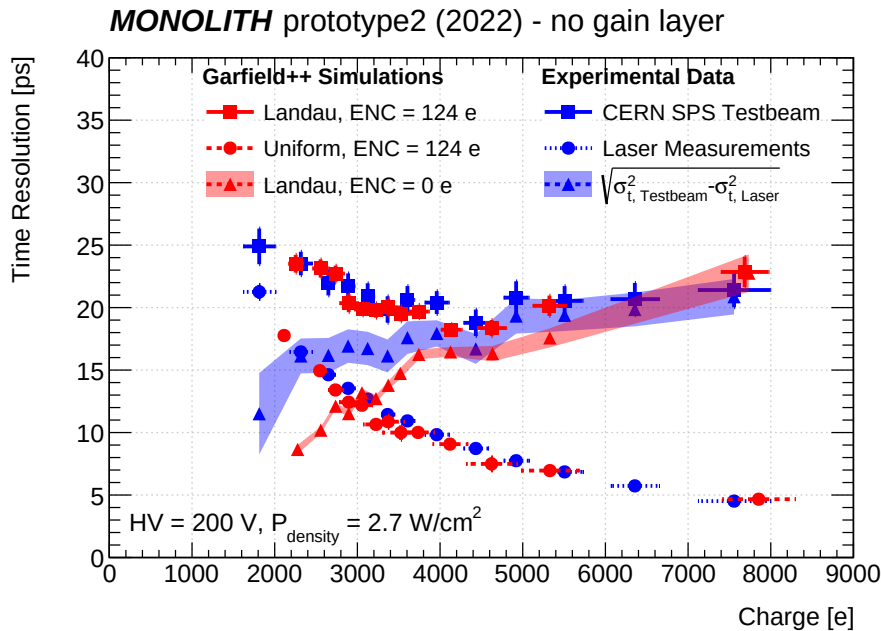


FIGURE 5.45: Time resolution as a function of the injected charge: comparison of the test beam, femtosecond laser results, and three different simulations. The experimental data are the same of Figure 5.25, for bias voltage $HV = 200 \text{ V}$ and power density $P_{\text{density}} = 2.7 \text{ W/cm}^2$.

Figure 5.45 shows that the first simulation (*Landau*, $ENC = 124 \text{ e}$) correctly reproduces the time resolution measured at the CERN SPS test beam. The same figure

demonstrates that the second simulation (*Uniform*, $ENC = 124 e$) accurately reproduces the time resolution obtained from the femtosecond laser measurements. Finally, Figure 5.45 shows that the third simulation (*Landau*, $ENC = 0 e$) correctly reproduces the increasing trend of the charge-collection noise contribution as a function of the collected charge. Still, it does not match the data corresponding to $\sqrt{\sigma_{t,\text{Testbeam}}^2 - \sigma_{t,\text{Laser}}^2}$.

A simple explanation for this discrepancy is that the initial assumption of no correlation between the two noise contributions was incorrect. If no correlations were present, the sum in quadrature of the time resolution due solely to the charge-collection noise, σ_{Landau} (*Landau*, $ENC = 0 e$), and the time jitter due solely to electronics noise, σ_{jitter} (*Uniform*, $ENC = 124 e$), would yield the total time resolution σ_t (*Landau*, $ENC = 124 e$). However, it was confirmed that $\sqrt{\sigma_{\text{Landau}}^2 + \sigma_{\text{jitter}}^2} < \sigma_t$, thus confirming the presence of a non-negligible correlation between the two contributions.

5.5.4 Non-Gaussianity of the Distributions of Time of Arrival

Throughout all the test beam campaigns, the distributions of the time of arrival difference between the DUT and the time reference consistently exhibited non-Gaussian features, specifically tails comprising approximately 2 % to 10 % of the total distribution. These tails are typically asymmetric, extending towards later times, indicating delayed responses recorded by the DUT.

Such non-Gaussian tails were observed under several prototypes and analysis configurations. They were present during the first test beam of the second MONOLITH prototype without a gain layer, where the time of arrival was extracted using a fixed threshold of $7\sigma_V$ and a time-walk correction was applied (see Figure 5.11). They also appeared in the second test beam on the same prototype, both in irradiated and non-irradiated ASICs, when using a constant fraction threshold at 50 % of the signal amplitude without applying any time-walk correction (see Figure 5.37). Similar tails were observed in the test beam of the PicoAD proof-of-concept prototype, again using a fixed threshold of $7\sigma_V$ with time-walk correction applied (see Figure 4.29). Finally, they were also reproduced in Garfield++ simulations, where a constant fraction threshold of 50 % was used without time-walk correction (see right panel of Figure 5.45). For the Garfield++ simulations, only the DUT was present, indicating that such non-Gaussian tails could not be attributed to the time reference but rather to delayed responses recorded by the DUT.

Investigation of Non-Gaussian Tails

To investigate the non-Gaussian tails and understand the potential causes, we conducted three additional Garfield++ simulations under progressively simplified configurations:

- **Fast Amplifier Delta Response:** We set the correction factor of the amplifier delta response to zero, effectively speeding up the amplifier shaping and confirmed that the non-Gaussian tails persisted; moreover, their fraction increased from about 5 % to 15 %. These results indicated that the amplifier delta response is not the source of the tails, but a fast amplifier delta response makes the tails more evident.
- **Electronics noise disabled:** We set the ENC to zero and confirmed that the non-Gaussian tails persisted; moreover, their fraction increased from about 15 % to 20 %. These results indicated that electronics noise is not the source of the

tails, and that the ENC's Gaussian smearing helps make the time-of-arrival distribution more Gaussian.

- **Uniform electric field:** We imposed a constant electric field of $5.4 \text{ V}/\mu\text{m}$ (the maximum value shown in Figure 5.41) throughout the sensor depth between $0 \mu\text{m}$ and $60 \mu\text{m}$, ensuring a uniform drift velocity. The non-Gaussian tails remained, but their fraction dropped back to about 16 %, ruling out field non-uniformity as their cause. This configuration also highlighted that the regions near the collection electrode ($0 \mu\text{m}$ – $5 \mu\text{m}$) and near the substrate interface ($50 \mu\text{m}$ – $60 \mu\text{m}$) contribute most strongly to the tails.
- **Landau fluctuations disabled:** We redistributed the charge per event uniformly, thereby removing the Landau fluctuations. This modification eliminated the non-Gaussian tails, confirming that they originate from the Landau charge deposition.

Causes of the Non-Gaussian Tails

Therefore, we returned to the simulation with electronics noise disabled and a uniform electric field to study the causes of the non-Gaussian tails. We saved the distributions of primary clusters and their sizes in electrons.

The left panel of Figure 5.46 shows the simulated distribution of time of arrival at a constant fraction of 50 % of the signal height with fast amplifier delta response, no electronics noise, and uniform electric field. The right panel of Figure 5.46 shows the distribution of maximum cluster size corresponding to different cuts of the time of arrival distribution.

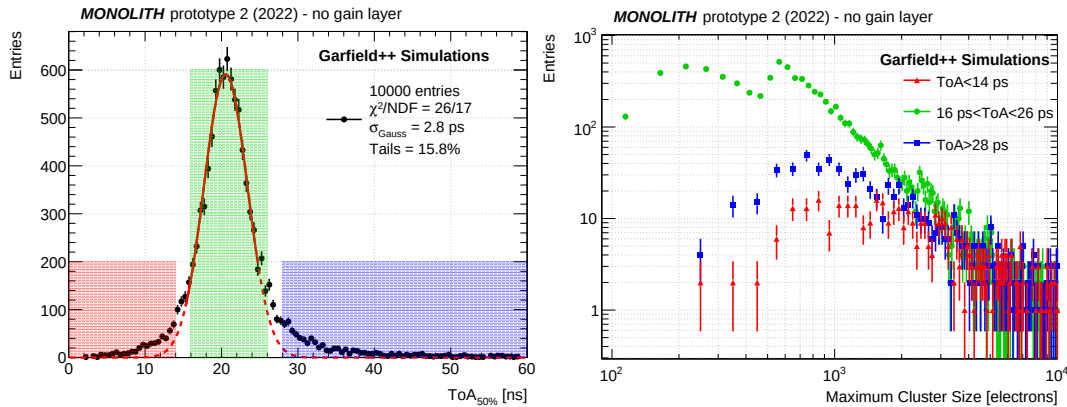


FIGURE 5.46: (Left) Distribution of time of arrival at a constant fraction of 50 % of the signal height from a Garfield++ simulation with fast amplifier delta response, no electronics noise, and uniform electric field. The three regions shaded in color were selected as the left tail, the core, and the right tail. (Right) Distribution of maximum cluster size for the same simulation: the color codes identify different cuts of the time-of-arrival distribution.

Figure 5.46 reveals a clear correlation between the non-Gaussian timing tails and events containing at least one charge cluster larger than 200 electrons. However, the central portion of the time-of-arrival distribution spans a comparably wide range of maximum cluster sizes, indicating that a large cluster is a necessary condition for an event to appear in the tail. Still, it is not sufficient on its own.

Therefore, we also studied the initial depth position of the largest cluster. Figure 5.47 shows the correlation between arrival time and the largest cluster's initial depth position.

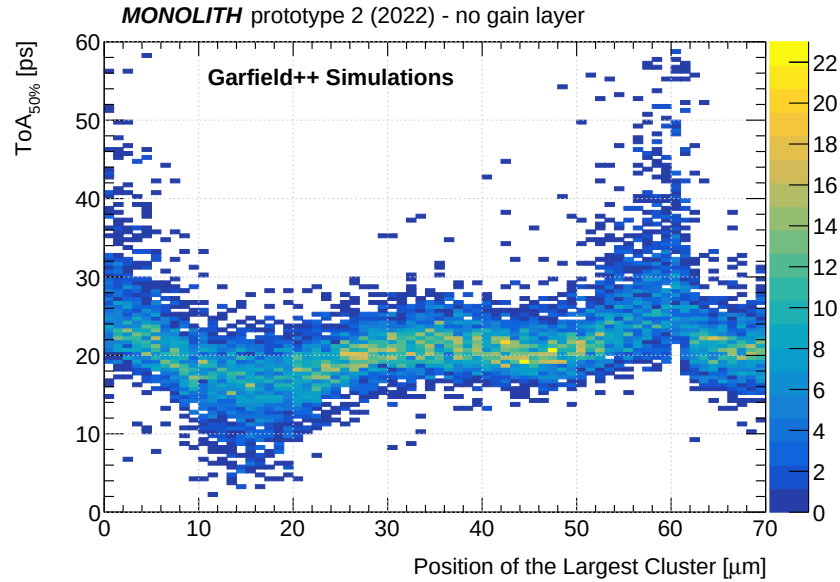


FIGURE 5.47: Time of arrival at a constant fraction of 50 % of the signal height as a function of the position of the largest cluster. The result is from a Garfield++ simulation with a fast amplifier delta response, no electronic noise, and a uniform electric field between 0 μm and 60 μm , plus 10 μm of substrate with zero electric field.

Figure 5.47 shows a clear correlation between the arrival time and the cluster's depth containing the most charge. When a cluster exceeding 200 electrons occurs in the peripheral regions, near the collection electrode (0 μm –5 μm) or at the substrate interface (50 μm –60 μm), it delays the signal, shifting the time of arrival into the right-hand tail of the distribution. Conversely, if such a large cluster is generated at depths between 10 μm and 20 μm , it accelerates the signal rise, causing the time of arrival to occur earlier, into the left-hand tail of the distribution.

Figure 5.47 further shows that when a cluster exceeding 200 electrons is located in the substrate region with zero electric field (60 μm –70 μm), the time of arrival remains in the core of the distribution. Without an electric field, such a cluster does not drift and therefore does not influence the signal rise. As a result, the induced current is governed by the other drifting charge clusters, yielding a time of arrival that falls within the distribution's central (Gaussian) portion.

Interpretation of the Non-Gaussian Tails

To interpret the dependence of the time of arrival on the depth of the largest cluster (Figure 5.47), we modeled the initial charge distribution as a uniform MIP-like background plus a single large cluster of more than 200 electrons at depth $x \in [0, 60]$ μm . The resulting total induced current is:

$$i(x, t) = i_{\text{MIP}}(t) + i_{\text{SC}}(x, t), \quad (5.14)$$

where $i_{\text{MIP}}(t)$ is given by Equation 3.7 for the uniform MIP-like deposition, and $i_{\text{SC}}(x, t)$ is given by Equation 3.6 for the single cluster at depth x .

The front-end electronics then shape this current via convolution with the amplifier delta response $h(t)$ (Equation 3.10). Applying a threshold V_{th} to the output defines the time of arrival, ToA, through:

$$V_{th} = \int_0^{ToA} h(ToA - t) i(x, t) dt. \quad (5.15)$$

In principle, solving this equation for ToA as a function of x , i.e., $ToA(x)$, would reproduce the curve in Figure 5.47. However, the integral equation is analytically intractable, so we opted for a numerical solution.

The function $ToA(x)$ is governed by the amplifier peaking time τ_p , the detector thickness d , and the voltage threshold V_{th} , which were all set to the values used in the simulation underlying Figure 5.47. In addition, $ToA(x)$ depends on the single-cluster size N , which we fixed to the mean value of the right panel of Figure 5.46, namely $N_{SC} = 950$ electrons.

Figure 5.48 displays the numerical evaluation results superimposed on the average y-axis values from Figure 5.47.

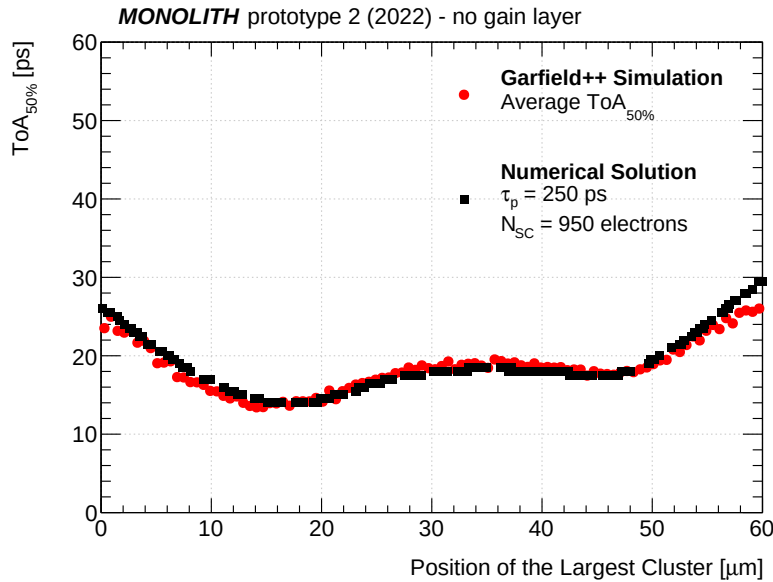


FIGURE 5.48: Time of arrival at a constant fraction of 50 % of the signal height as a function of the position of the largest cluster. The red dots are the profile of the distribution in Figure 5.47. The black square are the result of a numerical solution of Equation 5.15, with amplifier peaking time $\tau_p = 250$ ps, and single cluster size $N_{SC} = 950$ electrons.

Figure 5.48 demonstrates a strong agreement between the full Garfield++ simulation and the numerical solution of Equation 5.15. The resulting curve exhibits a double minimum, which arises from the interplay between the center-of-gravity arrival times of the induced currents and the amplifier's finite peaking time. Since the amplifier peaking time, $\tau_p = 250$ ps, is comparable to the average center-of-gravity times, the convolution with the delta response is essential for shaping the output voltage waveform. By contrast, in [101], the output signal was treated as a constant pulse equal to the amplifier delta response shifted by the center-of-gravity time, reflecting the assumption of a much longer amplifier integration time relative to the induced-current timescales.

Chapter 6

MONOLITH Prototype 2 with a Gain Layer

This chapter presents the main results obtained from laboratory measurements and test-beam studies, which were used to validate and optimize the silicon sensor design of the second MONOLITH prototype with a gain layer. The ASIC design and the gain layer layout were discussed in Section 1.2.7.

Section 6.1 covers laboratory measurements performed before beam tests. The energy-calibration runs with radioactive sources, particularly a ^{55}Fe X-ray source and a ^{90}Sr β^- source, are analyzed to extract the avalanche gain given by the gain layer.

Section 6.2 presents the results of a test-beam campaign, where sensors are exposed to minimum-ionizing particles at a CERN SPS test beam. A high-resolution silicon-pixel telescope provides precise track references, while a fast-sampling oscilloscope reads the Device Under Test (DUT). We present the efficiency results as a function of the signal threshold, bias voltage, and power density, with a focus on variations across the pixel cell. Then, we measure the time resolution by using the time of threshold crossing relative to a fast time reference, after applying time-walk corrections to improve temporal precision.

Some of the results presented in this chapter have been published previously. In particular, the test beam results are presented in [8].

6.1 Laboratory Measurements

The gain response of the second MONOLITH prototype featuring a gain layer was characterized using a ^{55}Fe X-ray source and a ^{90}Sr β^- source. Measurements were conducted inside a climate chamber at two ambient temperatures: -5°C , matching the conditions of the subsequent test beam campaign, and 20°C . The analog pixel outputs were read out by an oscilloscope with a 1 GHz analog bandwidth, and the resulting waveforms were recorded for offline analysis. The measurements were conducted in the clean room laboratory at the University of Geneva during the first months of 2024.

As described in Section 4.2, the ^{55}Fe source emits X-ray photons with an energy of 5.9 keV. Depending on the absorption depth, avalanche multiplication is initiated either by electrons or holes, resulting in different values for the impact ionization gain since their impact ionization coefficients differ [111]. Specifically, X-rays absorbed below the gain layer (in the absorption region) generate electron-driven avalanches. In contrast, those absorbed above the gain layer (in the drift region) produce hole-driven gain.

As reported in Section 4.2 and confirmed by TCAD simulations (see Figure 4.11), transient space-charge effects reduce the electric field in the gain layer avalanche

development, therefore reducing the electron gain for ^{55}Fe X-ray photons. For this reason, a different method was used with this prototype to calculate the electron gain, which also involved a ^{90}Sr β^- source.

A ^{90}Sr β^- source behaves like a MIP in silicon (see Section 3.2.1), depositing on average $\tilde{q}_{\mu\text{m}}$ electron-hole pairs per μm . Electrons generated below the gain layer (in the absorption region of thickness d_{abs}) drift through the gain layer toward the collection electrode, undergoing avalanche multiplication with gain G_e . Conversely, holes generated above the gain layer (in the drift region of thickness d_{drift}) drift away from the collection electrode and experience avalanche multiplication with gain G_h . Within a 10 % uncertainty on both region thicknesses, the total collected charge is then:

$$Q = G_e \tilde{q}_{\mu\text{m}} d_{\text{abs.}} + G_h \tilde{q}_{\mu\text{m}} d_{\text{drift.}} \quad (6.1)$$

The analog front-end electronics convert the collected charge into a voltage signal according to the Shockley–Ramo theorem (see Section 3.3). We assume the output amplitude measured by the oscilloscope, amp , is directly proportional to the collected charge, Q , via the preamplifier gain, G_{preamp} : $\text{amp} = G_{\text{preamp}} Q$. Substituting Q from Equation 6.1 and solving for the electron gain, G_e , yields:

$$G_e = \frac{\text{amp}}{G_{\text{preamp}}} \frac{1}{\tilde{q}_{\mu\text{m}} d_{\text{abs.}}} - G_h \frac{d_{\text{drift}}}{d_{\text{abs.}}} \quad (6.2)$$

The preamplifier gain G_{preamp} was determined using ^{55}Fe and ^{109}Cd radioactive sources, as well as an infrared laser, all of which yielded consistent results. Even at the minimum bias voltage required for complete depletion, a residual electron multiplication is present, as evidenced by the double-peak structure in the ^{55}Fe spectrum. To minimize avalanche gain during calibration, we ensured that photon-induced charge clusters were generated in the drift region, used the lowest gain layer dose, and set the high voltage, HV , to the depletion threshold. Under these conditions, the measured preamplifier gain was $G_{\text{preamp}} = 49 \text{ mV/fC}$, which can be overestimated by up to 10 % due to residual avalanche gain that also holes can have.

The hole-initiated avalanche gain, G_h , was determined using a ^{55}Fe source following the procedure of Section 4.2. In the absence of any observed space-charge suppression for holes, G_h was taken as the mean charge corresponding to the first peak of the ^{55}Fe spectrum divided by the preamplifier gain. Although the preamplifier gain remains constant with bias voltage HV , G_h varies; therefore, we recorded ^{55}Fe spectra at each operating point and extracted G_h for all bias settings.

After determining the preamplifier gain, G_{preamp} , and the hole gain, G_h , we recorded voltage-amplitude spectra using the ^{90}Sr source at various bias voltages and temperatures for ASICs with different gain layer doses. Each spectrum was fitted with a Landau distribution to extract its Most Probable Value (MPV), which we take as the average amplitude amp in Equation 6.2. Substituting these amplitudes into Equation 6.2 then yielded the electron gain G_e for each set of operating conditions.

Figure 6.1 presents the electron gain G_e as a function of sensor bias voltage for operating temperatures of -5°C and 20°C . The data correspond to ASICs featuring a $15\mu\text{m}$ -thick drift layer and a $5\mu\text{m}$ -thick absorption layer, fabricated with three gain-layer doses: 3, 3.5, and 4. The ASIC with the highest dose (dose 4) is the same device tested in the beam campaign described in Section 6.2.

Figure 6.1 demonstrates that the electron gain G_e rises with increasing bias voltage HV . As expected, sensors with lower gain-layer implant doses exhibit lower gain at the same high voltage (HV). Furthermore, reducing the temperature from 20°C to -5°C shifts the G_e – HV curve by approximately 30 V, effectively increasing G_e at any

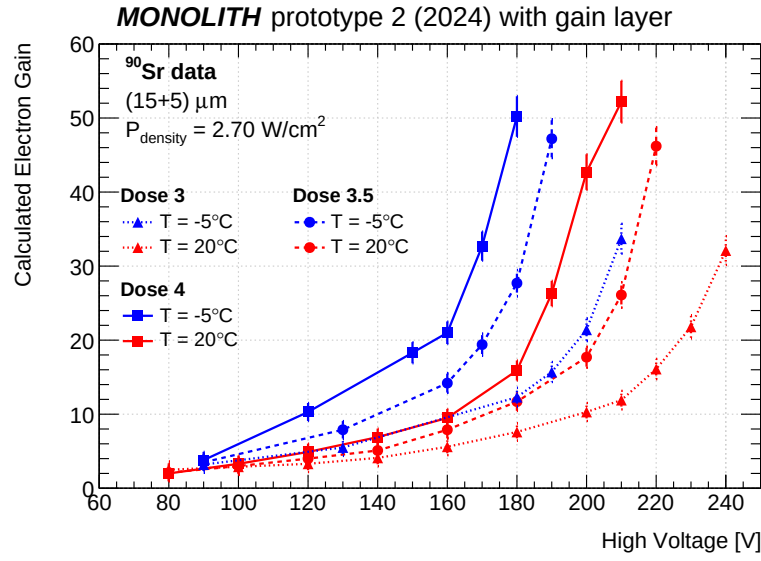


FIGURE 6.1: Electron gain of a ^{90}Sr source as a function of the bias voltage HV for two different temperatures and three different gain layer doses of the prototype with a 15 μm -thick drift layer and a 5 μm -thick absorption layer. The electron gain was calculated using Equation 6.2.

fixed bias. All measurements were taken up to 10 V below the sensor's breakdown voltage, ensuring safe operation. At that limit, devices with doses of 3.5 and 4 achieve similar maximum gains of around 50, whereas the dose-3 device only reaches a gain of nearly 32, likely due to constraints in the ASIC design rather than the gain layer. Consequently, doses 3.5 and 4 are functionally interchangeable by adjusting HV .

6.2 Test Beam

This section presents the measurements of the second MONOLITH prototype with a gain layer performed during a CERN SPS test-beam campaign with pions with momentum 120 GeV/ c . The measurements were conducted in the experimental *North Area* at the Preveessin site of CERN during April and May 2024.

First, we briefly discuss the experimental setup used for this test beam, which is the same as the one already used for the test beam of the second MONOLITH prototype without a gain layer for an ASIC not irradiated (see Section 5.2.1) and irradiated ASICs (see Section 5.4.4).

Second, we present the efficiency and noise-hit ratio results, along with their dependence on bias voltage, power density, and particle position. These measurements show that introducing a gain layer in this prototype improved efficiency and made it more uniform across the pixel area.

Finally, we present the results of time resolution after a time-walk correction is applied to the data. As for efficiency, the time resolution is also mapped into the available parameter space of bias voltage, power density, and the particle's position. These results also demonstrate that introducing a gain layer in this prototype was beneficial in terms of time resolution, resulting in a more uniform signal across the pixel area. The most significant improvement in time resolution was achieved at low power densities, where the introduction of the gain layer significantly enhances performance.

6.2.1 Experimental Setup

Data were taken at the CERN SPS test beam facility using pions with a momentum of 120 GeV/c. The experimental setup is the same as the previous test beams, except for some differences in the connections of the oscilloscopes, and can be seen in Figure 5.4. It comprises:

- the monolithic silicon ASIC device under test (DUT). Only one ASIC flavor was tested after being selected based on laboratory measurements. It was a second MONOLITH prototype with a gain layer of the highest dose with an epitaxial layer of 5 μm and 15 μm ;
- the UNIGE FE-I4 test beam telescope [114] to reconstruct the tracks of the pions and give the trigger to the two oscilloscopes;
- two Photonis PP2365AC micro-channel plate photomultiplier tubes (called MCP0 and MCP1) to provide a precise reference time, which were mounted outside the FE-I4 telescope downstream of the pion beam;
- an oscilloscope Lecroy WaveMaster 820 Zi-A (40 GSamples/s and set to 4 GHz analog bandwidth); its four channels were used to read the two polarities of the differential signal from one of the analog channels of the DUT as well as the two MCPs;
- an oscilloscope Lecroy WaveRunner 9404 (20 GSamples/s and set to 4 GHz analog bandwidth); its four channels were used to read the positive single-ended polarities of the remaining three analog pixels of the DUT and to check the shape of the trigger given by the telescope.

The DUT was placed in the middle of the six planes of the telescope in a cooling box and was operated through all the data acquisition at a temperature of -5°C . A region of interest of $250\ \mu\text{m} \times 250\ \mu\text{m}$ centered around the four analog pixels of the DUT was imposed on the first telescope plane. The coincidence between this region and the last plane of the telescope triggered the system.

For each triggered event, the waveforms of the DUT and the two MCPs were acquired by the oscilloscope in a time window of 500 ns, with the DUT and MCP signals crossing the threshold around the center of the time window. The first 200 ns, which did not contain the signals, was used to measure the baseline voltage noise and its standard deviation, σ_V .

Figure 6.2 shows an example of waveforms of the DUT and the MCP0, zoomed around the signal peak of the DUT. Each dot represents a sample in amplitude acquired with 25 ps time binning.

6.2.2 Data Samples

During the test-beam campaign, data were acquired at different working conditions of the ASIC, varying the sensor bias voltage and the power density of the front-end electronics. The minimum value of the sensor bias voltage was $HV = 90\ \text{V}$, corresponding to an electron gain of 4. The maximum value was 180 V, which is about 20 V before the sensor breakdown, corresponding to an electron gain of 50. The front-end electronics were operated with absorbed currents between 3 and 160 μA , corresponding to power densities between 0.05 and 2.6 W/cm^2 .

Table 6.1 lists the data samples collected. For each operating point, a few million tracks were recorded. The offline selection criteria were the same as those applied in

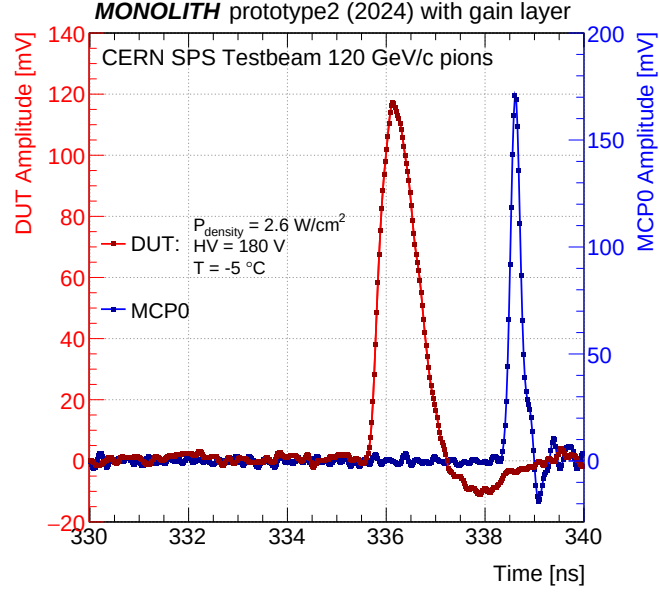


FIGURE 6.2: Example of waveforms acquired by the oscilloscope in a 10 ns time interval around the DUT signal (dots interpolated with red segments) and the MCP0 signal (dots interpolated with blue segments). The dots represent the amplitude values acquired with a 25 ps time binning, while the segments connecting the dots show the linear interpolation used during data analysis.

Section 5.2.1, i.e., only one reconstructed telescope track per event with a recorded hit in all six planes of the telescope and a reconstructed $\chi^2/NDF < 1.6$. Moreover, two more requests were made for accurate time-walk correction: only one hit in the MCP0 and the MCP1 to ensure a precise time reference, and only DUT signals whose amplitudes were not saturated by the oscilloscope’s readout scale.

6.2.3 Signal Amplitudes

For each dataset, the DUT signal amplitudes were distributed in bins of the distance from the hit position to the pixel center and fitted with a Landau function, which allowed for the extraction of the most probable value (MPV) of the Landau distribution. Figure 6.3 illustrates a comparison of the Landau MPV for the current prototype and the proof-of-concept version (see Section 4.3.3).

In the current design, a 20 % decrease in MPV is observed from the pixel center to the inter-pixel region, compared to a nearly eightfold reduction in the earlier version. This improvement results from two key sensor modifications: increasing the drift layer thickness from 10 μA to 15 μA and raising the active layer resistivity from 50 $\Omega\text{ cm}$ to 350 $\Omega\text{ cm}$ in the current prototype.

6.2.4 Efficiency

The efficiency was measured using the sample of all events with the telescope track crossing the DUT plane within the area of the four analog pixels. The tracks extrapolated within 15 μm from the external edge of the surface of the four pixels were kept in the sample to account for the telescope’s spatial resolution [114]. An event was defined as efficient if it had a signal in the DUT with an amplitude larger than 7 times the voltage noise σ_V .

HV [V]	Electron Gain	P_{density} [W/cm ²]	Threshold [electrons] ([mV])	Selected Events	Efficiency [%]
90	4	0.05	895 (3)	8'499	43.7 ^{+1.5} _{-2.8}
		0.1	950 (4)	4'191	70.3 ^{+1.0} _{-2.0}
		0.3	790 (5)	12'666	95.5 ^{+0.3} _{-0.5}
		1.0	725 (7)	11'022	96.8 ^{+0.3} _{-0.6}
		2.6	840 (9)	12'339	98.4 ^{+0.2} _{-0.4}
120	10	0.05	895 (3)	11'406	69.9 ^{+0.4} _{-0.7}
		0.1	950 (4)	8'199	90.5 ^{+0.4} _{-0.8}
		0.3	790 (5)	10'425	99.1 ^{+0.1} _{-0.3}
		1.0	725 (7)	4'542	99.3 ^{+0.2} _{-0.4}
		2.6	840 (9)	31'524	99.73 ^{+0.04} _{-0.09}
150	18	0.05	895 (3)	14'505	93.7 ^{+0.2} _{-0.4}
		0.1	950 (4)	9'558	99.1 ^{+0.1} _{-0.3}
		0.3	790 (5)	34'416	99.65 ^{+0.03} _{-0.07}
		1.0	725 (7)	9'567	99.77 ^{+0.04} _{-0.12}
		2.6	840 (9)	44'430	99.98 ^{+0.01} _{-0.03}
160	21	0.3	790 (5)	16'290	99.84 ^{+0.02} _{-0.05}
		1.0	725 (7)	9'009	99.83 ^{+0.03} _{-0.07}
		2.6	840 (9)	7'704	99.97 ^{+0.02} _{-0.06}
170	33	0.05	895 (3)	17'364	99.27 ^{+0.07} _{-0.15}
		0.1	950 (4)	9'612	99.73 ^{+0.07} _{-0.18}
		0.3	790 (5)	4'161	99.8 ^{+0.1} _{-0.3}
		1.0	725 (7)	9'798	99.85 ^{+0.07} _{-0.17}
		2.6	840 (9)	11'487	99.96 ^{+0.04} _{-0.12}
180	50	0.05	895 (3)	13'083	99.76 ^{+0.04} _{-0.09}
		0.1	950 (4)	10'167	99.88 ^{+0.04} _{-0.12}
		0.3	790 (5)	9'807	99.92 ^{+0.03} _{-0.10}
		1.0	725 (7)	9'069	99.84 ^{+0.05} _{-0.13}
		2.6	840 (9)	53'973	99.99 ^{+0.01} _{-0.02}

TABLE 6.1: Summary of the 28 datasets acquired. The gain and threshold values in electrons were obtained using data with ⁵⁵Fe and ⁹⁰Sr radioactive sources. The datasets with efficiency smaller than 99 % are highlighted in grey and were not considered for the time resolution measurements.

The left panel of Figure 6.4 shows the efficiency map for the four analog pixels. While the efficiency in the five internal inter-pixel regions remains constant and is larger than 95 %, it degrades in the 15 μm area around the 14 external edges, confirming the influence of the telescope's spatial resolution.

To obtain an unbiased measurement of the efficiency, the sample of telescope tracks was restricted to those that crossed the DUT inside the two triangles obtained by connecting the centers of the four analog pixels (shown by dashed red segments in Figure 6.4). The right panel of Figure 6.4 shows the efficiency map of this unbiased sample, with a color scale ranging from 96 % to 100 %.

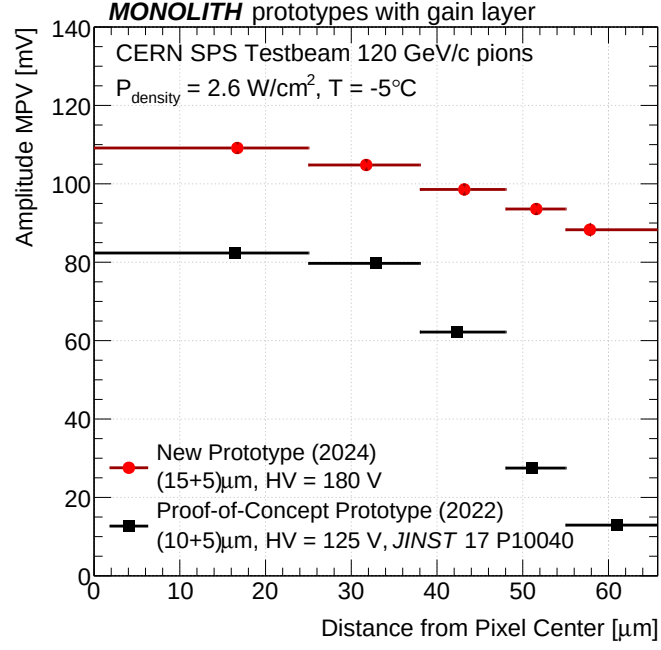


FIGURE 6.3: Comparison of the MPV of the signal amplitude as a function of the distance from the pixel center for the current prototype (red dots) and the proof-of-concept prototype (black squares).

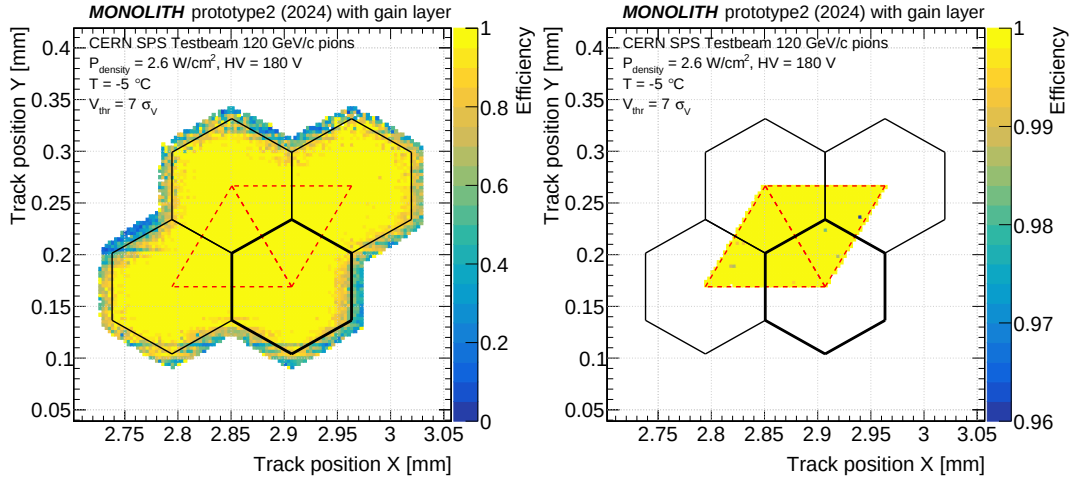


FIGURE 6.4: (Left) Map of the efficiency of the four analog pixels. The effect of the test beam telescope's finite pointing resolution is visible at the external edges of the pixels. (Right) The map of efficiency is restricted to the region defined by the two triangles formed by connecting the centers of the pixels, which is not biased by the telescope resolution. The results refer to the sample taken with $HV = 180 \text{ V}$ and $P_{\text{density}} = 2.6 \text{ W/cm}^2$. Only the bottom-right pixel (highlighted by thicker lines) was connected to the 25 ps sampling oscilloscope and thus used to compute the time resolution.

The dataset shown in Figure 6.4, acquired with the DUT operated at $HV = 180 \text{ V}$ and $P_{\text{density}} = 2.6 \text{ W/cm}^2$, gives an efficiency of $(99.99 \pm 0.01) \%$. The usual Clopper-Pearson frequentist approach [119] with a confidence level of 95 % was used to estimate the uncertainties.

Threshold Scan

Figure 6.5 shows the efficiency and the noise-hit rate as a function of the threshold for the same dataset with the DUT operated at $HV = 180$ V and $P_{\text{density}} = 2.6$ W/cm².

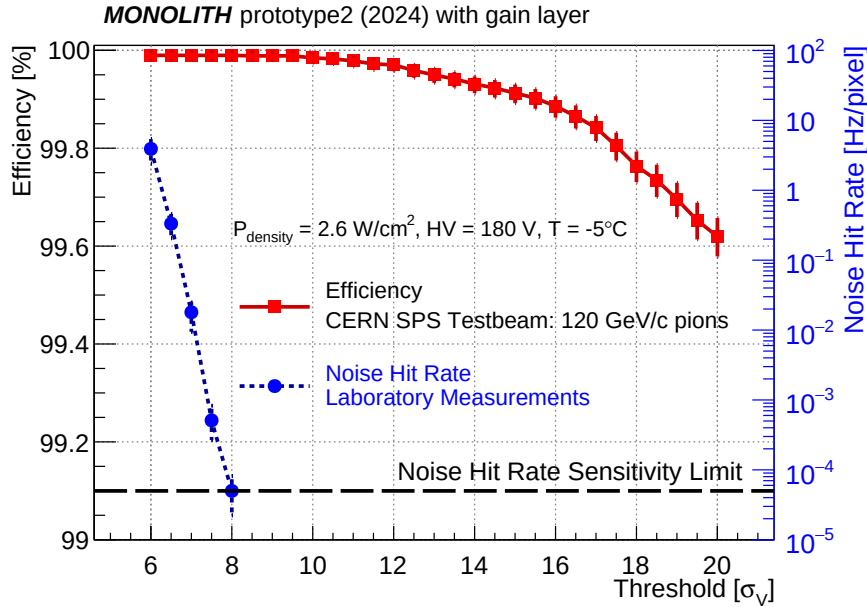


FIGURE 6.5: Efficiency (red squares) and noise hit rate (blue dots) as a function of the threshold for the dataset taken with $HV = 180$ V and $P_{\text{density}} = 2.6$ W/cm². The threshold is given in units of the voltage noise σ_V .

Remarkably, even for a threshold 20 times the voltage noise σ_V , the efficiency remains larger than 99.6 %. The threshold of $7\sigma_V$ utilized throughout this paper is well within the efficiency plateau and corresponds to a noise-hit rate per pixel of 2×10^{-2} Hz. Variation of the threshold by one unit of σ_V produces a change in noise hit rate of approximately two orders of magnitude.

Bias Voltage and Power Density Scans

The datasets' efficiency is reported in Table 6.1 and in Figure 6.6.

As expected, efficiency increases with the power density of the preamplifier analog front-end electronics and with the sensor bias voltage. The former is due to increased front-end gain, and the latter is due to a larger avalanche multiplication in the gain layer. The measurements show that for $P_{\text{density}} \geq 0.1$ W/cm² the detector reaches the plateau of efficiency at $HV = 150$ V. At $P_{\text{density}} = 0.05$ W/cm², a sensor bias voltage of 170 V is needed to obtain an efficiency larger than 99 %.

Efficiency in Different Regions of the Pixel

Figure 6.7 shows the efficiency in bins of the distance from the pixel center for the dataset taken with the highest $P_{\text{density}} = 2.6$ W/cm².

The efficiency is well above 99.9 % and does not depend on the hit position for $HV = 150$ V and $HV = 180$ V. At $HV = 120$ V, the efficiency decreases to 99.5 % near the pixel edge, likely due to charge sharing between adjacent pixels.

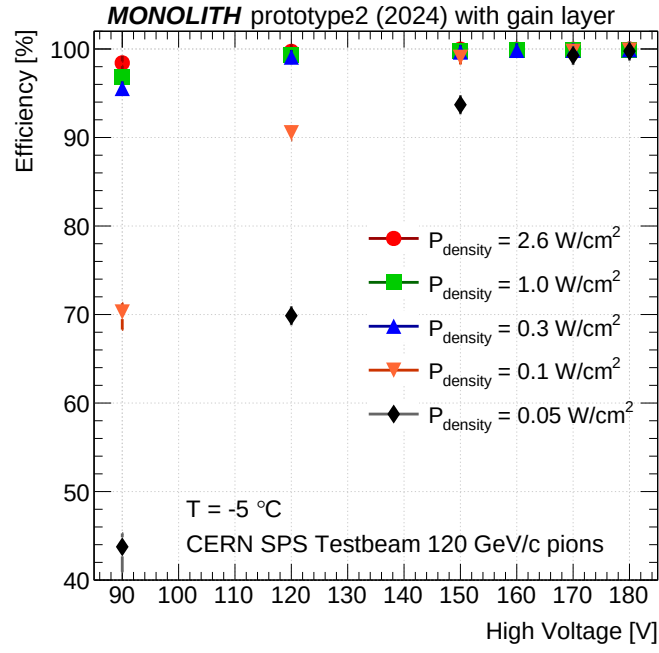


FIGURE 6.6: Efficiency as a function of the bias voltage of the sensor for five values of the power density of the preamplifier. All the datasets reported in Table 6.1 are shown.

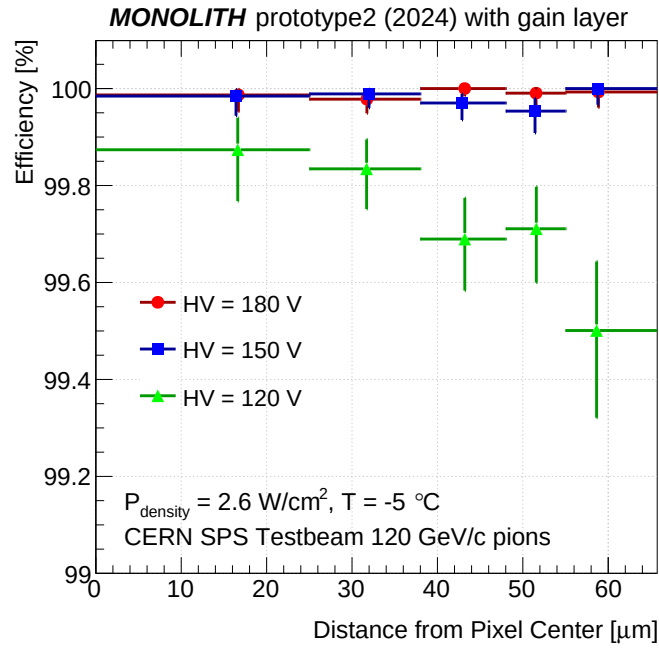


FIGURE 6.7: Efficiency in bins of the distance of the hit position from the pixel center. The data show the results obtained with $P_{\text{density}} = 2.6 \text{ W/cm}^2$ for three values of the sensor bias voltage. In each bin, the data points are plotted at the mean track distance from the center of the pixel.

6.2.5 Time Resolution

The time resolution was measured using the same sample of events as that used to calculate the efficiency in Figure 6.4. Since only the bottom-right analog pixel of the DUT in Figure 6.4 was connected to the oscilloscope sampling at 25 ps, only the subsample of hits in this analog pixel was considered for further analysis, which

corresponds to approximately one-third of the events used for the measurement of the efficiency. The time-resolution measurement was restricted to the datasets in Table 6.1 with efficiency larger than 99 %.

The oscilloscope samplings were interpolated linearly to measure the time of arrival (ToA) of the signals. In the case of the two MCPs, the ToA was taken as the time at which the interpolated signal crossed 50 % of the maximum value of the amplitude, thus correcting most of the time walk. In the case of the DUT, the ToA was taken as the time at which the interpolated signal crossed a fixed voltage threshold. The threshold value V_{th} was set at the integer value in mV closest to seven times the voltage noise σ_V measured in that dataset. As reported in Table 6.1, the threshold ranged from 3 to 9 mV, depending on the front-end electronics working point. Finally, the ToA value was determined for each DUT signal by linearly interpolating between the two oscilloscope samplings with amplitudes just below and above the threshold.

Time-Walk Correction

The signals from the DUT were corrected for time walk by fitting the difference $ToA_{DUT} - ToA_{MCP0}$ within bins of the DUT signal amplitude using Gaussian functional forms. The mean values obtained from these fits were subsequently linearly interpolated to derive the time-walk correction. Figure 6.8 shows the distributions of $ToA_{DUT} - ToA_{MCP0}$ as a function of the signal amplitude of the DUT (left panel) or the MCP0 (right panel).

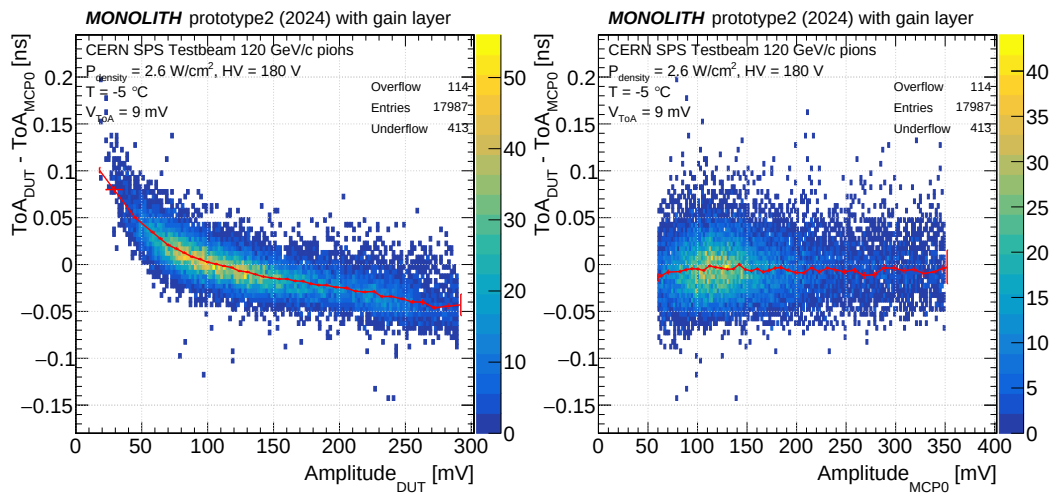


FIGURE 6.8: Difference between the times of arrival of the DUT and the MCP0 as a function of the amplitude of the DUT (left) or the MCP0 (right). The red crosses show the mean values of the ToA difference, as obtained by a Gaussian fit in each vertical slice of the distributions. The red segments are the linear interpolations between the mean values.

The red line in the panel at the left of Figure 6.8 shows the time-walk correction function that was applied to the DUT in the case of the dataset taken with $HV = 180$ V and $P_{density} = 2.6$ W/cm². The panel on the right of Figure 6.8 shows the ToA difference as a function of the amplitude of the MCP0 and confirms that the MCP0 did not need further time-walk correction, since the distribution is flat for all the MCP0 amplitudes.

Methods to Measure the Time Resolution

The dataset with the largest statistics was used to compute the time resolution of the MCP0 using the same method adopted in the previous test beams (see, for example, Equation 4.2). Namely, the time resolution for each of the three ToA differences, constructed from pairs among DUT, MCP0, and MCP1, was measured by fitting the distribution with a Gaussian functional form, and the resulting system of three equations with three unknowns was solved. This method yielded a time resolution for MCP0 of $\sigma_{t,\text{MCP0}} = (5.4 \pm 0.4)$ ps.

Once this value was established, the time resolution of the DUT at each operating working condition was determined by performing a Gaussian fit on the $\text{ToA}_{\text{DUT}} - \text{ToA}_{\text{MCP0}}$ distribution and then subtracting the measured $\sigma_{t,\text{MCP0}}$ in quadrature.

An alternative method was also used, which involved taking the arrival time for the DUT at 10 % of the maximum signal amplitude. This method, which eliminates the need for explicit time-walk correction, yields time resolutions nearly equivalent to those obtained after time-walk correction when using a fixed signal voltage threshold.

The two panels on the left of Figure 6.9 show the $\text{ToA}_{\text{DUT}} - \text{ToA}_{\text{MCP0}}$ distributions before the time-walk correction of the DUT signals for the datasets acquired at $HV = 180$ V with $P_{\text{density}} = 0.3$ W/cm² (top) and 2.6 W/cm² (bottom). The two panels at the center of Figure 6.9 show the time-walk corrected $\text{ToA}_{\text{DUT}} - \text{ToA}_{\text{MCP0}}$ distributions for the same operating condition. Finally, the two right-side panels of Figure 6.9 present the results of the time resolution computed by taking the time at 10 % of the maximum signal amplitude.

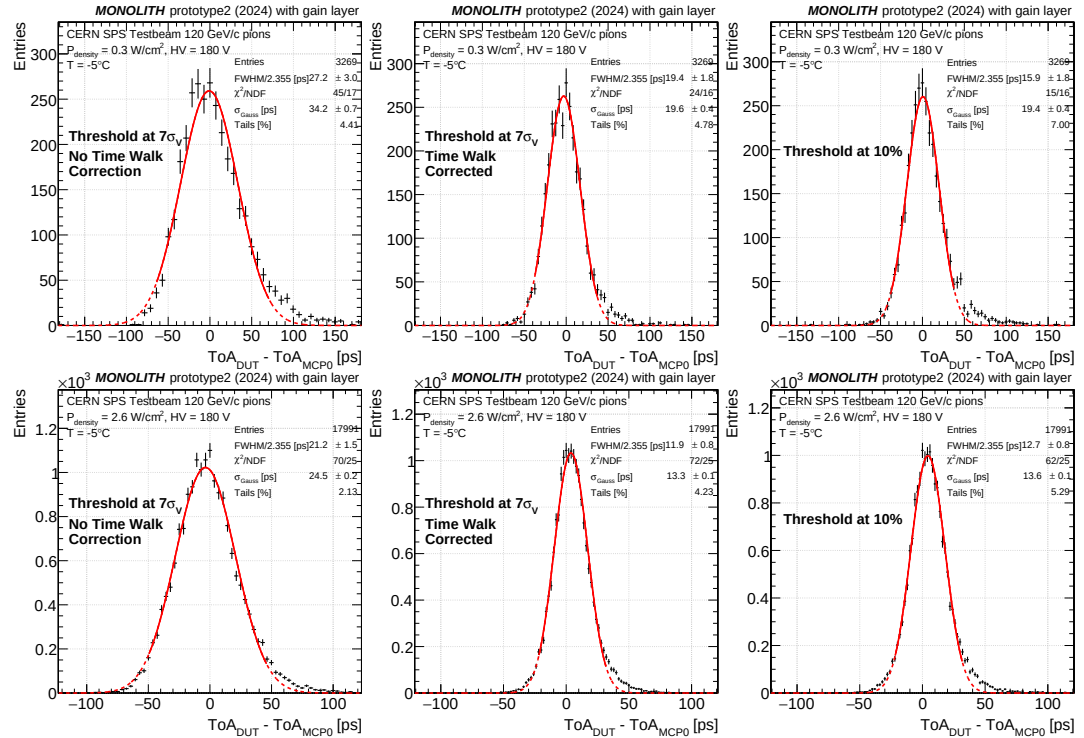


FIGURE 6.9: ToA difference between the DUT and the MCP0 before (left panel) and after (center panel) time-walk correction when the DUT signal time is taken at a threshold $V_{\text{th}} = 7\sigma_V$. The right panels show the ToA difference obtained when the DUT signal time is taken at a threshold $V_{\text{th}} = 10\%$ of the signal height. The two working points shown were taken at $HV = 180$ V and front-end electronics power density of 0.3 (top) and 2.6 W/cm² (bottom). The distributions were fitted with a Gaussian function up to 20 % of the maximum.

It should be noted that the $\text{ToA}_{\text{DUT}} - \text{ToA}_{\text{MCP0}}$ distributions present non-Gaussian tails at large positive values. As explained in Section 5.5, we verified that these delayed DUT signal times are caused by fluctuations in primary charge, which produce large clusters in the lower region of the absorption layer, positioned far from the collection electrode.

To account for the presence of these tails and evaluate their percentage, the Gaussian fits were performed up to 20 % of the maximum height of the distribution. The fraction of non-Gaussian tails was then computed from the ratio between the total number of entries of the histogram and the integral of the Gaussian function resulting from the fit. In all datasets analyzed, the tails were found to vary between 2 % and 10 %. Therefore, the time resolutions quoted here refer to 90 % or more of the recorded signals.

Table 6.2 reports the measured time resolutions for all datasets with efficiency larger than 99 %, together with the working condition of bias voltage HV and power density P_{density} , and the electron gain measured with radioactive sources in the laboratory. All three methods are reported (no time-walk correction and $V_{\text{th}} = 7\sigma_V$, time-walk correction and $V_{\text{th}} = 7\sigma_V$, no time-walk correction and $V_{\text{th}} = 10\%$).

HV [V]	Electron Gain	P_{density} [W/cm ²]	σ_T (no TWC) [ps]	σ_t (TWC) [ps]	σ_t (10 %) [ps]
120	10	0.3	78.8 ± 1.7	32.5 ± 0.7	34.3 ± 0.7
		1.0	50.9 ± 0.5	24.8 ± 0.8	26.8 ± 0.9
		2.6	42.0 ± 0.5	17.8 ± 0.3	19.4 ± 0.2
150	18	0.1	98.1 ± 2.1	44.3 ± 0.9	45.9 ± 0.8
		0.3	55.9 ± 0.7	22.3 ± 0.3	23.4 ± 0.3
		1.0	36.9 ± 0.8	18.8 ± 0.5	20.0 ± 0.4
		2.6	32.5 ± 0.3	14.1 ± 0.2	14.8 ± 0.3
160	21	0.3	46.8 ± 1.0	21.3 ± 1.0	22.6 ± 0.4
		1.0	35.1 ± 0.8	18.1 ± 0.5	18.4 ± 0.4
		2.6	29.5 ± 0.7	13.7 ± 0.6	14.8 ± 0.6
170	33	0.05	88.6 ± 1.6	48.2 ± 1.1	47.1 ± 0.7
		0.1	65.2 ± 1.3	34.4 ± 1.0	33.7 ± 0.6
		0.3	38.1 ± 1.2	20.4 ± 0.7	20.4 ± 0.7
		1.0	29.9 ± 0.5	17.0 ± 0.6	17.4 ± 0.5
		2.6	24.9 ± 0.6	13.2 ± 0.4	13.4 ± 0.4
180	50	0.05	73.8 ± 1.4	41.6 ± 1.0	40.9 ± 0.7
		0.1	52.4 ± 1.2	28.4 ± 0.8	28.2 ± 0.6
		0.3	33.2 ± 0.7	18.8 ± 0.4	18.6 ± 0.4
		1.0	26.9 ± 0.6	15.3 ± 0.4	15.8 ± 0.4
		2.6	24.3 ± 0.2	12.1 ± 0.2	12.5 ± 0.2

TABLE 6.2: Electron gain and time resolution for the acquired datasets with an efficiency above 99 %. The time resolution was measured using several methods: fixed threshold without time-walk correction, fixed threshold with time-walk correction (used as the reference for the summary plots), and threshold at a constant fraction of the amplitude at 10 %.

Comparison of the time resolution measurements in Table 6.2, obtained before and after time-walk correction, indicates that the correction significantly enhances the time resolution in all cases by approximately a factor of two. On the other hand,

the results obtained without time-walk correction and with a threshold at a constant fraction of the amplitude at 10 %, are substantially compatible with the results with time-walk correction and fixed threshold, despite being systematically worse of about 5-10 %.

Bias Voltage and Power Density Scans

The time resolution results for all the acquired datasets with efficiency larger than 99 % reported in Table 6.2 are also summarized in Figure 6.10.

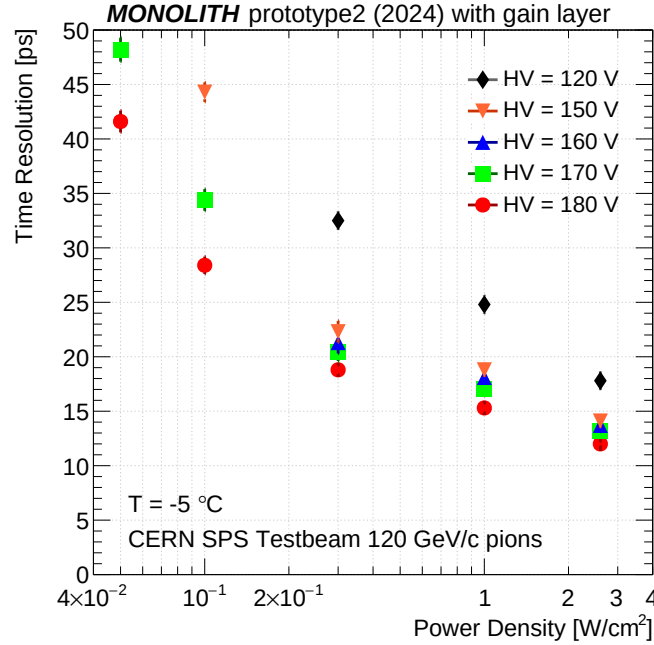


FIGURE 6.10: Time resolution as a function of the power density supplied to the preamplifier for five sensor bias voltages. Only the datasets in Table 6.1 with operating conditions that provided efficiency larger than 99 % were considered for the measurement of the time resolution.

A time resolution better than 50 ps was measured in all cases. Moreover, as expected, the time resolution improves with increasing preamplifier power density and sensor bias voltage.

At sensor bias voltage $HV = 180$ V (red dots in Figure 6.10), the monolithic ASIC prototype with PicoAD sensor provides outstanding time resolution at low and high power density, namely $\sigma_{t,DUT} = (41.6 \pm 1.0)$ ps at $P_{\text{density}} = 0.05$ W/cm², (18.8 ± 0.4) ps at 0.3 W/cm², to reach (12.1 ± 0.3) ps at 2.6 W/cm².

Time Resolution in Different Regions of the Pixel

The time resolution was analyzed in bins based on the distance of the hit position from the pixel center. The left panel in Figure 6.11 shows the results obtained using five distance bins, defined to contain similar statistics, for the three acquired sensor bias voltage values at the highest $P_{\text{density}} = 2.6$ W/cm².

The left panel in Figure 6.11 shows a worsening of the time resolution from the center to the edge of the pixel of approximately 3 ps for $HV = 180$ V and 5 ps for $HV = 120$ V, corresponding to approximately 30 %.

The right panel of Figure 6.11 presents a comparison between the results of the current prototype and those of the PicoAD proof-of-concept prototype (see Section 4.3.5).

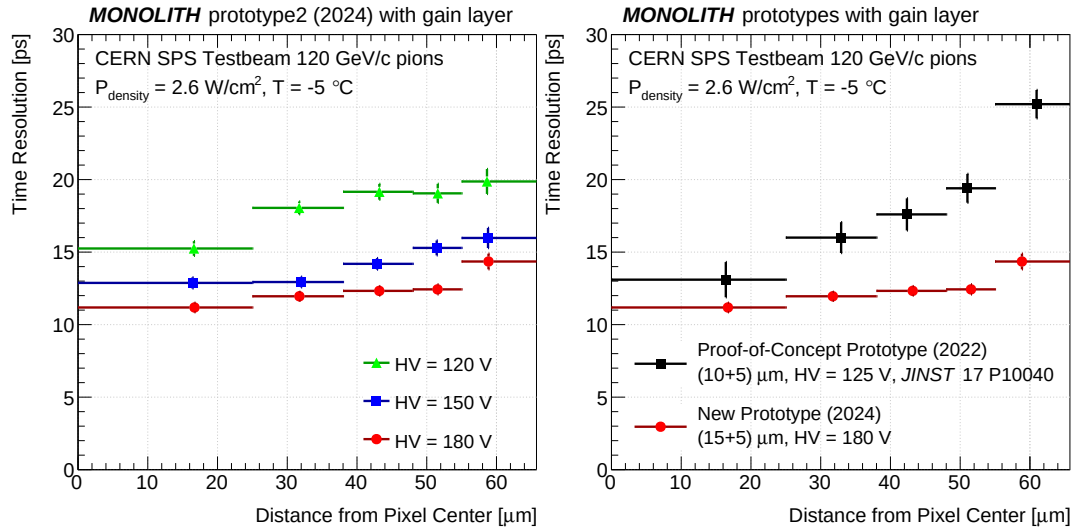


FIGURE 6.11: Time resolution as a function of the distance from the pixel center. (Left) Scan for the three different sensor bias voltages. (Right) Comparison between this prototype and the proof-of-concept prototype. In each bin, the data points are plotted at the mean track distance from the center of the pixel. The values from the proof-of-concept prototype are the ones reported in Section 4.3.5.

In the proof-of-concept prototype, the average time resolution was 17 ps, with values ranging from 13 ps at the pixel center to 25 ps at the interpixel region, indicating a twofold worsening from center to edge. The reduced variation in time resolution as a function of the hit position in the current prototype is attributed to the decreased dependence of amplitude on the hit position, as shown in Figure 6.3.

The PicoAD proof-of-concept prototype achieved a time resolution of (17.3 ± 0.4) ps using a signal-processing technique referred to as *CFD-like* in Section 4.3.5. This method involved shifting the signal by 200 ps (half of its rise time), subtracting the shifted signal from the original, and determining the ToA at 25 % of the resulting signal amplitude. This signal-processing approach was implemented to mitigate low-frequency noise components in the prototype, which originated from feedback injection through the power supply lines. Its adoption led to a significant improvement in time resolution compared to the (24.1 ± 0.4) ps achieved using an amplitude-based analysis similar to the method employed in this paper. In this updated prototype, where a redesigned front-end electronics system effectively resolved the feedback injection issue responsible for the low-frequency noise, the same method yielded only a marginal improvement in time resolution of approximately 10 % and was discontinued.

Part III

Conclusions

Chapter 7

Conclusions

This thesis has presented the development, characterization, and validation of successive monolithic silicon detector prototypes for the MONOLITH project. It has also demonstrated the feasibility and performance potential of front-end electronics read-out implemented in SiGe BiCMOS technology and the PicoAD sensor concept. The results of extensive test beam campaigns and complementary simulations confirm that sub-20 picoseconds timing resolution and near-perfect detection efficiency are achievable across a wide range of power densities and bias voltages.

PicoAD Proof-of-Concept Prototype

A monolithic proof-of-concept prototype for the MONOLITH project, featuring the PicoAD sensor concept, was produced in 130 nm SiGe BiCMOS technology with hexagonal pixels and a pitch of 100 μm . The prototype, characterized by a fully-depleted double-junction $(n-p)_{\text{pixel}}-(n-p)_{\text{gain}}$ structure, was tested both in the clean rooms laboratory at the University of Geneva and at the CERN SPS facility in 2021 with a beam of pions with 180 GeV/c momentum.

The laboratory measurements confirm the results of the TCAD simulations on the impact-ionization gain, showing that the proof-of-concept prototype works as expected. Some probe station current-voltage characteristics reveal insights into the working principle of the gain layer. A measured gain of 20 is achieved for electron-initiated avalanches using a ^{55}Fe radioactive source.

The test beam measurements show that, at a sensor bias voltage $HV = 125\text{ V}$, the measured detection efficiency is compatible with 99.9 % when the power density is varied between 0.4 and 2.7 W/cm^2 . In this prototype, the time resolution depends significantly on the distance from the center of the pixel: at a power density of 2.7 W/cm^2 and a bias voltage of 125 V, the time resolution varies between 13 ps at the center of the pixel and 25 ps at the edge of the pixel, while in the full pixel area, including the inter-pixel regions, it is $(17.3 \pm 0.4)\text{ ps}$. At a reduced power consumption of 0.9 and 0.4 W/cm^2 , the overall-pixel time resolution is $(25.1 \pm 0.7)\text{ ps}$ and $(30.2 \pm 0.3)\text{ ps}$, respectively.

These results show that the PicoAD sensor concept works and that the addition of a continuous gain layer in a second $p-n$ junction far from the pixel matrix provides a fully efficient monolithic sensor at an affordable power consumption that can improve significantly the already remarkable performance obtained with a SiGe BiCMOS front-end in a monolithic sensor without a gain layer [58].

Second MONOLITH Prototype without a Gain Layer

A second prototype for the MONOLITH project, this time without a gain layer, was produced in 2022 with the SG13G2 SiGe BiCMOS IHP process. This prototype was

tested in the laboratory at the University of Geneva, yielding a high amplifier gain of up to (60 ± 2) mV/fC, an equivalent noise charge of (124 ± 5) electrons, and a fast rise time of (375 ± 4) ps.

The ASIC was tested with a beam of 120 GeV/c pions at the CERN SPS at several sensor bias voltages and front-end power consumption density. The detection efficiency is consistent with 99.8 % for all working points analyzed using a voltage threshold seven times the voltage noise. This efficiency does not depend on the position of the hit within the pixel surface, demonstrating considerable efficiency also in the inter-pixel regions. The time resolution obtained by computing the signal time of arrival at the voltage threshold by a simple linear interpolation of the acquired signal samplings is 20 ps at bias voltage of 200 V and front-end power density of 2.7 W/cm^2 . The prototype can be operated with a time resolution of 30 ps at 0.36 W/cm^2 and 80 ps at 0.04 W/cm^2 . The measured time resolutions have little dependency on the position of the hit within the pixel area, with a slight degradation towards the edge of the pixel.

The time resolution was also studied using pulses from a femtosecond laser. The prototype was operated at a sensor bias voltage of 200 V and a front-end power density ranging from 0.36 to 2.7 W/cm^2 . At the highest power density, the time resolution was found to vary between 45 ps and 3 ps for laser pulses producing amplitudes between 20 mV and 180 mV, corresponding approximately to 0.4 and 3.5 times the most probable value of the charge deposited by a minimum-ionizing particle, respectively. Comparison with the time resolution obtained at the same working point using data taken in a test beam with minimum-ionizing particles shows that the charge-collection noise contribution to the time resolution increases with the signal amplitude and dominates the time resolution at high signal amplitudes.

Samples of these second MONOLITH prototypes were irradiated at the CYRIC facility in Japan with 70 MeV protons up to a maximum fluence of $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$. The ASICs could be operated after bypassing the radiation-damaged electronic components of the read-out boards. Data were taken with a ^{90}Sr radioactive source at a temperature of -35°C , a sensor bias voltage of 200 V, and a power density of 0.9 W/cm^2 to characterize the irradiated ASICs together with three not irradiated ASICs used for reference.

The time jitter of each signal was calculated as the ratio between the voltage noise σ_V and the signal slope measured at the $7\sigma_V$ threshold, and the timing performance of each pixel was estimated by the time jitter of the most probable signal amplitude. The results obtained with this method for the ASICs that were not irradiated reproduced the timing performance at the 20 ps level measured at the test beam. No degradation in timing performance was observed up to a proton fluence of $6 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$. At $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$ the time jitter of the most probable signal amplitude was found to be $\sigma_t^{^{90}\text{Sr}} = (56.6 \pm 16.3)$ ps. A new working point specific to this proton fluence, which includes larger sensor bias voltage and larger analog voltage supplied to the preamplifier, brings the time jitter to $\sigma_t^{^{90}\text{Sr}} = (39.6 \pm 6.6)$ ps.

The same ASICs were tested at the CERN SPS using a beam of 120 GeV/c pions. Before proton irradiation, the detection efficiency at a sensor bias voltage of 200 V and front-end electronics power density of 0.9 W/cm^2 was measured to be $99.96^{+0.01}_{-0.02}$ %. For a proton fluence $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$, the detection efficiency in the same conditions was measured to be (96.2 ± 0.3) %; an increase of the bias voltage to 300 V provided $99.69^{+0.06}_{-0.07}$ %. These efficiency values also include the inter-pixel regions. The time of arrival of the signals used in the data analysis was defined as the time at which the signal amplitude reached 50 % of its maximum value, and it was calculated using

linear interpolation of the oscilloscope samples. No further time-walk correction was applied to the data. The time resolution at power density of 0.9 W/cm^2 and bias voltage of 200 V is $(20.0 \pm 1.0) \text{ ps}$ before proton irradiation. At the highest proton fluence studied of $1 \times 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$, the time resolution became $(53.1 \pm 3.4) \text{ ps}$. Increasing the sensor bias voltage to 300 V improved the result to $(45.3 \pm 1.6) \text{ ps}$.

These results demonstrate that SiGe BiCMOS processes can be utilized for producing high-time-resolution pixelated silicon detectors without an internal gain layer, making them suitable for future colliders and other disciplines that involve high-radiation environments.

Finally, a fully working simulation within the Garfield++ framework was set up, and it is now available for further study of improvements to the SiGe BiCMOS technology used in the MONOLITH project. The simulation successfully reproduced the results obtained from the test beam experimental data. The simulation also disentangles the time resolution into its primary contributions: charge-collection fluctuations (Landau noise) and electronics time jitter. We isolated each component by selectively disabling ENC in Garfield++ and employing uniform versus MIP-like charge depositions. We uncovered non-Gaussian tails in the time-of-arrival distributions at the few-percent level. A systematic simulation study revealed that these tails originate from occasional large charge clusters formed at a specific depth of the sensor.

Second MONOLITH Prototype with a Gain Layer

A third prototype, with the same ASIC design as the second one but featuring a gain layer similar to the PicoAD proof-of-concept prototype, was produced in 2024. The detection efficiency and time resolution were measured using a $120 \text{ GeV}/c$ pion beam at a CERN SPS test beam.

Data were taken at sensor bias voltages between 90 and 180 V and a front-end power density between 0.05 and 2.6 W/cm^2 . A fast oscilloscope was used to read the analog channels in the pixel matrix and record the data. At a sensor bias voltage of 180 V, which is 20 V below the breakdown, and power density of 0.3 W/cm^2 , the measured detection efficiency is $99.92^{+0.03}_{-0.10} \%$. When measuring the signal time at a fixed threshold of seven times the voltage noise, the time resolution is $(33.2 \pm 0.7) \text{ ps}$ before time-walk correction and improves to $(18.8 \pm 0.4) \text{ ps}$ after correction. Using the signal time at 10 % of the maximum amplitude, the time resolution remains at $(18.6 \pm 0.4) \text{ ps}$ without the need for further time-walk correction. At the same sensor bias voltage and the highest front-end power used of 2.6 W/cm^2 , the detection efficiency is $99.99^{+0.01}_{-0.02} \%$ and the time resolution after time-walk correction is $(12.1 \pm 0.3) \text{ ps}$.

These results demonstrate that monolithic ASICs with PicoAD sensor and SiGe HBT front-end electronics are outstanding candidates for a broad range of applications with power requirements spanning two orders of magnitude. They are well-suited for large-area applications, such as particle physics 4D trackers, where power density consumption must stay well below 1 W/cm^2 . The PicoAD's full-fill factor and ultra-thin design suit applications requiring a fully active, low-material-budget timing detector. It is also ideal for timing applications with fewer pixels, like direct-TOF LiDAR systems, where a power density of a few W/cm^2 is feasible. With its excellent noise performance, radiation tolerance, and compatibility with standard CMOS processing, SiGe HBT is a mature, scalable technology for the next generation of fast-timing layers at future high-luminosity colliders.

Bibliography

- [1] L. Paolozzi et al. “Picosecond Avalanche Detector — Working Principle and Gain Measurement with a Proof-of-concept Prototype”. In: *Journal of Instrumentation* 17.10 (Oct. 2022), P10032. DOI: [10.1088/1748-0221/17/10/P10032](https://doi.org/10.1088/1748-0221/17/10/P10032).
- [2] G. Iacobucci et al. “Testbeam Results of the Picosecond Avalanche Detector Proof-of-concept Prototype”. In: *Journal of Instrumentation* 17.10 (Oct. 2022), P10040. DOI: [10.1088/1748-0221/17/10/P10040](https://doi.org/10.1088/1748-0221/17/10/P10040).
- [3] M. Milanesio et al. “Gain Measurements of the First Proof-of-concept PicoAD Prototype with a ^{55}Fe X-ray Radioactive Source”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1046 (Jan. 2023), p. 167807. ISSN: 0168-9002. DOI: [10.1016/j.nima.2022.167807](https://doi.org/10.1016/j.nima.2022.167807).
- [4] S. Zambito et al. “20 ps Time Resolution with a Fully Efficient Monolithic Silicon Pixel Detector without Internal Gain Layer”. In: *Journal of Instrumentation* 18.03 (Mar. 2023), P03047. DOI: [10.1088/1748-0221/18/03/P03047](https://doi.org/10.1088/1748-0221/18/03/P03047).
- [5] M. Milanesio et al. “Radiation Tolerance of SiGe BiCMOS Monolithic Silicon Pixel Detectors without Internal Gain Layer”. In: *Journal of Instrumentation* 19.01 (Jan. 2024), P01014. DOI: [10.1088/1748-0221/19/01/P01014](https://doi.org/10.1088/1748-0221/19/01/P01014).
- [6] M. Milanesio et al. “Time Resolution of a SiGe BiCMOS Monolithic Silicon Pixel Detector without Internal Gain Layer with a Femtosecond Laser”. In: *Journal of Instrumentation* 19.04 (Apr. 2024), P04029. DOI: [10.1088/1748-0221/19/04/P04029](https://doi.org/10.1088/1748-0221/19/04/P04029).
- [7] T. Moretti et al. “Testbeam Results of Irradiated SiGe BiCMOS Monolithic Silicon Pixel Detector without Internal Gain Layer”. In: *Journal of Instrumentation* 19.07 (July 2024), P07036. DOI: [10.1088/1748-0221/19/07/P07036](https://doi.org/10.1088/1748-0221/19/07/P07036).
- [8] L. Paolozzi et al. “Testbeam Characterization of a SiGe BiCMOS Monolithic Silicon Pixel Detector with Internal Gain Layer”. In: *Journal of Instrumentation* 20.04 (Apr. 2025), P04001. DOI: [10.1088/1748-0221/20/04/P04001](https://doi.org/10.1088/1748-0221/20/04/P04001).
- [9] D. J. Griffiths. *Introduction to Elementary Particles*. 2nd, Revised. Weinheim: Wiley-VCH, 2008. ISBN: 978-3-527-40601-2. DOI: [10.1002/9783527618460](https://doi.org/10.1002/9783527618460).
- [10] S. Navas et al. “Review of particle physics”. In: *Phys. Rev. D* 110.3 (Aug. 2024), p. 030001. DOI: [10.1103/PhysRevD.110.030001](https://doi.org/10.1103/PhysRevD.110.030001).
- [11] F. Hartmann. *Evolution of Silicon Sensor Technology in Particle Physics*. Vol. 275. Springer Tracts in Modern Physics. Springer, 2017. ISBN: 978-3-319-64434-9, 978-3-319-64436-3. DOI: [10.1007/978-3-319-64436-3](https://doi.org/10.1007/978-3-319-64436-3).
- [12] Enrico Fermi. “On the Quantization of the Monoatomic Ideal Gas”. In: *arXiv Condensed Matter* (Dec. 1999). English Translation of the Original 1926 Paper. arXiv: [cond-mat/9912229](https://arxiv.org/abs/cond-mat/9912229).
- [13] S. N. Bose. “Plancks Gesetz und Lichtquantenhypothese”. In: *Z. für Physik* 26 (Dec. 1924), pp. 178–181. DOI: [10.1007/BF01327326](https://doi.org/10.1007/BF01327326).

- [14] S.-I. Tomonaga. "On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields". In: *Progress of Theoretical Physics* 1.2 (Aug. 1946), pp. 27–42. DOI: [10.1143/PTP.1.27](https://doi.org/10.1143/PTP.1.27).
- [15] J. Schwinger. "Quantum Electrodynamics. I. A Covariant Formulation". In: *Physical Review* 74.10 (Nov. 1948), pp. 1439–1461. DOI: [10.1103/PhysRev.74.1439](https://doi.org/10.1103/PhysRev.74.1439).
- [16] R. P. Feynman. "Space-Time Approach to Quantum Electrodynamics". In: *Physical Review* 76.6 (Sept. 1949), pp. 769–789. DOI: [10.1103/PhysRev.76.769](https://doi.org/10.1103/PhysRev.76.769).
- [17] E. Fermi. "Tentativo di una teoria dei raggi β ". In: *Il Nuovo Cimento* 9 (Dec. 1934), pp. 1–19. DOI: [10.1007/BF02959820](https://doi.org/10.1007/BF02959820).
- [18] C. D. Anderson and S. H. Neddermeyer. "Cloud Chamber Observations of Cosmic Rays at 4300 Meters Elevation and Near Sea-Level". In: *Phys. Rev.* 50 (Aug. 1936), pp. 263–271. DOI: [10.1103/PhysRev.50.263](https://doi.org/10.1103/PhysRev.50.263).
- [19] F. Reines and C. L. Cowan Jr. "Detection of the Free Neutrino". In: *Physical Review* 92 (Nov. 1953), pp. 830–831. DOI: [10.1103/PhysRev.92.830](https://doi.org/10.1103/PhysRev.92.830).
- [20] G. Danby et al. "Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos". In: *Physical Review Letters* 9 (July 1962), pp. 36–44. DOI: [10.1103/PhysRevLett.9.36](https://doi.org/10.1103/PhysRevLett.9.36).
- [21] S. L. Glashow. "Partial-Symmetries of Weak Interactions". In: *Nuclear Physics* 22.4 (Feb. 1961), pp. 579–588. DOI: [10.1016/0029-5582\(61\)90469-2](https://doi.org/10.1016/0029-5582(61)90469-2).
- [22] S. Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (Nov. 1967), pp. 1264–1266. DOI: [10.1103/PhysRevLett.19.1264](https://doi.org/10.1103/PhysRevLett.19.1264).
- [23] A. Salam. "Weak and Electromagnetic Interactions". In: *Conf. Proc. C* 680519 (May 1968), pp. 367–377. DOI: [10.1142/9789812795915_0034](https://doi.org/10.1142/9789812795915_0034).
- [24] F.J. Hasert et al. "Search for Elastic Muon-Neutrino Electron Scattering". In: *Physics Letters B* 46.1 (Sept. 1973). Gargamelle Collaboration, pp. 121–124. DOI: [10.1016/0370-2693\(73\)90494-2](https://doi.org/10.1016/0370-2693(73)90494-2).
- [25] F.J. Hasert et al. "Observation of Neutrino-Like Interactions without Muon or Electron in the Gargamelle Neutrino Experiment". In: *Physics Letters B* 46.1 (Sept. 1973). Gargamelle Collaboration, pp. 138–140. DOI: [10.1016/0370-2693\(73\)90499-1](https://doi.org/10.1016/0370-2693(73)90499-1).
- [26] G. 't Hooft and M. Veltman. "Regularization and Renormalization of Gauge Fields". In: *Nuclear Physics B* 44.1 (July 1972), pp. 189–213. DOI: [10.1016/0550-3213\(72\)90279-9](https://doi.org/10.1016/0550-3213(72)90279-9).
- [27] S. L. Glashow, J. Iliopoulos, and L. Maiani. "Weak Interactions with Lepton-Hadron Symmetry". In: *Physical Review D* 2 (Oct. 1970), pp. 1285–1292. DOI: [10.1103/PhysRevD.2.1285](https://doi.org/10.1103/PhysRevD.2.1285).
- [28] J. J. Aubert et al. "Experimental Observation of a Heavy Particle J ". In: *Phys. Rev. Lett.* 33 (Dec. 1974), pp. 1404–1406. DOI: [10.1103/PhysRevLett.33.1404](https://doi.org/10.1103/PhysRevLett.33.1404).
- [29] J. E. Augustin et al. "Discovery of a Narrow Resonance in e^+e^- Annihilation". In: *Phys. Rev. Lett.* 33 (Dec. 1974), pp. 1406–1408. DOI: [10.1103/PhysRevLett.33.1406](https://doi.org/10.1103/PhysRevLett.33.1406).
- [30] J. H. Christenson et al. "Evidence for the 2π Decay of the K_2^0 Meson". In: *Physical Review Letters* 13.4 (July 1964), pp. 138–140. DOI: [10.1103/PhysRevLett.13.138](https://doi.org/10.1103/PhysRevLett.13.138).

- [31] M. Kobayashi and T. Maskawa. “CP-Violation in the Renormalizable Theory of Weak Interaction”. In: *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657. DOI: [10.1143/PTP.49.652](#).
- [32] S.W. Herb et al. “Observation of a Dimuon Resonance at 9.5 GeV in 400 GeV Proton - Nucleus Collisions”. In: *Phys. Rev. Lett.* 39 (Aug. 1977), pp. 252–255. DOI: [10.1103/PhysRevLett.39.252](#).
- [33] F. Abe et al. “Observation of Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab”. In: *Physical Review Letters* 74.14 (Apr. 1995). CDF Collaboration, pp. 2626–2631. DOI: [10.1103/PhysRevLett.74.2626](#).
- [34] S. Abachi et al. “Observation of the Top Quark”. In: *Physical Review Letters* 74.14 (Apr. 1995). D0 Collaboration, pp. 2632–2637. DOI: [10.1103/PhysRevLett.74.2632](#).
- [35] D. J. Gross and F. Wilczek. “Ultraviolet Behavior of Non-Abelian Gauge Theories”. In: *Phys. Rev. Lett.* 30 (June 1973), pp. 1343–1346. DOI: [10.1103/PhysRevLett.30.1343](#).
- [36] H. D. Politzer. “Reliable Perturbative Results for Strong Interactions?” In: *Phys. Rev. Lett.* 30 (June 1973), pp. 1346–1349. DOI: [10.1103/PhysRevLett.30.1346](#).
- [37] G. Arnison et al. “Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV”. In: *Phys. Lett. B* 122 (Feb. 1983). UA1 Collaboration, pp. 103–116. DOI: [10.1016/0370-2693\(83\)91177-2](#).
- [38] M. Banner et al. “Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN $p\bar{p}$ Collider”. In: *Phys. Lett. B* 122 (Mar. 1983). UA2 Collaboration, pp. 476–485. DOI: [10.1016/0370-2693\(83\)91605-2](#).
- [39] G. Arnison et al. “Experimental observation of lepton pairs of invariant mass around 95 GeV/c² at the CERN SPS collider”. In: *Physics Letters B* 126.5 (July 1983). UA1 Collaboration, pp. 398–410. DOI: [10.1016/0370-2693\(83\)90188-0](#).
- [40] P. Bagnaia et al. “Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN $p\bar{p}$ collider”. In: *Physics Letters B* 129.1 (Sept. 1983). UA2 Collaboration, pp. 130–140. DOI: [10.1016/0370-2693\(83\)90744-X](#).
- [41] M. L. Perl et al. “Evidence for Anomalous Lepton Production in e^+e^- Annihilation”. In: *Physical Review Letters* 35.22 (Dec. 1975), pp. 1489–1492. DOI: [10.1103/PhysRevLett.35.1489](#).
- [42] K. Kodama et al. “Observation of Tau Neutrino Interactions”. In: *Physics Letters B* 504.3 (Apr. 2001), pp. 218–224. DOI: [10.1016/S0370-2693\(01\)00307-0](#).
- [43] P. W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13 (Oct. 1964), pp. 508–509. DOI: [10.1103/PhysRevLett.13.508](#).
- [44] G. Aad et al. “Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS detector at the LHC”. In: *Phys. Lett. B* 716 (Sept. 2012). The ATLAS Collaboration, pp. 1–29. DOI: [10.1016/j.physletb.2012.08.020](#).
- [45] S. Chatrchyan et al. “Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC”. In: *Phys. Lett. B* 716 (Sept. 2012). The CMS Collaboration, pp. 30–61. DOI: [10.1016/j.physletb.2012.08.021](#).

- [46] V. C. Rubin and W. K. Ford Jr. "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions". In: *Astrophysical Journal* 159 (Feb. 1970), pp. 379–403. DOI: [10.1086/150317](https://doi.org/10.1086/150317).
- [47] D. Clowe et al. "A Direct Empirical Proof of the Existence of Dark Matter". In: *Astrophysical Journal Letters* 648 (Aug. 2006), pp. L109–L113. DOI: [10.1086/508162](https://doi.org/10.1086/508162).
- [48] The Planck Collaboration et al. "Planck 2018 Results. VI. Cosmological parameters". In: *Astronomy and Astrophysics* 641 (Sept. 2020), A6. DOI: [10.1051/0004-6361/201833910](https://doi.org/10.1051/0004-6361/201833910).
- [49] G. F. Smoot et al. "Structure in the COBE Differential Microwave Radiometer First-Year Maps". In: *Astrophysical Journal Letters* 396 (Sept. 1992), p. L1. DOI: [10.1086/186504](https://doi.org/10.1086/186504).
- [50] A. G. Riess et al. "Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant". In: *Astron. J.* 116 (Sept. 1998), pp. 1009–1038. DOI: [10.1086/300499](https://doi.org/10.1086/300499).
- [51] A. D. Sakharov. "Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe". In: *Pisma Zh. Eksp. Teor. Fiz.* 5 (Dec. 1967), pp. 32–35. DOI: [10.1070/PU1991v034n05ABEH002497](https://doi.org/10.1070/PU1991v034n05ABEH002497).
- [52] G. F. Giudice. "Naturalness after LHC-7 and LHC-8". In: *PoS EPS-HEP2013* (Mar. 2013), p. 163. DOI: [10.22323/1.180.0163](https://doi.org/10.22323/1.180.0163).
- [53] Y. Fukuda et al. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Physical Review Letters* 81.8 (Aug. 1998), pp. 1562–1567. DOI: [10.1103/PhysRevLett.81.1562](https://doi.org/10.1103/PhysRevLett.81.1562).
- [54] M. Benoit et al. "100 ps Time Resolution with Thin Silicon Pixel Detectors and a SiGe HBT Amplifier". In: *Journal of Instrumentation* 11.03 (Mar. 2016), P03011. DOI: [10.1088/1748-0221/11/03/P03011](https://doi.org/10.1088/1748-0221/11/03/P03011).
- [55] P. Valerio et al. "A Monolithic ASIC Demonstrator for the Thin Time-of-Flight PET Scanner". In: *Journal of Instrumentation* 14.07 (July 2019), P07013. DOI: [10.1088/1748-0221/14/07/P07013](https://doi.org/10.1088/1748-0221/14/07/P07013).
- [56] L. Paolozzi et al. "Time Resolution and Power Consumption of a Monolithic Silicon Pixel Prototype in SiGe BiCMOS Technology". In: *Journal of Instrumentation* 15.11 (Nov. 2020), P11025. DOI: [10.1088/1748-0221/15/11/P11025](https://doi.org/10.1088/1748-0221/15/11/P11025).
- [57] H. Rücker, B. Heinemann, and A. Fox. "Half-Terahertz SiGe BiCMOS technology". In: *2012 IEEE 12th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems* (Mar. 2012), pp. 133–136. DOI: [10.1109/SiRF.2012.6160164](https://doi.org/10.1109/SiRF.2012.6160164).
- [58] G. Iacobucci et al. "Efficiency and Time Resolution of Monolithic Silicon Pixel Detectors in SiGe BiCMOS Technology". In: *Journal of Instrumentation* 17.02 (Feb. 2022), P02019. DOI: [10.1088/1748-0221/17/02/P02019](https://doi.org/10.1088/1748-0221/17/02/P02019).
- [59] G. Iacobucci. *MONOLITH - Monolithic Multi-Junction Picosecond Avalanche Detector for Future Physics Experiments and Applications*. H2020 ERC Advanced Project. 2020. URL: <https://cordis.europa.eu/project/id/884447>.
- [60] G. Iacobucci L. Paolozzi and P. Valerio. "Multi-Junction Pico-Avalanche Detector". In: *European Patent EP3654376A1, US Patent US2021280734A1* (Nov. 2018).
- [61] H. F. W. Sadrozinski, A. Seiden, and N. Cartiglia. "4D Tracking with Ultra-Fast Silicon Detectors". In: *Reports on Progress in Physics* 81.2 (Dec. 2017), p. 026101. DOI: [10.1088/1361-6633/aa94d3](https://doi.org/10.1088/1361-6633/aa94d3).

- [62] The CMS Collaboration. “A MIP Timing Detector for the CMS Phase-2 Upgrade”. In: *CERN-LHCC-2019-003; CMS-TDR-020* (Mar. 2019). DOI: [10.17181/CERN.QZ28.FLHW](https://doi.org/10.17181/CERN.QZ28.FLHW).
- [63] The ATLAS Collaboration. “Technical Design Report: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade”. In: *CERN-LHCC-2020-007; ATLAS-TDR-031* (June 2020). DOI: [10.17181/CERN.FOZZ.ZP3Q](https://doi.org/10.17181/CERN.FOZZ.ZP3Q).
- [64] I. Perić et al. “High-Voltage CMOS Active Pixel Sensor”. In: *IEEE Journal of Solid-State Circuits* 56.8 (Aug. 2021), pp. 2488–2502. DOI: [10.1109/JSSC.2021.3061760](https://doi.org/10.1109/JSSC.2021.3061760).
- [65] J. Kemmer and G. Lutz. “New detector Concepts”. In: *Nucl. Instrum and Meth. A* 253.3 (Jan. 1987), pp. 365–377. DOI: [10.1016/0168-9002\(87\)90518-3](https://doi.org/10.1016/0168-9002(87)90518-3).
- [66] M. V. C. Marinas and M. Vos. “The Belle-II DEPFET Pixel Detector: A Step Forward in Vertexing in the superKEKB Flavour Factory”. In: *Nucl. Instrum and Meth. A* 650.1 (Sept. 2011), pp. 59–63. DOI: [10.1016/j.nima.2010.12.116](https://doi.org/10.1016/j.nima.2010.12.116).
- [67] A. Dorokhov et al. “High Resistivity CMOS Pixel Sensors and their Application to the STAR PXL detector”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 650.1 (Sept. 2011). International Workshop on Semiconductor Pixel Detectors for Particles and Imaging 2010, pp. 174–177. DOI: [10.1016/j.nima.2010.12.112](https://doi.org/10.1016/j.nima.2010.12.112).
- [68] N. Berger. “The Mu3e Experiment”. In: *Nuclear Physics B - Proceedings Supplements* 248-250 (Mar. 2014). 1st Conference on Charged Lepton Flavor Violation, pp. 35–40. DOI: [10.1016/j.nuclphysbps.2014.02.007](https://doi.org/10.1016/j.nuclphysbps.2014.02.007).
- [69] G. Aglieri Rinella et al. “The ALPIDE pixel sensor chip for the upgrade of the ALICE Inner Tracking System”. In: *Nucl. Instrum. Meth. A* 845 (Feb. 2017), p. 583. DOI: [10.1016/j.nima.2016.05.016](https://doi.org/10.1016/j.nima.2016.05.016).
- [70] R. Cardella. “CMOS detector and system developments for LHC detector upgrades”. Available at <https://repository.cern/records/hf5rf-mz156>. PhD thesis. University of Oslo, Nov. 2019.
- [71] R. Arcidiacono et al. “State-of-the-Art and Evolution of UFSD Sensors Design at FBK”. In: *Nucl. Instrum. Meth. A* 978 (2020), p. 164375. DOI: [10.1016/j.nima.2020.164375](https://doi.org/10.1016/j.nima.2020.164375).
- [72] R. Heller et al. “Characterization of BNL and HPK AC-LGAD Sensors with a 120 GeV Proton Beam”. In: *Journal of Instrumentation* 17.05 (May 2022), P05001. DOI: [10.1088/1748-0221/17/05/p05001](https://doi.org/10.1088/1748-0221/17/05/p05001).
- [73] M. Tornago et al. “Resistive AC-Coupled Silicon Detectors: Principles of Operation and First Results from a Combined Analysis of Beam Test and Laser Data”. In: *Nucl. Instrum. Meth. A* 1003 (2021), p. 165319. DOI: [10.1016/j.nima.2021.165319](https://doi.org/10.1016/j.nima.2021.165319).
- [74] A. Lai et al. “First Results of the TIMESPOT Project on Developments on Fast Sensors for Future Vertex Detectors”. In: *Nucl. Instrum. Meth. A* 981 (2020), p. 164491. DOI: [10.1016/j.nima.2016.05.016](https://doi.org/10.1016/j.nima.2016.05.016).
- [75] A. Lampis et al. “10 ps Timing with Highly Irradiated 3D Trench Silicon Pixel Sensors”. In: *Journal of Instrumentation* 18.01 (Jan. 2023), p. C01051. DOI: [10.1088/1748-0221/18/01/C01051](https://doi.org/10.1088/1748-0221/18/01/C01051).

- [76] F. Cadoux et al. "The 100 μ PET Project: A Small-Animal PET Scanner for Ultra-High Resolution Molecular Imaging with Monolithic Silicon Pixel Detectors". In: *Nucl. Instrum. Methods Phys. Res., A* 1048 (Mar. 2023), p. 167952. DOI: [10.1016/j.nima.2022.167952](https://doi.org/10.1016/j.nima.2022.167952).
- [77] J. Boyd. *The FASER W-Si High Precision Preshower Technical Proposal*. Tech. rep. Geneva: CERN, Mar. 2022. DOI: cds.cern.ch/record/2803084.
- [78] S.M. Sze, Y. Li, and K.K. Ng. *Physics of Semiconductor Devices*. 3rd ed. Wiley, 2012. ISBN: 9781119429111. DOI: [10.1002/9781119429111](https://doi.org/10.1002/9781119429111).
- [79] B. Streetman and S. Banerjee. *Solid State Electronic Devices*. 7th ed. Pearson Education, 2016. ISBN: 9780133356113. DOI: elibrary.pearson.de/book/99.150005/9781292060767.
- [80] M. Rudan. *Physics of Semiconductor Devices*. 7th ed. Springer International Publishing, 2018. ISBN: 9783319631547. DOI: [10.1007/978-3-319-63154-7](https://doi.org/10.1007/978-3-319-63154-7).
- [81] C. Canali et al. "Electron and Hole Drift Velocity Measurements in Silicon and their Empirical Relation to Electric Field and Temperature". In: *IEEE Transactions on Electron Devices* 22.11 (Nov. 1975), pp. 1045–1047. DOI: [10.1109/T-ED.1975.18267](https://doi.org/10.1109/T-ED.1975.18267).
- [82] Synopsys Inc. *Sentaurus TCAD: Technology Computer-Aided Design*. Available at <https://www.synopsys.com/silicon/tcad.html>. Synopsys. Mountain View, CA, USA, 2023.
- [83] C. Jacoboni et al. "A Review of some Charge Transport Properties of Silicon". In: *Solid-State Electronics* 20.2 (Feb. 1977), pp. 77–89. DOI: [10.1016/0038-1101\(77\)90054-5](https://doi.org/10.1016/0038-1101(77)90054-5).
- [84] K. G. McKay. "Avalanche Breakdown in Silicon". In: *Phys. Rev.* 94 (May 1954), pp. 877–884. DOI: [10.1103/PhysRev.94.877](https://doi.org/10.1103/PhysRev.94.877).
- [85] E. Gatti and P. F. Manfredi. "Processing the signals from solid-state detectors in elementary-particle physics". In: *La Rivista del Nuovo Cimento* 9.1 (Jan. 1986), pp. 1–146. DOI: [10.1007/BF02822156](https://doi.org/10.1007/BF02822156).
- [86] S. Centro. "Signal Amplification for High Capacitance Ionization Chambers". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 315.1 (May 1992), pp. 404–410. DOI: [10.1016/0168-9002\(92\)90736-N](https://doi.org/10.1016/0168-9002(92)90736-N).
- [87] A. Picardi. "Fast Timing SiGe BiCMOS Monolithic Silicon Pixel ASIC for HEP and Medical Imaging". Available at <https://cds.cern.ch/record/2921627>. PhD thesis. University of Geneva, Oct. 2024.
- [88] K. G. McKay. "Electron-Hole Production in Germanium by Alpha-Particles". In: *Phys. Rev.* 84 (Nov. 1951), pp. 829–832. DOI: [10.1103/PhysRev.84.829](https://doi.org/10.1103/PhysRev.84.829).
- [89] S. Goto and Z. Takagi. "Silicon Surface Barrier Detector". In: *Journal of Nuclear Science and Technology* 1.8 (Dec. 1961). DOI: [10.1080/18811248.1964.9732130](https://doi.org/10.1080/18811248.1964.9732130).
- [90] G. Andersson-Lindström and B. Zausig. "Silicon Surface Barrier Detectors; Fabrication, Test Methods, Properties and some Applications". In: *Nuclear Instruments and Methods* 40.2 (Mar. 1966), pp. 277–290. DOI: [10.1016/0029-554X\(66\)90387-9](https://doi.org/10.1016/0029-554X(66)90387-9).
- [91] P. Weilhammer. "Experience with Si Detectors in NA32". In: *Workshop on New Solid State Devices for High-energy Physics* (May 1986). DOI: [10.1016/0370-2693\(83\)90256-3](https://doi.org/10.1016/0370-2693(83)90256-3).

- [92] K. Borer et al. "Construction and Performance of a 1 m² Silicon Detector in UA2". In: *2nd Topical Seminar on Perspectives for Experimental Apparatus at Future High-energy Machines* (July 1987). DOI: [10.1016/0168-9002\(87\)90967-3](https://doi.org/10.1016/0168-9002(87)90967-3).
- [93] The ATLAS Collaboration. *ATLAS: Technical Proposal for a General-Purpose pp Experiment at the Large Hadron Collider at CERN*. LHC technical proposal. Available at <https://cds.cern.ch/record/290968>. Geneva: CERN, 1994.
- [94] The CMS Collaboration. *CMS, the Compact Muon Solenoid: Technical Proposal*. LHC technical proposal. Available at <https://cds.cern.ch/record/290969>. Geneva: CERN, 1994.
- [95] H. Bethe. "Zur Theorie des Durchgangs Schneller Korpuskularstrahlen durch Materie". In: *Annalen der Physik* 397.3 (Sept. 1930), pp. 325–400. DOI: [10.1002/andp.19303970303](https://doi.org/10.1002/andp.19303970303).
- [96] L. Landau. "On the Energy Loss of Fast Particles by Ionization". In: *Collected Papers of L.D. Landau*. Ed. by D. TER HAAR. Pergamon, 1965, pp. 417–424. ISBN: 978-0-08-010586-4. DOI: [10.1016/B978-0-08-010586-4.50061-4](https://doi.org/10.1016/B978-0-08-010586-4.50061-4).
- [97] R. Veenhof and the Garfield++ development team. *Garfield++: Simulation of Gaseous and Semiconductors Detectors*. Available at <https://garfieldpp.web.cern.ch/garfieldpp/>. 2017.
- [98] I. B. Smirnov. "Modeling of Ionization Produced by Fast Charged Particles in Gases". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 554.1 (Dec. 2005), pp. 474–493. DOI: [10.1016/j.nima.2005.08.064](https://doi.org/10.1016/j.nima.2005.08.064).
- [99] H. Bichsel. "Straggling in Thin Silicon Detectors". In: *Rev. Mod. Phys.* 60 (July 1988), pp. 663–699. DOI: [10.1103/RevModPhys.60.663](https://doi.org/10.1103/RevModPhys.60.663).
- [100] W. W. M. Allison and J. H. Cobb. "Relativistic Charged Particle Identification by Energy Loss". In: *Annual Review of Nuclear and Particle Science* 30. Volume 30, 1980 (Dec. 1980), pp. 253–298. DOI: [10.1146/annurev.ns.30.120180.001345](https://doi.org/10.1146/annurev.ns.30.120180.001345).
- [101] W. Riegler and G. Aglieri Rinella. "Time Resolution of Silicon Pixel Sensors". In: *Journal of Instrumentation* 12.11 (Nov. 2017), P11017–P11017. DOI: [10.1088/1748-0221/12/11/p11017](https://doi.org/10.1088/1748-0221/12/11/p11017).
- [102] H. Schindler. "Microscopic Simulation of Particle Detectors". Available at <https://cds.cern.ch/record/1500583>. PhD thesis. Vienna, Tech. U., Dec. 2012.
- [103] M. A. Green. "Self-Consistent Optical Parameters of Intrinsic Silicon at 300K Including Temperature Coefficients". In: *Solar Energy Materials and Solar Cells* 92.11 (Nov. 2008), pp. 1305–1310. DOI: [10.1016/j.solmat.2008.06.009](https://doi.org/10.1016/j.solmat.2008.06.009).
- [104] National Institute of Standards and Technology. *X-Ray Form Factor, Attenuation, and Scattering Tables*. URL: <https://physics.nist.gov/PhysRefData/FFast/html/form.html>.
- [105] W. Shockley. "Currents to Conductors Induced by a Moving Point Charge". In: *Journal of Applied Physics* 9.10 (Oct. 1938), pp. 635–636. DOI: [10.1063/1.1710367](https://doi.org/10.1063/1.1710367).
- [106] S. Ramo. "Currents Induced by Electron Motion". In: *Proceedings of the IRE* 27.9 (Sept. 1939), pp. 584–585. DOI: [10.1109/JRPROC.1939.228757](https://doi.org/10.1109/JRPROC.1939.228757).
- [107] L. Rolandi, W. Riegler, and W. Blum. *Particle Detection with Drift Chambers*. Springer Berlin, Heidelberg, 2008. DOI: [10.1007/978-3-540-76684-1](https://doi.org/10.1007/978-3-540-76684-1).

- [108] Cadence Design Systems, Inc. *Spectre Circuit Simulator*. Available at https://www.cadence.com/en_US/home/tools/custom-ic-analog-rf-design/circuit-simulation.html. Cadence Design Systems. San Jose, CA, USA, 2025.
- [109] FormFactor, Inc. *Cascade CM300xi Probe Station*. Available at <https://www.formfactor.com/products/probe-systems/semiautomated>. FormFactor. Livermore, CA, USA, 2023.
- [110] Keysight Technologies. *Keysight B1505A Power Device Analyzer / Curve Tracer*. Available at <https://www.keysight.com/us/en/products/parameter-device-analyzers-curve-tracer>. Keysight Technologies. Santa Rosa, CA, USA, 2024.
- [111] Y. Okuto and C. R. Crowell. "Ionization coefficients in semiconductors: A nonlocalized property". In: *Phys. Rev. B* 10 (Nov. 1974), pp. 4284–4296. DOI: [10.1103/PhysRevB.10.4284](https://doi.org/10.1103/PhysRevB.10.4284).
- [112] W. Maes, K. De Meyer, and R. Van Overstraeten. "Impact Ionization in Silicon: a Review and Update". In: *Solid-State Electronics* 33.6 (June 1990), pp. 705–718. DOI: [10.1016/0038-1101\(90\)90183-F](https://doi.org/10.1016/0038-1101(90)90183-F).
- [113] C. R. Crowell and S. M. Sze. "Temperature Dependence of Avalanche Multiplication in Semiconductors". In: *Applied Physics Letters* 9.6 (Sept. 1966), pp. 242–244. DOI: [10.1063/1.1754731](https://doi.org/10.1063/1.1754731). eprint: <https://doi.org/10.1063/1.1754731>.
- [114] M. Benoit et al. "The FE-I4 Telescope for Particle Tracking in Testbeam Experiments". In: *Journal of Instrumentation* 11.07 (July 2016), P07003. DOI: [10.1088/1748-0221/11/07/p07003](https://doi.org/10.1088/1748-0221/11/07/p07003).
- [115] Francesco Gramuglia et al. "Sub-10 ps Minimum Ionizing Particle Detection With Geiger-Mode APDs". In: *Front. in Phys.* 10 (Nov. 2022), p. 849237. DOI: [10.3389/fphy.2022.849237](https://doi.org/10.3389/fphy.2022.849237).
- [116] Rene Brun and Fons Rademakers. "ROOT — An Object Oriented Data Analysis Framework". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389.1 (1997), pp. 81–86. DOI: [10.1016/S0168-9002\(97\)00048-X](https://doi.org/10.1016/S0168-9002(97)00048-X).
- [117] M. Vicente Barreto Pinto. "New pixel-detector technologies for the ATLAS ITk upgrade and the CLIC vertex detector". Available at <https://www.unige.ch/dpnc/en/master-phd/phd/>. PhD thesis. University of Geneva, Dec. 2019.
- [118] H. Jeffreys. "An Invariant Form for the Prior Probability in Estimation Problems". In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 186.1007 (Sept. 1946), pp. 453–461. DOI: [10.1098/rspa.1946.0056](https://doi.org/10.1098/rspa.1946.0056).
- [119] C. J. Clopper and E. S. Pearson. "The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial". In: *Biometrika* 26.4 (Dec. 1934), pp. 404–413. DOI: [10.1093/biomet/26.4.404](https://doi.org/10.1093/biomet/26.4.404).
- [120] S. Abroug, F. Saadallah, and N. Yacoubi. "Optical and Thermal Properties of Doped Semiconductor". In: *European Physical Journal-special Topics - EUR PHYS J-SPEC TOP* 153 (Jan. 2008), pp. 29–32. DOI: [10.1140/epjst/e2008-00386-7](https://doi.org/10.1140/epjst/e2008-00386-7).
- [121] M. Ullàn al. "Evaluation of Silicon-Germanium (SiGe) Bipolar Technologies for use in an Upgraded ATLAS Detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 604.3 (June 2009), pp. 668–674. DOI: [10.1016/j.nima.2009.03.177](https://doi.org/10.1016/j.nima.2009.03.177).

- [122] J. D. Cressler. "Radiation Effects in SiGe Technology". In: *IEEE Transactions on Nuclear Science* 60.3 (Mar. 2013), pp. 1992–2014. DOI: [10.1109/TNS.2013.2248167](https://doi.org/10.1109/TNS.2013.2248167).
- [123] K. Nakamura et al. "Irradiation and Testbeam of KEK/HPK Planar p-type Pixel Modules for HL-LHC". In: *Journal of Instrumentation* 10.06 (June 2015), p. C06008. DOI: [10.1088/1748-0221/10/06/C06008](https://doi.org/10.1088/1748-0221/10/06/C06008).
- [124] G. D. Watkins. "Intrinsic Defects in Silicon". In: *Materials Science in Semiconductor Processing* 3.4 (Aug. 2000), pp. 227–235. DOI: [10.1016/S1369-8001\(00\)00037-8](https://doi.org/10.1016/S1369-8001(00)00037-8).
- [125] M. Bruzzi. "Radiation Damage in Silicon Detectors for High-Energy Physics Experiments". In: *IEEE Transactions on Nuclear Science* 48.4 (Aug. 2001), pp. 960–971. DOI: [10.1109/23.958706](https://doi.org/10.1109/23.958706).
- [126] M. Moll. "Displacement Damage in Silicon Detectors for High Energy Physics". In: *IEEE Transactions on Nuclear Science* 65.8 (Apr. 2018), pp. 1561–1582. DOI: [10.1109/TNS.2018.2819506](https://doi.org/10.1109/TNS.2018.2819506).