Ensuring superior performance: Characterizations of novel silicon detectors for High-Luminosity LHC and beyond

Dissertation
zur Erlangung des Doktorgrades

an der Fakultät für Mathematik und Naturwissenschaften
Fachbereich Physik
der Bergische Universität Wuppertal

vorgelegt von
Daniil D. Rastorguev
aus Moskau

Hamburg
2025

Supervisors	Dr. Paul Schütze Prof. Dr. Katerina Lipka
Members of the committee	Prof. Dr. Katerina Lipka Prof. Dr. Dirk Lützenkirchen-Hecht Prof. Dr. Wolfgang Wagner Dr. Paul Schütze Dr. Simon Spannagel
Date of the disputation	October 24th, 2025

Declaration

I hereby declare on oath that I have written the present dissertation on my own and have not used any aids and resources other than those acknowledged. The written version submitted corresponds to the one provided electronically. The presented dissertation in this or similar form has not been previously accepted in a doctoral procedure or assessed as unsatisfactory.

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebene Hilfsmittel und Quelle benutzt habe. Die eingerichtete schriftliche Fassung entspricht der auf dem elektronischen Speichermedium. Die Dissertation wurde in der vorgelegten oder einer ähnlichen Form nicht schon einmal in einem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt.

Daniil Rastorguev
Hamburg, 01 August 2025

Abstract

Modern particle physics is extremely reliant on advanced instrumentation, in particular, detectors. One of the key technologies for current and future experiments is semiconductor detectors. As experiments require higher and higher performance of detectors, continuous R&D is necessary to refine designs and improve performance.

The present thesis encompasses diverse R&D efforts on novel silicon detectors and discusses particular detector testing methods, with the two main research directions being developments of sensor characterization techniques by the means of pulsed lasers, and developments for the Phase-2 upgrade of the CMS experiment at the LHC.

The first part of the thesis is devoted to the laser techniques. The commissioning and upgrade of the Laserbox, an experimental setup for testing silicon sensors via charge injection with pulsed lasers, is presented. Furthermore, an approach for Monte-Carlo simulations of such laser injection experiments was developed using the Allpix^2 framework. These simulations were then validated by a comparison with the experimental data, obtained with the Laserbox. It was shown that the simulation is capable of accurately reproducing signal shapes, induced in silicon sensors in these experimental conditions.

The Laserbox was also used to study the DESY digital silicon photomultiplier (dSiPM) prototype, a novel monolithic pixelated photo-detector with CMOS SPADs as sensitive cells. A characterization campaign centering on timing features of the device was conducted. The time resolution of the dSiPM was found to be 53 ± 4 ps under optimal conditions. Meanwhile, the localized charge deposition with the laser allows one to resolve micrometer-scale features of the tested device, which revealed in-pixel variations of the dSiPM characteristics linked to the pixel cell layout.

The second part of the thesis covers the CMS Upgrade, discussing two particular aspects of production and testing of PS modules for the CMS Phase-2 Outer Tracker. First, the development of mechanical construction procedures for the modules and establishment of the robot-assisted assembly pipeline are discussed. These procedures achieve a micrometer-level precision during the assembly, which is crucial for the functioning of the novel p_T-discrimination feature of these modules. Second, a qualification campaign for the modules at the DESY II test beam facility is reported, with a focus on detection and p_T-discrimination efficiency. It was shown that the module is able to select tracks with a specified p_T at an efficiency of $98\pm0.2\%$, whereas outside the selection region this efficiency drops to below 1%. This campaign proves the production readiness of the module design from the particle detection functionality point of view.

Zusammenfassung

Die moderne Teilchenphysik ist in hohem Maße auf fortschrittliche Messgeräte, insbesondere Detektoren, angewiesen. Eine der Schlüsseltechnologien für aktuelle und zukünftige Experimente sind Halbleiterdetektoren. Da die Anforderungen an die Leistungsfähigkeit dieser Detektoren stetig steigen, ist kontinuierliche Forschung und Entwicklung notwendig, um deren Designs zu optimieren und die Sensoreigenschaften zu verbessern.

Die vorliegende Arbeit umfasst verschiedene Themen innerhalb der Forschung und Entwicklung zu neuartigen Siliziumdetektoren und diskutiert spezifische Charakterisierungsmethoden. Die beiden Hauptschwerpunkte liegen zum einen auf der Entwicklung von Sensorkarakterisierungstechniken mittels gepulster Laser und zum anderen auf Beiträgen zum Phase-2-Upgrade des CMS-Experiments am LHC.

Der erste Teil der Arbeit behandelt die zuvor genannten Lasertechniken. Es wird die Inbetriebnahme und Aufrüstung der Laserbox, einer Versuchsanordnung zur Charakterisierung von Siliziumsensoren mittels Ladungsinjektion mit gepulsten Lasern, vorgestellt. Darüber hinaus wurde ein Ansatz zur Monte-Carlo-Simulation solcher Laserinjektionsexperimente mit dem Allpix^2-Framework entwickelt. Diese Simulationen wurden durch den Vergleich mit Messdaten aus der Laserbox validiert. Es konnte gezeigt werden, dass die Simulation die unter diesen Bedingungen in Siliziumsensoren induzierten Signalformen präzise reproduziert.

Die Laserbox wurde weiterhin zur Charakterisierung des digitalen Silizium-Photomultiplier-Prototyps (dSiPM) von DESY eingesetzt, einem neuartigen monolithischen, pixelierten Photonendetektor mit CMOS-SPADs als empfindliche Zellen. Eine Charakterisierung des Detektorprototyps zeigte, dass die zeitliche Auflösung des dSiPM unter optimalen Bedingungen 53 ± 4 ps beträgt. Gleichzeitig ermöglicht die lokalisierte Deposition von Ladungsträgern mittels Laser die örtliche Auflösung von Sensormerkmalen im Mikrometerbereich, wodurch Abweichungen der Eigenschaften des dSiPM innerhalb eines Pixels aufgezeigt und mit dem Layout einzelner Pixelzellen korreliert werden konnten.

Der zweite Teil der Arbeit befasst sich mit dem Upgrade des CMS-Experiments und behandelt insbesondere die zwei Aspekte der Produktion und der Prüfung von PS-Modulen für den Spurdetektor. Zunächst wird die Entwicklung von mechanischen Montageverfahren sowie der Aufbau einer robotergestützten Montagelinie beschrieben. Diese Verfahren erreichen bei der relativen Ausrichtung einzelner Komponenten eine Präzision im Mikrometerbereich, was für die neuartige Funktion dieser Module zur p_T-Diskriminierung entscheidend ist. Anschließend wird die Qualifizierung der Module mit Messungen am DESY II Teststrahl vorgestellt, mit Fokus auf der Messung der Effizienz der einzelnen Sensoren und der p_T-Diskriminierung. Es wurde gezeigt, dass das Modul Spuren oberhalb einer definierten p_T-Schwelle mit einer Effizienz von $98\pm0{,}2\%$ selektiert, während diese Effizienz außerhalb des gewählten p_T-Bereichs auf unter 1% abfällt. Diese Messungen belegen die Produktionsreife des aktuellen Moduldesigns im Hinblick auf die Teilchendetektion.

Contents

Abstract vi

Zusammenfassung vii

1 Introduction to silicon detectors for particle physics 1
1.1 Scattering experiments in high-energy physics . . . 1
1.2 Particle detectors in scattering experiments . . . 4
1.3 Interactions of particles with matter . . . 6
1.3.1 Interaction of charged particles with matter . . . 6
1.3.2 Interaction of photons with matter . . . 8
1.4 Design principles of ionization detectors . . . 9
1.5 Semiconductor ionization detectors . . . 11
1.5.1 Charge carriers in semiconductors . . . 11
1.5.2 Design principles of silicon sensors . . . 16
1.5.3 Resolution of tracking detectors . . . 20
1.5.4 Radiation damage . . . 21
1.6 Development, production and deployment of silicon detectors . . . 23
1.7 Scope of this work . . . 25

I Characterizing silicon devices with pulsed lasers 27

2 Injection lasers for sensor R&D 29
2.1 Transient Current Technique . . . 29
2.1.1 Overview of charge injection techniques . . . 29
2.1.2 Charge injection with pulsed lasers . . . 30
2.1.3 Spatial configurations of charge injection . . . 33
2.1.4 Extractable observables . . . 35
2.2 The Laserbox setup . . . 37
2.2.1 Setup components . . . 38
2.2.2 Beam spot characterization . . . 42
2.2.3 Applications . . . 44

3 Monte Carlo simulations of laser injection 45
3.1 Monte Carlo simulations for detector development . . . 45
3.1.1 General concepts of modeling the detector response . . . 45

3.1.2 The Allpix² framework . . . 47
3.1.3 DepositionLaser module . . . 49
3.2 Comparison of simulation and experiment . . . 53
3.2.1 Experimental dataset . . . 53
3.2.2 TCAD modelling of the sensor . . . 54
3.2.3 Allpix² simulation of the laser injection . . . 55
3.2.4 Results . . . 58

4 Timing performance of a digital silicon photomultiplier prototype 63
4.1 Timing with silicon sensors . . . 63
4.2 Silicon sensors with gain . . . 65
4.2.1 Ways to increase signals . . . 65
4.2.2 Silicon Photomultipliers . . . 66
4.3 The DESY Digital SiPM Prototype . . . 67
4.3.1 dSiPM specifications . . . 68
4.4 Characterizations of the timing performance with a pulsed laser . . . 70
4.4.1 Adjustment of the laser system . . . 71
4.4.2 TDC characterization . . . 74
4.4.3 Intrinsic SPAD time resolution . . . 75
4.4.4 Matrix delay measurements . . . 76
4.4.5 In-pixel timing resolution . . . 77

Summary of Part I 87

II Building and testing PS modules for the CMS Phase-2 Outer Tracker 89

5 CMS experiment at the LHC 91
5.1 The Large Hadron Collider . . . 91
5.2 Layout of the CMS detector . . . 94
5.2.1 Design philosophy . . . 95
5.2.2 Magnet . . . 95
5.2.3 Muon system . . . 95
5.2.4 Calorimeter . . . 96
5.2.5 Tracker . . . 98
5.2.6 Triggering and data handling . . . 100
5.3 CMS Phase-2 Outer Tracker . . . 101
5.3.1 The concept of p_T−discriminating modules . . . 101
5.3.2 Module design . . . 102
5.3.3 Tracker layout and mechanics . . . 105

6 Construction of PS modules 107
6.1 Challenges of PS module production . . . 107
6.2 Assembly sequence of the Phase-2 OT modules . . . 110
6.3 Bare module automated assembly . . . 112
6.3.1 The assembly machine . . . 112
6.3.2 Glues and the glue dispensing robot . . . 115

6.3.3 Sequence . . . 117
6.3.4 Development of the assembly procedures . . . 118
6.3.5 Results of the bare module assembly . . . 121

7 Test beam qualification of PS modules **123**
7.1 Test beams for detector R&D . . . 123
7.2 DESY II test beam facility . . . 124
7.2.1 DESY II accelerator and beamlines . . . 124
7.2.2 Instrumentation . . . 126
7.3 Data reconstruction with Corryvreckan framework . . . 131
7.3.1 General concepts and procedures . . . 131
7.3.2 Setup alignment . . . 133
7.3.3 Telescope resolution . . . 134
7.3.4 Observables definitions . . . 135
7.3.5 Trigger offset . . . 136
7.4 Results . . . 138
7.4.1 Detection efficiency studies . . . 138
7.4.2 Efficiency of an irradiated sensor . . . 142
7.4.3 In-pixel studies . . . 144
7.4.4 Cluster size and angle calibration . . . 146
7.4.5 Spatial residual and coordinate resolution . . . 147
7.4.6 Stub performance . . . 149
7.4.7 Direct measurement of the incidence angle . . . 150

Summary of Part II **153**

Conclusion **155**

A Allpix^2 configuration for TCT simulation **157**

Bibliography **159**

List of acronyms **171**

Acknowledgments **175**

Chapter 1

Introduction to silicon detectors for particle physics

1.1 Scattering experiments in high-energy physics

In particle physics, the structure of matter is investigated through particle scattering experiments. To resolve very fine structures, high momenta of colliding particles are necessary. Natural sources of high-energy particles are *radioactivity* and *cosmic rays*, while the controlled conditions can be achieved at experiments at particle *accelerators*.

The **radioactivity** is a phenomenon of spontaneous decays of certain naturally existent unstable isotopes. In these decays, the nuclei transform to other nuclei with an emission of additional particles. Most common types of such decays are *α-decay*, which is an emission of an α-particle from the nucleus, and *β^--decay* which is a transformation of a neutron in the nucleus to a proton that remains bound in the nucleus, and an electron and an electron antineutrino that are emitted. The new nuclei could be produced in an *excited* state with a subsequent release of the excess energy via emission of a photon. Radioactivity is of interest for nuclear physics, neutrino physics, as well as applied research and medicine. For most of the radioactive isotopes, however, the kinetic energy of the emitted particles does not exceed several MeV, making them unsuitable for studying processes happening on a characteristic spatial scale of elementary interactions.

Cosmic rays are free highly energetic particles originating from distant cosmic objects, e.g. supernovae explosions or black holes, or the Sun and traveling at relativistic speeds. Cosmic rays mostly consist of protons (90%) and α-particles (9%) [1]. For some rare events the energy could be as high as hundreds of EeV [2]. While most of the particles are deflected by the magnetic field of the Earth, high-energetic ones reach the atmosphere. These particles then undergo multiple interactions with gas nuclei in the atmosphere, producing an *atmospheric shower*, a narrow stream of secondary particles of various kinds, each carrying a fraction of the primary particle's energy. Out of shower particles, mostly muons reach the surface due to their comparably low interaction probability. In addition, there is a background flux of neutrinos of various energies, originating from distant supernovae explosions, solar nuclear reactions, and other sources. From the particle physics perspective, cosmic ray experiments are of historical significance, providing first observations of positrons [3], muons [4], and pions [5], and they are also widely used in neutrino physics. However, for precise tests of the structure of matter the experimental use of cosmic rays is limited due to their relatively low flux, limited variety of particles, and their random spatial and temporal

distribution. Yet, secondary cosmic muons are used in applied research for particle physics, e.g. in low-cost and scalable test infrastructure for particle detectors [6] and for spatial alignment of large-scale particle physics experiments [7].

In order to produce high-energy particles in a controlled environment and at a high rate, dedicated *particle accelerators* are built. To be used for an experiment, each particle needs to be accelerated to a desired momentum and guided to the interaction point. The particles are accelerated electromagnetically, thus, only charged particles are used: conventionally electrons, protons, and their antiparticles. Generation of high-energy particles of other types (including electrically neutral ones) is also possible and is achieved by accelerating long-lived charged particles first and then directing them to a special *converter* medium, so that secondary particles would be produced, carrying a part of momentum of the primary charged particle.

The effect of an electric field $\vec{E}$ and a magnetic field $\vec{B}$ on a point-like particle with mass m and charge q, traveling at a velocity $\vec{v}$ can be described with the Lorentz force:

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}). \tag{1.1}$$

For relativistic particles the force $\vec{F}$ should be considered in a relativistic sense:

$$\vec{F} = \dot{\vec{p}} = \frac{d}{dt}\gamma m\vec{v}, \tag{1.2}$$

where γ is the so-called *gamma-factor*:

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{1.3}$$

and all quantities should be given in one frame of reference.

A particle accelerator conventionally consists of a pipe with high vacuum maintained inside, so that collisions of accelerated particles with e.g. gas molecules are excluded. Charged particles are injected into the pipe and accelerated with electric field, parallel to their travel direction. Historically, electrostatic fields created by electrodes held at a certain potential difference or at alternating voltages [8] were used to accelerate charged particles. Modern machines use so-called *radio frequency (RF) cavities* as their main acceleration units. The RF-cavities are electromagnetic resonators, in which a high-amplitude standing electromagnetic wave is maintained, and the electric component of this wave is used as driving electric field.

A stream of moving particles is referred to as a *beam*. Due to the nature of the standing wave in the RF-cavities, most accelerators are unable to produce beams with a constant uniform particle flux. Instead, beams typically have a periodic temporal structure: short and compact bursts of particles, called *bunches*, are accelerated at distinct time intervals. For steering and focusing high energy beams magnetic fields are more effective than electric fields. Electromagnets with the field perpendicular to the beam are used to guide the particles. Dipole magnets deflect the beam in one direction, whereas magnets with higher number of poles (e.g., quadru- and sextupoles) affect off-axis particles in the beam, thus being able to focus or defocus it.

Two possible variants of particle accelerators are circular and linear accelerators. The beam pipe in **linear accelerators** (*linac*s) is a straight line. Particles are injected at one end and are accelerated during a single passage through the machine. Linacs do not require extremely strong dipole magnets to maintain the beam trajectory, however, energies achievable by linacs are limited by the length of the accelerator and number of RF-cavities, since particles pass each cavity only once.

Circular accelerators have a beam pipe arranged in a ring-like shape, in which particles circulate, passing RF-cavities multiple times and gradually gaining energy at each revolution. In order to guide the beam along the circular trajectory, the ring is equipped with multiple dipole magnets, that *bend* the beam. It can be derived from Eq. 1.1 that a particle with momentum p and charge q in a constant magnetic field $\vec{B} \perp \vec{p}$ would orbit at a constant radius

$$r = \frac{p}{|q|B}, \tag{1.4}$$

known as *Larmor radius*. Usually it is not technically feasible to create a uniform magnetic field along the whole circumference of an accelerator, so that the particle trajectory is a perfect circle. Instead, the beam pipe would consist of arc segments with separate magnets, connected with straight line segments, where RF-cavities and focusing magnets are usually placed. During acceleration the particles' momentum increases, and in order to keep the beam in a circular trajectory, the bending magnetic field of the dipoles needs to be gradually increased. Accelerators with this mechanism implemented are called *synchrotrons*. Additionally, as the particles' velocity increases, the revolution period decreases, frequencies of the RF-cavities must be adjusted as well such that bunches pass the cavities always at the same phase of the standing wave. Although, when particles injected initially are already ultra-relativistic ($v \simeq c$), the velocity change is negligible, thus changing the radio frequency is not necessary. Another phenomenon, significantly affecting operation of circular accelerators, is *synchrotron radiation*. A charged particle, moving with non-zero acceleration (either radial or tangential), would emit electromagnetic radiation and thus lose energy [9]. Assuming motion of a relativistic particle in an circular accelerator with a constant radius r, one obtains the energy loss per revolution [10]:

$$\delta E_{\text{rev}} = \frac{q^2 \gamma^4 v^3}{3\epsilon_0 r c^3}. \tag{1.5}$$

Since the energy loss per revolution for ultra-relativistic particles scales as $\gamma^4 \sim E^4$, whereas energy gained per revolution from RF acceleration is constant, synchrotron radiation is a major limiting factor on the maximum energy achievable in circular electron[1] accelerators. For heavier particles, e.g. protons, the maximum energy is usually defined by the maximal bending field strength, that can be generated by dipole magnets.

High-energy accelerators usually consist of multiple stages. Every stage is an independent accelerator that is optimized for a certain range of energies and acts as an injector for the following stage.

Properties of a two-particle scattering process by a large extent depend on the total energy of the system in the center-of-mass reference frame, which in turn is affected by the geometry of the experiment. Two corner-case kinematic configurations are *fixed-target* and *collider* experiments. In fixed-target experiments the beam interacts with a stationary target, offering lower center-of-mass energy, but higher event rates. In collider experiments both colliding particles are accelerated to high momenta and brought to a head-on collision, which creates the initial state with a significantly higher center-of-mass energy. However, the experimental realization of collider setups poses significant engineering challenges: two beams need to be prepared, aligned and focused such that bunches synchronously cross at the desired *interaction point* to ensure the high rate of collisions. Certain accelerators, e.g. HERA [11] and KEKB [12], are designed such that momenta of colliding bunches are not exactly equal. These machines are known as *asymmetric colliders*.

[1] Since for the same energy electrons would have a higher gamma-factor, compared to other particles.

1.2 Particle detectors in scattering experiments

Properties of all stable particles in the final state of interactions, i.e. the type, trajectory, momentum, and energy of the particles are measured by dedicated scientific instruments called *particle detectors*. With this information, properties of the primary interaction at the collision point can be inferred. Short-lived particles, however, would decay before reaching detectors. Nevertheless, they can still be detected indirectly by their long-lived decay products.

Multi-purpose detectors in accelerator experiments are aimed at measuring all properties of the final state. They usually consist of several subsystems, arranged in layers, each designed for measuring certain properties of particles. In fixed-target experiments, detector layers are placed behind the target in a so-called *forward spectrometer* geometry, as the interaction products are boosted in the direction of the beam. In collider experiments, layers are typically shaped as concentric cylinders, or *barrels*, placed around the interaction point, as the distribution of particles in such a configuration is expected to exhibit axial symmetry. Coverage of the full solid angle is ensured by closing the barrel with *endcaps*, flat circular layers with similar functions. Detectors are typically segmented, i.e. divided to cells. This ensures that the desired quantities can be measured independently for multiple particles, crossing the detector simultaneously.

While the exact design choices and technologies used vary widely from experiment to experiment, the general selection and arrangement of subsystems stays approximately the same. A layout of a multipurpose detector in a collider experiment and signatures of passage of different particles in it are showcased with the CMS detector [13] at the LHC accelerator complex [14] in the Figure 1.1 (presented in detail in Chapter 5). In the following, the typical subsystems are discussed.

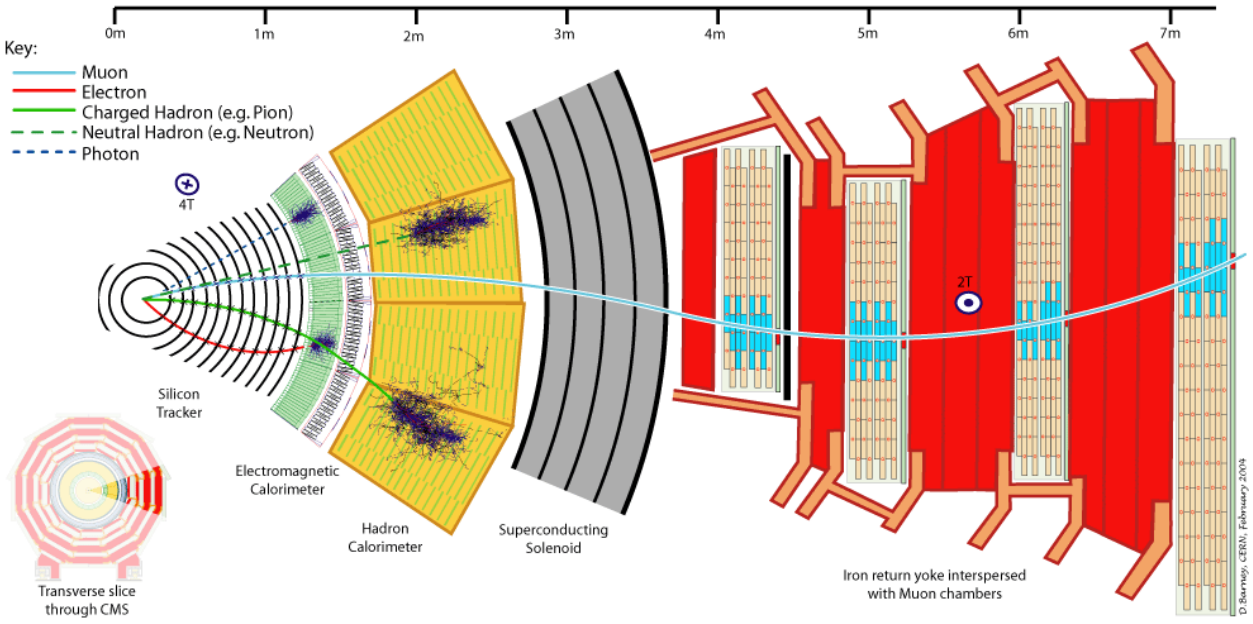


Figure 1.1: Layout of the CMS detector and passage of particles of different types through it. A sector of a transverse slice through the detector barrels is shown. From [15].

A **tracking detector** (or *tracker*) is usually the innermost subsystem. It determines trajectories of charged particles (*tracks*) passing through. To register particles, tracking detectors conventionally

rely on *ionizing energy loss* in sensitive planes (reviewed in Section 1.3). Tracking detectors are typically immersed in a magnetic field, so they are able to determine the momentum and the sign of the electric charge of passing charged particles. In a magnetic field a charged particle would have a helical trajectory, parametrized with a curvature radius r (as follows from Eq. 1.1, 1.4). At least three track coordinate measurements are required to determine the curvature and therefore the momentum, with the scheme being shown in Figure 1.2. For more complex arrangements with higher number of planes, a *fitting* procedure is used, i.e. numerical optimization algorithms are employed to find a helix line that is the best match to all measurements.

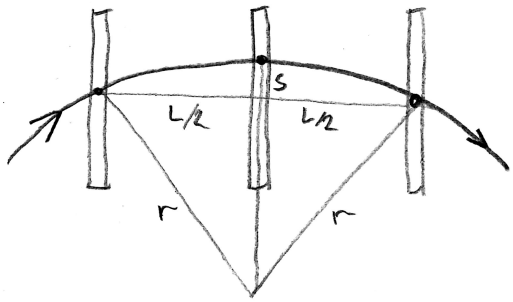


Figure 1.2: Momentum measurement with three planes, placed over distance L. The *sagitta* s is calculated, and the curvature radius is expressed as $r = L^2/8s$.

In addition, particle trajectories can be interpolated back to determine the *vertex* (i.e. point of origin). To date, *silicon* detectors (presented in Section 1.5), which are the main topic of this thesis, are superior in spatial and momentum resolution.

Systems for **particle identification** can be placed outside of the tracker.[1] The aim of these subsystems is to provide an auxiliary measurement of particles' velocity, which together with momentum readings from the tracker allows to infer particles' mass and to identify their type. One of the approaches, called *time-of-flight (ToF)* is a direct measurement of particle velocity by precisely measuring their travel time from the interaction point to a dedicated *timing layer*. Other approaches are based on registering electromagnetic emission, namely, *Cherenkov light* (e.g. RICH detectors [16]) and *transition radiation* (e.g. ATLAS TRT [17]), caused by charged particles crossing special radiators at a certain velocity.

A **calorimeter** measures particle energies via their full absorption in a dense medium. Light charged particles and photons deposit their energy via *electromagnetic showers*, which are consecutive chains of branching reactions of *bremsstrahlung* and *pair production* (as explained in Section 1.3). Hadrons, on the other hand, mostly deposit energy in *hadronic showers*, which are chains of strong interactions of hadrons with matter with creation of additional hadrons. To measure shower energy, calorimeters typically employ *scintillator* materials, that convert energy to a proportional amount of visible light, which in turn is picked up by photo-detectors. Calorimeters are often built in two separate parts, the inner *electromagnetic calorimeter (ECAL)* and the outer *hadronic calorimeter (HCAL)*, with each of the parts being optimized to measure energy of respective types of showers.

A **muon system** provides additional tracking and momentum information for muons. Since muons are highly penetrating particles, the muon system is usually placed outermost.

[1] Not present in the CMS detector (Figure 1.1).

1.3 Interactions of particles with matter

Particles of different types and energies interact with matter differently, which should be taken into account for designing particle detectors.

1.3.1 Interaction of charged particles with matter

Ionization losses

Tracking detectors are based exclusively on the phenomenon of *ionizing energy loss*. It is the major mechanism of interaction of charged particles with matter.[1] As the particle passes the medium, it undergoes a series of quasi-elastic interactions with atomic electrons, and in each interaction a fraction of the particle's energy is transferred to an electron. Electrons are then released from atomic shells, with excess energy, thus atoms of the medium become *ionized*.

The mean energy transferred by the particle to the medium via ionization per unit length $-\langle \mathrm{d}E/\mathrm{d}x \rangle$ can be obtained by integration over all kinematically allowed energy transfers per interaction, weighted by probability of such transfers, expressed by the Rutherford cross-section [18]. For a particle of mass M and charge $z = q/e$, passing the medium at the speed $\beta = v/c$, the mean energy loss is given by the *Bethe-Bloch formula* [19]:

$$-\langle \frac{\mathrm{d}E}{\mathrm{d}x} \rangle = K \cdot \frac{\rho Z}{A} \cdot \frac{z^2}{\beta^2} \cdot \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 T_{\max}}{I^2} - \underbrace{\beta^2 - \frac{\delta(\beta\gamma)}{2} - \frac{C(\beta\gamma, I)}{Z}}_{\text{correction terms}} \right], \tag{1.6}$$

where:

- $K = \dfrac{N_A e^4}{4\pi\epsilon_0^2 m_e c^2} \simeq 0.3 \, \dfrac{\mathrm{MeV} \cdot \mathrm{cm}^2}{\mathrm{mol}}$ is a constant;
- ρ, Z, A and I are parameters of the medium, respectively, density, atomic number, atomic mass and mean excitation energy;
- $T_{\max} = \dfrac{2m_e c^2 \beta^2 \gamma^2}{1 + 2\gamma m_e/M + (m_e/M)^2}$ is the maximal kinematically allowed energy transfer from the incoming particle to a free electron at rest [20];
- correction terms account for different effects, related to target electrons being in bound states in atomic shells of the medium:
  - term $-\beta^2$ for spin-orbital interactions of recoil electrons;
  - term $-\delta(\beta\gamma)/2$ for the density effect (i.e. shielding of the electric field of the particle by the medium), relevant for high values of $\beta\gamma$;
  - term $-C(\beta\gamma, I)/Z$ for the shell structure of the atoms, relevant for low values of $\beta\gamma$.

One major consequence that can be drawn from the Bethe-Bloch formula is that the dependency of the energy loss on the particle mass can be neglected in typical scenarios, and it is considered dependent only on the particle velocity. The sought value of $-\langle \mathrm{d}E/\mathrm{d}x \rangle$ for any particles is therefore

[1] However, for electrons and positrons the radiation loss is the primary mechanism due to their low mass.

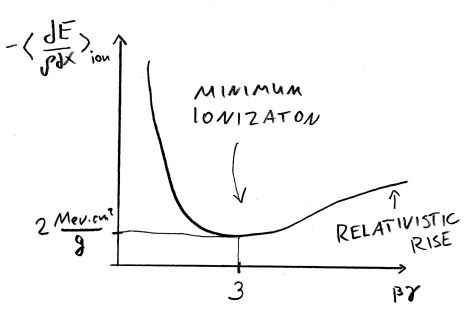


Figure 1.3: Mean energy loss per unit mass thickness as a function of $\beta\gamma$.

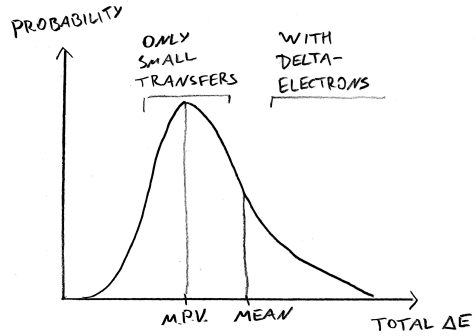


Figure 1.4: Distribution of the total energy loss in a thin layer.

conventionally presented as a function of the parameter $\beta\gamma$. For electrons, however, an additional minor correction needs to be made to account for non-distinguishability between the incoming and the shell electrons. Additionally, the energy loss is often normalized by the density of the medium and expressed as energy loss per unit of *mass thickness* $-\langle \mathrm{d}E/\rho\mathrm{d}x\rangle$. Since for most elements $Z/A \simeq 0.5\,\mathrm{mol/g}$, the density-normalized loss does not have a significant dependence on the chemical composition of the medium. The generalized dependence of $-\langle \mathrm{d}E/\rho\mathrm{d}x\rangle$ is schematically presented in Figure 1.3. Another important consequence that can be drawn is that ionization losses have a broad minimum at $\beta\gamma \simeq 3$ and a gradual increase proportional to $\ln\beta\gamma$ thereafter. Charged particles with $\beta\gamma \simeq 3$ and higher, as the logarithmic increase at higher velocities is relatively minor, are often called *minimum ionizing-particles (m.i.p.)*. Thus, the value at the minimum can serve as an approximation for the *mean* ionization loss of any high-energetic unitarily-charged particle in any medium:

$$-\langle \frac{\mathrm{d}E}{\rho\mathrm{d}x}\rangle_{\mathrm{mip}} \simeq 1 - 2\ \left[\frac{\mathrm{MeV}\cdot\mathrm{cm}^2}{\mathrm{g}}\right]. \tag{1.7}$$

For detector design applications, however, the quantity of interest is the *total* amount of energy lost over a *finite* path in the medium. One should consider that during the passage through matter the particle undergoes multiple statistically independent interactions with electrons in the medium. Whereas small energy transfers ($\mathcal{O}(\mathrm{eV})$) in each interaction are favored, there is a chance of a central collision with a significantly higher energy transfer, leading to a creation of a high-energy (keV-scale) recoil electron, called *δ-electron*. The distribution of the sum of individual contributions, or the total ionization energy loss, is given by a Landau-Vavilov distribution [21]. The general shape of such a distribution for a thin layer is shown in Figure 1.4. The distribution exhibits a significant asymmetry, with the main peak representing cases with no high-energy transfers, and the tail on the right describing more rare scenarios with a presence thereof. With an increase of the material thickness and number of interactions the distribution gets a more Gaussian-like shape.

Interactions with nuclei

Besides interactions with shell electrons, particles passing through matter also undergo Rutherford scattering on the nuclei. As the particles' mass is in general much smaller than the mass of the nucleus, the energy transfer in such a process is negligible, so the momentum transfer rather

changes the travel direction of the particle. Two effects, caused by interactions with nuclei, that should be accounted for when designing a tracking detector, are *radiation losses* and *multiple scattering*.

Radiation loss, or the *bremsstrahlung* effect, is the energy loss by emitting gamma-radiation, caused by charged particles undergoing acceleration [9] in the events of scattering on nuclei. As the effect depends on the particle's mass as $1/m^2$ [20], this is more relevant for light particles, i.e. electrons and positrons (and muons to some extent). The energy loss for electrons of energy E per unit distance is given by the relation [20]

$$-\frac{\mathrm{d}E}{\rho \mathrm{d}x}\bigg|_{\mathrm{rad}} = \frac{E}{X_0}, \tag{1.8}$$

where X_0, the *radiation length*, is a material-dependent constant. For different materials X_0 roughly scales as $1/Z^2$, and for most materials the density-weighted radiation length ρX_0 is typically in the range $10 - 50\,\mathrm{g/cm^2}$ [22, 23]. Since the radiation energy loss scales proportionally to E, whereas the ionization loss is proportional to $\ln \beta\gamma \simeq \ln E$, above a certain energy threshold bremsstrahlung will be the dominating mechanism of energy loss. The energy, at which $\mathrm{d}E/\rho\mathrm{d}x|_{\mathrm{rad}} = \langle \mathrm{d}E/\rho\mathrm{d}x\rangle|_{\mathrm{ion}}$, is called *critical energy*. Critical energy roughly scales as $1/Z$, and is of the order of dozens MeV for electrons and of hundred GeV for muons.

Multiple scattering is the effect of the deflection of particles from the original incidence direction. As the scattering arises from multiple independent random events, the total deflection angle after crossing material of a finite thickness can be approximated by a Gaussian distribution with a mean at 0. The standard deviation of this distribution, referred to as multiple scattering angle θ_{MS}, can be parameterized with the *Highland formula* [23]:

$$\theta_{\mathrm{MS}} = \frac{13.6\,\mathrm{MeV}}{c} \cdot \frac{z}{p\beta} \cdot \sqrt{\frac{x}{X_0}} \cdot \left(1 + 0.038 \ln \frac{x}{X_0}\right), \tag{1.9}$$

where z, p and β are electric charge (in e^-), momentum and velocity of the particle respectively, x is the distance crossed, and X_0 is the radiation length of the medium.

1.3.2 Interaction of photons with matter

Three main mechanisms of photon interaction with matter are the *photoeffect*, *Compton scattering* and *pair production*.

The **photoeffect** is the absorption of an incident photon by an atom with an emission of an electron. Photoeffect is the dominant mechanism of photon interaction with matter for energies below several tens of keV. The full energy of the photon is transferred to a single electron in this process. Assuming that the photon energy exceeds the binding energy [24] of a given electron, the electron is released from the atom with the kinetic energy $T = E_\gamma - E_{\mathrm{bind}}$. Interactions with inner shells (with higher E_{bind}) are favored, such that the kinetic energy of the released electron is minimal. Assuming a single atomic shell, the cross-section of photoeffect is higher for high-Z absorber materials. With an increase of the photon energy, the probability decreases rapidly. If the photon energy is less than the binding energy of the given shell, the cross-section is zero:

$$\sigma_{\mathrm{p.e.}}\big|_{\mathrm{shell}} \propto \frac{Z^5}{E_\gamma^{7/2}} \cdot \theta(E_\gamma - E_{\mathrm{bind}}), \tag{1.10}$$

where θ is the Heaviside step function. For lower photon energies ($\lesssim E_{\text{bind}}$), however, special considerations should be made regarding interactions of photons with electrons without the release of the latter, rather with a transition of an electron to a different bound energy level. If the matter is crystalline, its band structure has to be taken into account (introduced in Section 1.5.1).

Compton scattering is the interaction of a photon with a free electron (electrons of the medium can be considered quasi-free in this regard). In such a process, the photon is absorbed and a lower energy photon is re-emitted at an angle w.r.t. the incident photon trajectory with the excess energy being transferred to the electron. This is the main mechanism for absorption of photons in the energy range from tens of keV to MeV.

Pair production is the interaction of a photon with an electric field (mostly from nuclei of the matter) with an absorption of the photon and a creation of one electron-positron pair. This effect is allowed only if the photon energy exceeds the energy of the electron-positron pair products at rest: $E_\gamma \geq 2m_e c^2 = 1.022\,\text{MeV}$. For energies above this threshold, pair production becomes the prevalent type of interaction of photons with matter.

It should be mentioned that all these mechanisms of photon interaction involve a complete loss of the photon (or at least significant change of properties in case of Compton scattering) in a *single* interaction. For an ensemble, the Beer-Lambert law stands: if a beam of identical photons of intensity I_0 passes a distance x in the medium, its intensity I reduces over distance as

$$I(x) = I_0 \cdot \exp\left(-\frac{x}{\lambda}\right), \tag{1.11}$$

where the absorption depth λ is a function of material and the photon energy. Another aspect that all three interactions have in common is a release of an electron, which in turn would deposit its energy via ionization or bremsstrahlung.

Even though detection of photons is not the primary aim of semiconductor tracking detectors, they can be employed for photon detection as well. Important application domains are detection of visible light in scintillation or Cherenkov detectors and laser alignment of components in a large experiment [25].

1.4 Design principles of ionization detectors

Design of ionization detectors is based on registering ionization energy deposit in the sensitive medium, which is then associated with a particle passage. The detecting volume therefore needs to be sensitive to the ionized state of the material. The world's first design of an ionization detector is the *Geiger counter* [26], which consists of a conductive gas-filled tube and with a wire electrode. A passage of a charged particle crosses the tube causes a space discharge in the gas, detected as a pulse of the electric current in the voltage supply circuit. A notable feature of the Geiger counter is the direct and nearly instantaneous conversion of the ionization to electric signals. This is much desired particularly in modern experiments, as it allows for electronic data acquisition and processing systems with a high degree of automation.

Certain ionization detectors use a detection principle other than electric. First, detectors with the sensitive medium in a supercritical state, such as supercooled vapor in the *Wilson chamber* [27] or superheated liquid in the *bubble chamber* [28]. Energy deposition, caused by passing charged particles, is sufficient to locally trigger a phase transition, which is visible in unaided sight as a thin trace of condensation or evaporation spots. Second, *emulsion detectors*, employed in e.g.

OPERA [29] and FASER [30] experiments. In this approach, ionization causes a local chemical reaction in a special mixture (emulsion), typically produced in form of films or plates. After the emulsion is extracted from the experimental setup and processed, spots where the reaction took place can be distinguished by optical means.

Two common choices of materials for modern ionization detectors are gases and semiconductors. For both, the working principle is the registration of charge carriers in a material that is normally non-conductive, but allows for existence and free movement of charge carriers. All the energy deposited via ionization loss is primarily converted to the creation of charge carriers in the medium. If a charge carrier with high kinetic energy is created, it would also deposit its energy via consequent ionization. Total amount of charge carriers per unit distance is given by:

$$\frac{N_{\text{pairs}}}{\Delta x} = \frac{\langle \mathrm{d}E/\mathrm{d}x \rangle}{\langle \mathrm{E/pair} \rangle}. \tag{1.12}$$

In gases, ionization creates charge carriers in a form of electrons, released from atomic shells, and positive ions. With typical energies per electron-ion pair creation of 20-30 eV and densities of $\sim 1\,\mathrm{mg/cm^3}$, a minimum-ionizing particle would create approximately 50 electron-ion pairs per cm. In semiconductors, quasi-particles in the crystal lattice, namely, electrons (e)[1] and holes (h) act as charge carries. For silicon as an example with a mean *e-h* pair creation energy of 3.6 eV and density of 2.65 $\mathrm{g/cm^3}$, the approximate yield of charge carriers from an m.i.p. is 100 e-h pairs per µm.

In the external electric field, maintained in the detecting medium by a set of electrodes, the charge carriers drift to the electrodes, inducing an electric current, that can be picked up (as explained in Section 1.5.2). In order to provide tracking functionality of the detector, electrodes are often segmented and read out independently per segment, such that exact coordinates of the energy deposit can be inferred.

Both gases and semiconductors as detector materials are widely employed in particle physics experiments. Main advantages of semiconductor detectors are:

- Possibility of fine electrode segmentation, allowing for high spatial resolution (order of µm), which is beneficial for resolution on particle parameters;
- Fast response (sub-ns) and low dead-time (∼ns) due to charge carrier drift properties facilitating operation at high particle flux;
- Lower pair production energy, resulting lower Landau fluctuation for the same material budget ρx compared to gases;
- No necessity to maintain gas vessels;

with the disadvantages being:

- High cost per unit area, which makes their use in applications requiring high area coverage (e.g. muon systems) unreasonable;
- High power consumption;

[1] While commonly referred to as just *electrons*, electrons in bound states in the crystal lattice are quasi-particles and should be clearly distinguished from free electrons.

- Necessity for dedicated cooling systems for maintaining the detector functioning.

Whereas gas detectors are highly scalable in size and often provide high output signals (e.g. multi-wire proportional chambers [31]), semiconductors and in particular silicon are preferred for precision measurements at high rates. As of now, semiconductor detectors are the technology of choice for tracking detectors in most modern experiments.

1.5 Semiconductor ionization detectors

1.5.1 Charge carriers in semiconductors

Structure of semiconductors

The accurate description of the charge carriers in the semiconductor medium is crucial to use them in ionization detectors [20, 32]. The key feature that determines properties of semiconductor materials is an arrangement of atoms in a periodic grid, called *lattice*. Interaction of medium electrons with the periodic lattice potential results in *effective* laws of motion for bound electrons, significantly different from those for free electrons or from bound states in single atoms. A notable effect is the *band structure* of the electron spectrum, allowing electron states only with energies within certain intervals, called *bands* (unlike discrete levels in atoms and continuous spectrum for free electrons). With the finite density of states and the Pauli exclusion principle, bands can be fully *filled*, so that no more electrons within the given energy range can be added to the material. The exact arrangement of the bands and their occupation defines the type of the material. If at $T = 0$K there is a partially filled band, the substance is electrically conductive and is considered *metal*. If the uppermost filled band, called *valence* band, is completely filled and the next available band, called *conduction* band is completely empty, this substance is either a *semiconductor* or an *insulator*. The distinction between the last two is made by the *band gap* E_G – the width of the forbidden interval of energies between the conduction and valence bands. The exact value of the band gap affects phyiscal properties of the substance. Materials with $0 < E_G \lesssim 4$ eV are typically classified as semiconductors, and materials with $E_G \gtrsim 4$ eV are insulators. At $T = 0$K, neither semiconductors nor insulators are electrically conductive, because the valence band is completely filled, meaning that all the possible momentum states in this band are occupied, so no electron can change its momentum, thus collective electron motion to facilitate the electric current is not possible.

Figure 1.5 shows the schematic structure of the electron energy spectrum in silicon. On the x-axis the electron quasi-momentum $\hbar k$ is shown, which in a crystal lattice is periodic with a period $2\pi\hbar/a$, where a is the lattice period. The lower branch on the Figure 1.5 represents the valence band and the upper branch represents the conduction band. There are no possible states with energy inside the band gap. The width of the band gap in silicon is 1.12 eV.

If an electron is given a sufficient amount of energy ($\geq E_G$) via thermal excitation or via particle-induced ionization, it can transfer to the conduction band, leaving an unoccupied state in the valence band, called a *hole*. Holes can be treated as quasi-particles, that can effectively move and carry positive unit charge, even though this is in fact created by a collective motion of bound electrons, occupying all the states *but one*. Both electrons in the conduction band and holes act as free charge carriers in semiconductors. The effective masses of conduction electrons and holes can be computed from the shape of the branches of the electron spectrum, e.g. in silicon: $m_e^* = 1.14 m_e$ and $m_h^* = 1.01 m_e$ [33].

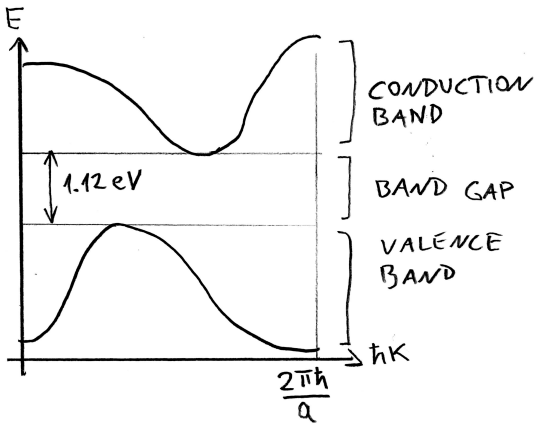


Figure 1.5: Schematic structure of the electron spectrum in silicon. Two uppermost bands are shown. Adapted from [32].

The net concentration n_i of thermally created charge carriers, both electrons (n_e) and holes (n_h), in a pure, or *intrinsic* semiconductor, as a function of temperature T, is given by:

$$n_i = \sqrt{n_e n_h} = 2 \cdot \left(\frac{kT}{2\pi\hbar^2} \cdot \sqrt{m_e^* m_h^*} \right)^{\frac{3}{2}} \cdot \exp\left(-\frac{E_G}{2kT} \right) \tag{1.13}$$

For silicon at room temperature this yields $n_i \sim 10^{10}/\mathrm{cm}^3$. Meanwhile, the amount of *e-h* pairs, generated by ionization is of the order of 10^4 per m.i.p., assuming a typical thickness of a silicon sensor of $\mathcal{O}(100\,\mu\mathrm{m})$. Therefore, the number of additional charge carriers is negligible compared to thermally created ones. There exist techniques to suppress intrinsic carrier generation in order to better distinguish the actual signal. These methods are often specific to the semiconductor material:

- High-purity germanium detectors can be operated at low temperatures [34];
- In gallium arsenide detectors one may exploit certain bulk impurities that suppress thermal carrier generation [35];
- Diamond has high E_G value, therefore thermal carrier generation is intrinsically low even at room temperatures [36];
- For silicon the common method is use of *p-n junctions* that create a so-called *depleted zone*.

p-n junction in silicon

The properties and concentration of charge carriers in silicon can be altered to a major extent by introducing *dopants* in the crystal. Other elements can be added to the material, so that a controlled fraction of silicon atoms in the lattice is replaced by atoms of other kind. The key properties of the doping are defined by its group in the periodic table. Silicon itself belongs to group IV and its atom has 4 electrons in the outer shell.

Group-III elements (typically boron), when added to silicon, are called *acceptor* dopings. Acceptor doping results in an additional empty energy level in the band gap, close to the edge of the

valence band. At non-zero temperatures, electrons from the valence band transition to this acceptor level, thus leaving empty states in the valence band, effectively creating excess holes. Such a configuration is called *p-type* silicon.

Group-V elements (typically arsenic), when added to silicon, are called *donor* dopings. Donor doping results in an additional filled energy level in the band gap, close to the edge of the conduction band. At non-zero temperatures, electrons from the donor level transition to the conduction band, effectively creating excess electrons. Such a configuration is called *n-type* silicon.

In a doped, or *extrinsic*, semiconductor, an excess of a certain type of charge carriers is present. Since the material as a whole is electrically neutral, it can be thought of as a lattice with *space charge* with some free charge carries present: p-type has a negatively charged lattice with excess holes, and n-type has positively charged lattice with excess electrons. For typical doping concentrations of $10^{12} - 10^{19}/\mathrm{cm}^3$ the amount of doping-induced carriers is significantly nigher than the amount of the thermally created ones, and all the carries have predominantly one type. Carrier concentrations in extrinsic semiconductors are:

- in p-type: $n_h \simeq N_a$, $n_h \gg n_e$;
- in n-type: $n_e \simeq N_d$, $n_e \gg n_h$;

where N_a and N_d are the concentrations of the acceptor and donor dopings respectively. The electric conductivity of extrinsic silicon is much higher than of the intrinsic one, and is mostly provided by a single type of carriers. In presence of both types of dopants, the resulting silicon bulk type is determined by the dopant with a higher concentration, and the effective doping concentration reads

$$N_{\mathrm{eff}} = |N_a - N_d|. \tag{1.14}$$

Doped silicon can be used to create a specific structure, called a *p-n junction*. If p- and n-type bulks are brought in contact[1], holes from the p-bulk and electrons from the n-bulk recombine with each other, thus reducing concentrations of free carriers at the contact surface. As the excess carriers do not compensate for the lattice space-charge anymore, an electric field is created by the negative space charge in the p-bulk and the positive space charge in the n-bulk, that counteracts the diffusion. The equilibrium state is reached eventually, featuring an insulating zone of finite thickness, totally free of charge carriers and referred to as the *depleted zone*. The potential difference created by the p-n junction is called *built-in voltage*; in silicon, $V_{\mathrm{bi}} \simeq 0.6\,\mathrm{V}$ for typical doping concentrations. The structure of a p-n junction and spatial distributions of physical parameters are shown in Figure 1.6.

If electric terminals are attached to both sides, p and n, the equilibrium of the depleted zone can be shifted by applying an external voltage. The electrical properties of the junction are as follows (see Figure 1.7):

- If a positive potential is applied to the p-side (*forward bias*), the depleted zone shrinks. As soon as the external voltage exceeds the built-in potential , the thickness of the depleted zone reaches zero, and the p-n junction becomes conductive, since there is no more insulating layer in the middle.

[1]Typically done by reverting a layer on top of a doped bulk by adding a high amount of doping of the opposite type.

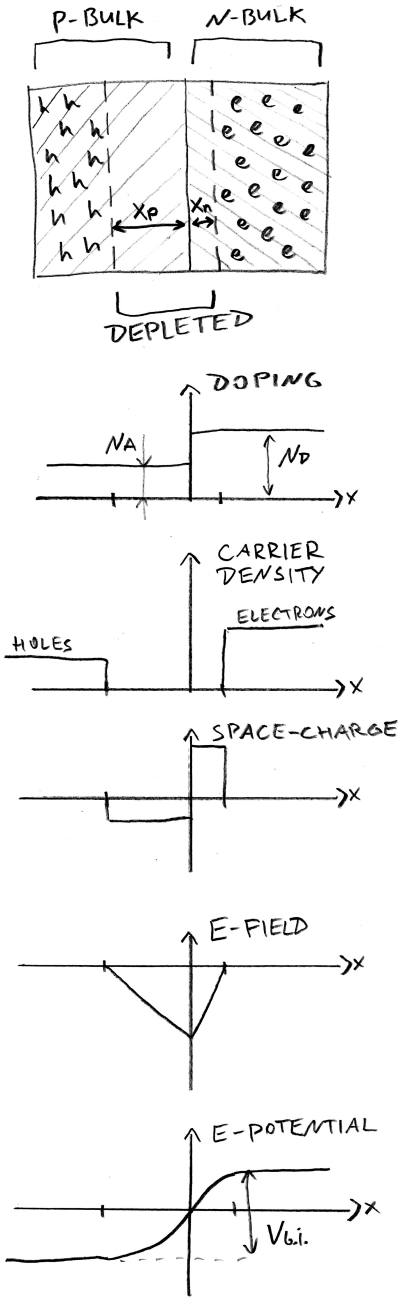


Figure 1.6: An unbiased p-n junction and schematic spatial distributions of electrostatic properties of the bulk as functions of the coordinate across the interface. Adapted from [20].

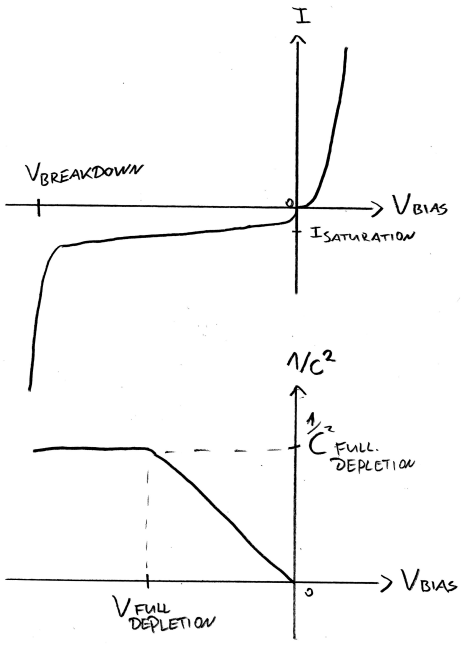


Figure 1.7: Electric properties of a p-n junction as functions of the external voltage. The positive voltage denotes a forward bias.
Top: current-voltage characteristic. Bottom: capacitance-voltage characteristic. The inverse square of capacitance is shown to emphasize the dependency of Eq. 1.17.

- Applying a positive potential to the n-side (*reverse bias*) results in growth of the depleted zone. The junction remains generally insulating, with only a minor *leakage* current flowing between terminals. The exact origin of the leakage current is from charge carriers, diffusing to the depletion zone from the non-depleted bulk [20, 37], and generation of electron-hole pairs in the depleted zone (either by certain defects or intrinsic thermal). If a certain threshold (highly dependent on exact properties of the junction, for typical designs is of order 10-100 V) is exceeded, a *breakdown* occurs, resulting in rapid increase of the reverse current.

The extent of the depleted zone from the interface towards the p-type (x_p) and to the n-type bulk (x_n) is given by:

$$\begin{aligned} x_p &= \sqrt{\frac{2\epsilon\epsilon_0}{e} \cdot \frac{N_d}{N_a(N_a+N_d)} \cdot (V_{\text{bi}} - V_{\text{ext}})}, \\ x_n &= \sqrt{\frac{2\epsilon\epsilon_0}{e} \cdot \frac{N_a}{N_d(N_a+N_d)} \cdot (V_{\text{bi}} - V_{\text{ext}})}, \end{aligned} \tag{1.15}$$

where the external voltage V_{ext} is defined positive for forward bias, and $V_{ext} < V_{bi}$. The p-n junction as a whole remains electrically neutral:

$$\frac{x_p}{N_a} = \frac{x_n}{N_d}. \tag{1.16}$$

The p-n junction with an insulating depleted zone in the middle, surrounded with a conductive doped bulk, acts as a plate capacitor. Being inversely proportional to the width of the depleted zone, its capacity c per unit area decreases if a reverse bias is applied:

$$c = \frac{\epsilon\epsilon_0}{x_p + x_n} \underset{V_{\text{bias}} \gg V_{\text{bi}}}{\propto} \frac{1}{\sqrt{|V_{\text{bias}}|}} \tag{1.17}$$

The depleted zone at the p-n junction is therefore an optimal choice for an ionization detector. As shown before, it is free of intrinsic charge carriers, but allows for existence and movement thereof. If a e-h pair is created, its recombination is suppressed due to indirect band structure of silicon. In order to satisfy the conservation of energy and quasi-momentum simultaneously, additional momentum contribution is required from a *phonon*, a quantum of the lattice vibrations. At room temperature this results in carrier lifetimes of $\sim$ms, which is orders of magnitude higher than the characteristic charge collection time.

Absorption of optical photons in silicon

The band structure of silicon should be taken into account to infer properties of interaction of optical photons via photoeffect. Photons of the optical and near-UV range ($E_\gamma = \frac{hc}{\lambda} \sim$ few eV) would typically produce a single e-h pair per photon. In such a regime, the absorption probability for a photon with a given energy is determined by the availability of a transition requiring this exact amount of energy. For $E_\gamma < E_G$ ($\lambda > 1107$ nm) there are no corresponding transitions in intrinsic silicon, thus this process is fully suppressed at $T = 0$K, and is mostly suppressed at nonzero temperatures. For $E_\gamma \simeq E_G$, where energy transition with $\Delta E = E_G$ is in principle allowed, it also requires additional quasi-momentum input from the lattice. A further increase of E_γ gradually relaxes requirements on phonon absorption, thus the cross section-increases until *mean* energy of e-h

pair production of 3.6 eV ($\equiv$ 344 nm) is reached. The absorption length (inversely proportional to the photoeffect cross-section) for photons of visible, IR-, and UV-light is shown in Figure 1.8. With further increase in energy the interaction probability declines proportionally to $E^{-7/2}$ (as given by Eq. 1.10).

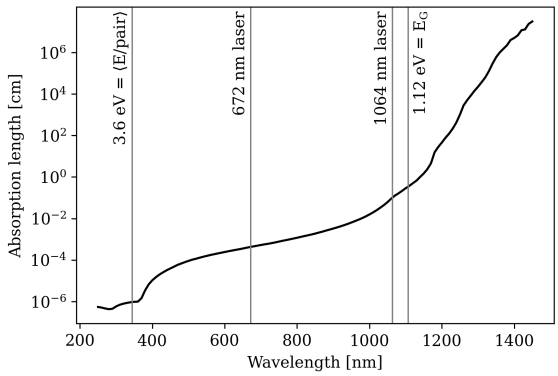


Figure 1.8: Depth of optical photons in silicon as a function of the wavelength [38].

1.5.2 Design principles of silicon sensors

Silicon manufacturing

Actual designs and implementations of silicon sensors are to a major extent driven by available manufacturing technologies. The standard industrial production pipeline consists of the following steps [39]:

1. Mono-crystalline silicon is produced in form of cylindrical ingots. Dopants can be added, such that the entire crystal is of p- or n-type. Different crystal growth techniques include the *Czochralski process*, the continious vapor deposition (CVD) and the float zone method (FZ). For particle physics applications, FZ silicon is the common choice because this technique allows to achieve the smallest amount of impurities in the entire crystal, that would later on affect the performance of the detector.

2. Silicon ingots are cut into thin ($\sim$hundreds of μm) discs, called *wafers*. Typical production wafers are available in diameters of 6 or 8 inches.

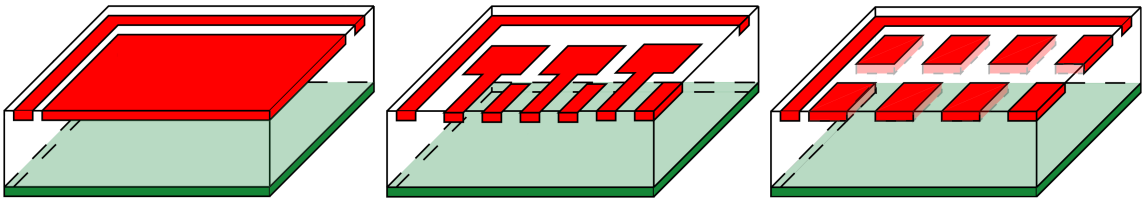
3. Further processing of wafers is performed with the *lithography* process. Different treatments (usually in a form of applying chemical agents) are performed on wafers, altering a thin layer on the surface. For example, coatings of metal, insulator or semiconductor can be deposited on the surface, layers of material can be removed, additional doping can be added to the thin surface layer of silicon. The key element of the lithography process is that all the modifications mentioned above can be performed on selected regions on a wafer thanks to the use of photo-activated masks. Modern technology allows for forming structures on wafers with features as small as several nanometers. By repeating different lithography steps, one can create complex arrangements, that would serve as the particle sensors or as electronic components.

4. Complete products, e.g. sensors, are cut out from the wafers.

Sensor layout

In the following, the design with the p-type bulk is discussed. The design and the working principles remain the same for the n-bulk.

As a basic material for detectors, wafers with relatively low concentration ($\sim 10^{11} - 10^{12}$ /cm^3) of p-type doping are used. Regions with higher (up to 10^{18} /cm^3) doping, called *implants*, are created on its surfaces via lithography. The frontside usually has n$^+$-type implants, with the plus sign denoting a much higher concentration of the doping, compared to the bulk. Such configuration is called *n$^+$-in-p*. The frontside implants together with the p-bulk form a p-n junction, which during operation of the detector is kept at a reverse bias, creating the depletion region. Since the doping concentration in the implants is much higher than in the bulk, the depletion region grows almost exclusively to the bulk, from the frontside towards the *backside*, and not to the implants (as given by Eq. 1.16). An important operational aspect is the *full depletion*. At a certain reverse bias voltage, the depletion region extends to the full thickness of the bulk. At this point, the junction capacity (Eq. 1.17) stops increasing with the further increase of the bias voltage. Detectors are almost exclusively operated at a bias voltage above the full depletion voltage, as maximizing the active thickness is crucial for the detector performance. Frontside implants commonly feature segmentation in shapes of pads, strips, or pixels (see Figure 1.9), depending on the required spatial resolution. The frontisde implants are typically used as *readout* implants, i.e. electric signals from ionizing particles are picked up from them. Segmentation of the detector provides spatial resolution. The backside in this layout is fully covered with a p$^+$-implant to provide ohmic electric contact to the bulk [39] in order to create a second terminal for applying the bias voltage.



(a) Pad. **(b)** Strips. **(c)** Pixels.

Figure 1.9: Types of segmentation of silicon detectors. Adapted from [20].

Similarly, a *p$^+$-in-n* design is possible. Another less common configuration is *n$^+$-in-n* with n-bulk, n$^+$-implants and a p$^+$-backside. The depletion region in this layout would grow from back to front.

A scheme and working principle of a segmented silicon detector is shown in Figure 1.10, together with the readout electronics.

Motion of the carriers

The depleted p-n junction configuration allows for the collection and detection of charge carriers, created by a passage of an ionizing particle, thus allowing for the particle itself to be registered. In the depleted region, charge carriers are exposed to the electric field, created by the lattice space charge and by the external bias potential. The motion of electrons and holes in a semiconductor is

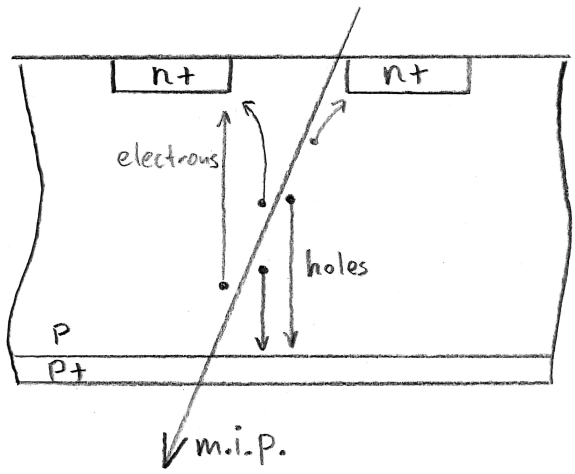


Figure 1.10: Scheme and working principle of an n^+-in-p segmented silicon detector.

described by the Drude model [40]. In an electric field $\vec{E}$ charge carriers drift at a constant velocity:

$$\vec{v}_{e,h} = -\mu_{e,h} \cdot \vec{E} \tag{1.18}$$

where the coefficients μ_e and μ_h are called electron and hole *mobility* respectively. The mobility is a function of temperature and doping concentration. The law stands for fields below saturation $E < E_{\text{sat}} \sim 10^4 - 10^5$ V/cm, at which carrier the velocity saturates and the linear relation is not valid anymore, thus formally the mobility is dependent on the electric field. For lower electric fields, typically used in sensors, an approximation $\mu = \text{const}(E)$ is sufficient. For silicon at room temperature and doping concentrations, typical for the bulk, carrier mobilities are $\mu_e \simeq 1300\,\text{cm}^2/(\text{V}\cdot\text{s})$ and $\mu_h \simeq 450\,\text{cm}^2/(\text{V}\cdot\text{s})$.

Besides the electric-field-induced drift, each carrier undergoes a process of random walk, or *diffusion*. Over time t, each carrier with charge q randomly shifts by some distance, with the mean deviation along axes given by the Einstein-Smoluchowski relation:

$$\sqrt{\langle \Delta x^2 \rangle} = \sqrt{\langle \Delta y^2 \rangle} = \sqrt{\langle \Delta z^2 \rangle} = \sqrt{\frac{2\mu k T t}{q}}, \tag{1.19}$$

where k is the Boltzmann constant. Whereas the motion of charge carriers is dominated by the electric field, the carriers' trajectories may deviate from the field lines due to diffusion. Under typical circumstances, for both electrons and holes, Eq. 1.19 yields the deflection of the order of 1 µm per 100 µm traveled.

Signal formation

Movement of charge carriers in the depleted bulk induces electric current on the electrodes. Instantaneous current, induced on n-th electrode by a charge q at a coordinate $\vec{r}$ and moving at a velocity $\vec{v}$, is given by the Shockley-Ramo theorem [41, 42]:

$$i_n = q\vec{v} \cdot \vec{\nabla} W_n(\vec{r}), \tag{1.20}$$

where W_n is the weighting potential of the n-th electrode. $W_n(\vec{r})$ is a dimensionless scalar potential, describing the spatial configuration of the electrode. W_n is defined as a potential in volts, that would be established if the n-th electrode is kept at a unit potential and the rest are kept at zero. Therefore, on the surface of the n-th electrode $W_n = 1$, and on the surfaces of other electrodes and at infinity $W_n = 0$. For an infinite flat capacitor without segmentation and planes located at $z = 0$ and $z = d$, the weighting potential for the electrode at $z = d$ is a linear function of the coordinate across the capacitor:

$$W(x, y, z) = \begin{cases} \frac{z}{d} & \text{if } z \in [0, d], \\ 0 & \text{elsewhere.} \end{cases} \tag{1.21}$$

For detectors with segmented electrodes in a form of small (width $\ll$ sensor thickness) implants, for each given electrode its weighting potential is nearly zero in the major part of the bulk, and only starts to increase in a close proximity (distance $\sim$ width) of the implant. An example of a weighting potential distribution in a strip sensor is shown on Figure 3.9.

By integrating Eq. 1.20 along the particle trajectory, one obtains the total charge, induced on the electrode n, as one charge carrier travels from $\vec{r_1}$ to $\vec{r_2}$:

$$\delta Q_n = q(W_n(\vec{r_2}) - W_n(\vec{r_1})). \tag{1.22}$$

As a consequence, signals from m.i.p.s in segmented silicon detectors are mostly generated by one type of charge carriers. The majority of carriers are generated in the bulk region, where the weighting potential is nearly zero. Only the carriers that drift towards the readout implants undergo a change in the potential along their trajectory so that they induce current on readout implants. In this regard, p^+-in-n sensors are referred to as *hole-collecting* and n^+-in-p as *electron-collecting*.

Principles of signal readout

The combined contributions of the induced current, created by individual charge carriers, create a distinct electric pulse at the readout electrodes. The duration of the pulse for a typical silicon sensor (thickness of few hundred μm, V_{bias} of several hundred volts) is of the order of nanoseconds. An electric circuit, called *front-end*, is employed to pick up and process these signals in order to detect a passing particle. Two main parts of the front-end are an *amplifier* and *digitization logics*. The amplifier increases the level of the induced signal (currents of some nA), such that it can be reliably processed with standard electronics. The digitization stage analyzes analog outputs of the amplifier to register signals, caused by particle-induced ionization.

The most basic digitization technique, called *binary readout*, is a comparison of the signal with a fixed *threshold*. If the signal level exceeds a certain pre-set value, this is interpreted as a particle passage. The exact threshold value has to be carefully chosen such that fluctuations on the amplifier baseline (e.g. caused by thermal carrier generation at the semiconductor bulk, or intrinsic noise of the amplifier) are unlikely to cross the threshold, in contrast to the expected signal from ionization. The binary readout does not provide any information on the properties of the signal. However, the signal amplitude or integral might be required in some cases. For instance, *weighting* of the signals in adjacent electrodes can be performed to interpolate the exact particle passage position, if several channels are triggered [20]. Common digitization techniques involve the use of *charge-to-digital converters (QDCs)* to measure the pulse integral or the *time-over-threshold (ToT)* method to

measure the duration ($\sim$ amplitude) of the pulse. For studies of certain properties of the sensor itself, high-frequency sampling of the full pulse can be performed (as presented in Chapter 2.1).

In many modern silicon particle detectors, each channel is equipped with an independent front-end circuit in order to facilitate operation at high repetition rates. Front-end electronics are often implemented as dedicated chips, or *application-specific integrated circuits (ASICs)*. The front-end chips are connected directly to readout implants (*DC-coupling*), or a capacitor is added in between to filter out constant currents (*AC-coupling*). Given the very fine segmentation of modern sensors (100 μm and below), thorough engineering is required to establish a connection to every sensor channel. For strip sensors, the front-end is typically placed near one end of the sensor and electrically attached to it with *wire-bonds*. For pixel sensors, the front-end ASICs cover the front-side of the sensor entirely, and are attached via the *flip-chip bump-bonding* technique [43] to form a so-called *hybrid* pixel detector.

Recent advances in lithography allow for integration of electronic components directly into the sensor wafer. Thus, a sensor and a fully-capable front-end can be implemented as a single part. This technology, called *monolithic active pixel sensors (MAPS)* is a very perspective research direction for particle detectors, and already has large-scale applications, e.g. in ALICE ITS2 [44]. The benefits of this technology include low material budget and lower power consumption of detectors, facilitating the design of high-precision trackers.

1.5.3 Resolution of tracking detectors

Characteristics of individual silicon detectors by a large extent define the performance of the entire tracker. Two important parameters of particles measured by trackers are the *transverse momentum* p_T and the impact parameter d_0. The transverse momentum[1] is derived from the curvature of the reconstructed trajectory. Combined with the angle θ between the trajectory and the magnetic field, it allows the full momentum of the particle to be determined. The impact parameter is the shortest distance between the trajectory and the interaction point.

The major sources of uncertainty in measurements of these quantities are the finite precision of coordinate measurement in each detector plane and multiple scattering (introduced in Section 1.3.1) in detectors, which distorts the trajectory. Here we consider a model case [20] with the following assumptions:

- Multiple ($N \gtrsim 10$) identical flat thin detectors are placed over a distance L and uniformly spaced;
- Each detector has a thickness x, consists of material with the radiation length X_0, and has a spatial resolution σ_meas;
- the constant magnetic field $\vec{B}$ is parallel to the detector planes;
- The particle is unitarily charged and has a sufficiently high momentum, so that $r \gg L$.

The relative error of p_T, related to the coordinate measurement uncertainty is given by the Gluckstern formula [45]:

$$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}}\Big|_\mathrm{meas} \simeq \frac{p_\mathrm{T} \cdot \sigma_\mathrm{meas}}{0.3\, L^2 B} \sqrt{\frac{720}{N+4}}, \tag{1.23}$$

[1] I.e. the component of the momentum transverse to the beam direction.

and for the multiple scattering component by the relation [20]:

$$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}}\Big|_{\mathrm{MS}} \simeq \frac{0.05}{\beta L B}\sqrt{\frac{Nx}{X_0 \sin\theta}}, \tag{1.24}$$

where p_T is given in GeV, B in T, and all the distances in m. The combuned resolution in p_T is then expressed as:

$$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}} = \sqrt{\left(\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}}\right)^2_{\mathrm{meas}} + \left(\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}}\right)^2_{\mathrm{MS}}} \equiv \sqrt{(a p_\mathrm{T})^2 + b^2} \tag{1.25}$$

Similar logic also stands for the impact parameter, and the contributions to the uncertainty read [20]:

$$\sigma_{d_0}\big|_{\mathrm{meas}} \propto \frac{\sigma_{\mathrm{meas}}}{\sqrt{N}} \quad \text{and} \quad \sigma_{d_0}\big|_{\mathrm{MS}} \propto \frac{1}{\beta p_\mathrm{T}}\sqrt{\frac{Nx}{X_0 \sin\theta}}. \tag{1.26}$$

At lower momenta, the uncertainties are dominated by scattering, whereas at higher momenta the coordinate uncertainty term is the main contribution (Eq. 1.25). The size L and magnetic field B of the tracking detector are often limited mechanically, and the number of planes has a counteracting effect on the two contributions to the resolution. Therefore for optimizing the resolution of the tracker it is universally beneficial to increase the performance of the individual detectors, i.e. to reduce coordinate uncertainty σ_{meas}, and the thickness x of each plane.

1.5.4 Radiation damage

A very important aspect of operating silicon detectors in an environment with high particle flux is *radiation damage*, i.e. accumulated defects in the silicon material, that affect the functioning of the detector. An ability of a device or a material to withstand radiation damage and to retain its properties is called *radiation hardness* or *radiation tolerance*.

Major effects and their implications on silicon detectors are presented in this section only qualitatively; a detailed review can be found e.g. in [46, 47].

Ionizing energy loss

Energy depositions in a form of creating electron-hole pairs are typically do not cause lasting damage to silicon. All the deposited charge is either collected in presence of the electric field, or recombined. Total amount of ionizing radiation, or *dose* is expressed in *Gray* (1 Gy = 1 J/kg). For the inner detector layers of the HL-LHC experiments (discussed in Section 5.1), the total ionizing dose is estimated to be of the order of 10 MGy.

The ionizing energy loss, however, can result in persistent damage to the Si-SiO_2 interface, which is commonly used as protection layer for the bulk, or in *metal-oxide-semiconductor field effect transistors (MOSFETs)*. Charge deposits at the interface may create trapped holes, leading to a buildup of static charge, that may render transistors non-functional. Thus, the readout electronics (either separate readout chips or embedded transistors in MAPS) are most affected by the ionizing energy loss. Radiation hardness of MOSFETs and other elements can be significantly increased by reducing their size, which is possible by using modern lithography technologies.

Non-ionizing energy loss

Another type of radiation damage to silicon is caused by *non-ionizing energy loss (NIEL)*, i.e. by quasi-elastic interactions of incoming particles with atoms in silicon bulk. Such interactions may create various impurities or irregularities (e.g. displaced, missing or excess nodes) in the lattice, called *defects*. Accumulation of defects may introduce additional energy levels for charge carriers, which, in turn, may alter macroscopic properties of silicon and thus affect functioning of the detector.

Non-ionizing dose is commonly measured in n_{eq}/cm^2 (equivalent 1 MeV neutrons per unit area). Although being caused by different particles of different energies, the damage can be scaled to that of reference particles (1 MeV neutrons). Even though this scaling (called NIEL scaling) is approximate, it is universally used. For the inner detector layers of the HL-LHC experiments, the total non-ionizing dose is estimated to be up to the order of 10^{16} n_{eq}/cm^2.

The main effects of non-ionizing radiation damage are:

- **Change of effective doping.** Defects may act as dopants, or neutralize existing doping, thus changing the effective doping concentration (defined by Eq. 1.14) towards the p-type. In a lowly-doped n-bulk a *type inversion* is possible: accumulated defects can turn the material into p-type. For p^+-in-n detectors this changes the depletion characteristics, and the p-n junction moves to the backside of the sensor.[1] If the applied bias voltage is not enough to fully deplete the bulk, the detector stops functioning, because the depletion region growing from the backside does not reach the collection implants. This can be mitigated by use of detector designs, that explicitly account for type inversion (e.g. n^+-in-n, as in Section 5.2.5), or use of sensors with p-type bulk, which is mostly unaffected (as in Section 5.3). The change of the effective doping concentration with received neutron fluence is shown in Figure 1.11.

- **Increased leakage current.** Bulk defects may increase the thermal generation of e-h pairs, leading to a significant increase of the leakage current. Leakage currents have to be kept low, as they increase the noise in the detector, and may also lead to a *thermal runaway*.[2] This is usually controlled by operating detectors at low temperatures (typically $-30\,^{\circ}C$).

- **Carrier trapping.** Defects known as trapping centers may reduce the mean lifetime of charge carriers in silicon, so that it becomes comparable to characteristic charge collection time. This leads to a reduction of the signal level, as some charge carriers are trapped or recombined, thus they contribute less to the induced current. This effect can be reduced by operating detectors at higher bias voltages,[3] thus resulting in faster charge collection.

[1] Since the backside implant remains n^+ due to much higher initial doping.

[2] Leakage current through the silicon bulk causes resistive heating in a positive feedback loop.

[3] However, increased leakage currents and possible breakdowns should be taken into account.

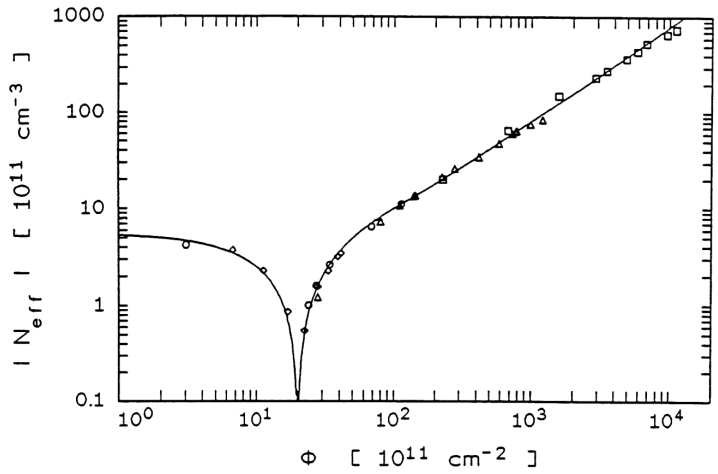


Figure 1.11: Absolute value of the effective doping concentration as a function of 1 MeV neutron dose. The wafer is initially n-doped at $N_d \simeq 5 \times 10^{11}/\text{cm}^3$. At the neutron fluence of $2 \times 10^{12}/\text{cm}^2$ the initial doping is compensated by p-type defects, and further irradiation keeps increasing the acceptor concentration N_a. Adapted from [48].

1.6 Development, production and deployment of silicon detectors

While silicon detectors show excellent characteristics in measurements of particle trajectories, major research efforts are usually required to create well-performing detector designs, which would in turn extend the physics reach of the experiments. The full development cycle for silicon detectors consists of multiple complex stages:

1. **Setting design goals.** Most of the developments of particle detectors are application-driven, e.g., detectors are developed to be employed in a specific experiment. Hence, requirements on the detector design originate from physical requirements of the foreseen experiment. Specifications of detectors themselves can be derived from the desired performance of the corresponding subsystem of the setup, e.g., of the tracker. Another less common goal setting strategy is the so-called *blue sky research*, which is a development of a certain technology in order to explore its limits and to estimate potential applications.

2. **Creation of a design.** Exact designs of silicon sensors should comply to known manufacturing limitations. The layout should be constructed such that it can be produced with the steps and processes, offered by the lithography industry. For silicon sensor designs, important aspects are availability of different materials of certain purities and possible complexity of implant arrangements.

3. **Verification of the design: simulations.** Modern software packages enable complete simulations of the manufacturing steps and of the performance of the final product (as explained in Chapter 3). The simulation studies allow to optimize the design on early stages, potentially reducing the required amount of costly prototyping iterations.

4. **Prototyping.** Given the manufacturing technique, the prototype production requires an equally high amount of effort as for setting up the full scale production, because the same

preparation of lithography masks is required regardless of the amount of silicon chips to be produced. Cost optimization strategies may take place, such as production of multiple prototype variations at once, in order to keep the number of production iterations as low as possible.

5. **Characterization of the prototypes.** Proper functioning of the detector and its performance have to be verified through extensive testing. This includes studies of the front-end and readout electronics in a laboratory environment, as well as tests of the particle detection performance under close-to-real use conditions, e.g. in a test beam environment (as presented Chapter 7).

6. **Reiteration.** If testing shows that all the specifications were met and no issues were discovered, the design is deemed complete and ready for use. Otherwise, the design has to be altered and steps 2–5 are then repeated.

The further steps only concern designs meant for large-scale deployment, e.g. CMS Outer Tracker, presented in Part II of this thesis.

7. **Mass production.** Whereas manufacturing of sensors themselves is easily scalable due to nature of the lithography technology, other steps, such as assembly of the readout and auxiliary structures, need to be adapted to facilitate reliable and consistent production of large batches of detectors (up to several thousand units in large-scale experiments).

8. **Quality control.** To ensure proper performance of the future experimental setup, every single detector unit in it has to meet the specifications. To uncover potential production issues and account for variance of sensor manufacturing, that is intrinsic to the lithography process, the separate components, sub-assemblies and complete detector modules are put to test. The testing procedure may be similar to characterization, but simplified in order to allow for high throughput.

9. **Deployment and commission.** The complete system is being installed at the experiment. Given the high complexity of setups, this poses engineering challenges. E.g., dedicated *integration* studies might be required to ensure proper functionality of the detector system as a part of a bigger setup.

1.7 Scope of this work

Given the nature of silicon particle detectors, their development involves significant efforts on prototype testing and characterizations, i.e. detailed studies of the devices' performance. In turn, such studies themselves often necessitate dedicated testing environments, specialized methods, and unique solutions. This thesis discusses several aspects of silicon detector development, including advancements of the research and development (R&D) methods and testing techniques, as well as applying them to build a large-scale detector that would be commissioned in the near future. By presenting projects at different stages of the development cycle, this thesis provides a holistic picture of characterization measurements as a part of the R&D and production chain.

Part I of this work is devoted to testing silicon sensors with pulsed lasers. This flexible and powerful technique enables studies of various properties of silicon sensors in great detail, which often yields valuable insights during early stages of sensor development. Implementation of the technique is discussed, as well as several aspects regarding simulating this process with Monte-Carlo methods, which aim to improve understanding of experimental outcomes. Additionally, a laser-based characterization of the digital SiPM [49], a novel silicon pixel photo-detector, is presented. Chapter 2 introduces the technique itself and the *Laserbox*, a laser setup which was commissioned by the author and used in this work. Chapter 3 presents Monte Carlo simulation approach of the technique with the Allpix^2 software, developed in the course of this work. Chapter 4 presents the characterization of the dSiPM chip.

Part II of this work is devoted to particular aspects of the upgrade for the Outer Tracker [50] of the Compact Muon Solenoid (CMS) detector [13] at the LHC. This part addresses challenges of a large-scale detector project at a late R&D stage. Production techniques for novel double-sensor silicon detector modules, improved and optimized throughout this work, are presented. Then, a qualification campaign for such modules at the DESY II test beam facility, performed during this doctoral project, is reported. The test beam environment allows one to test detectors under conditions, close to those in a particle physics experiment. Chapter 5 presents the CMS experiment itself and its upgrade. Details of the assembly and test beam characterization of silicon detector modules for the Outer Tracker upgrade are covered in Chapters 6 and 7 respectively.

The work accomplished in the course of the doctoral project and presented herein contributes to the field in manifold aspects:

- commissioning of a flexible multi-purpose laser injection test stand at DESY Hamburg, together with a validated simulation tool;
- proof of principle for a pixelated single-photon detector with picosecond timing capabilities, produced with CMOS technology, and insights on its temporal resolution;
- significant improvements of the assembly strategy for novel double-sensor modules, facilitating a timely and precise production of detector modules for the CMS experiment upgrade;
- confirmation of production readiness of the module design.

Part I

Characterizing silicon devices with pulsed lasers

Chapter 2

Injection lasers for sensor R&D

In this chapter, characterization methods for silicon sensors that employ charge injection with a pulsed laser are presented. Section 2.1 is devoted to basic principles and relevant experimental aspects of the Transient Current Technique. Section 2.2 introduces the *Laserbox*, a laser injection setup at DESY, commissioned and maintained by the author and then used for measurements, presented in this work.

2.1 Transient Current Technique

The *Transient Current Technique (TCT)* is a capable method to study several properties of silicon sensors. The main idea behind this technique lies in the analysis of shapes of electric pulses, induced by the drift of charge carriers through the sensor bulk. The pulse shape, according to the Shockley-Ramo theorem (Eq. 1.22), encodes information about the electric field, the weighting potential, and carrier lifetime in the sensor. Thus, bulk properties can be inferred for different regions within the sensor if the initial spatial distribution of carriers can be created in a controlled way. A commonly used method for controlled charge injection is illumination with pulsed lasers due to the flexibility of this technique. The electric pulses from sensors are typically picked up with low-distortion readout chains with high-bandwidth (GHz and higher) amplifiers, as well as digitizers with high sampling rates (e.g digital oscilloscopes).

2.1.1 Overview of charge injection techniques

In order to properly characterize a particle detector, one needs to verify that it produces a proper response to impinging particles. In case of segmented silicon detectors, it is often required to test every channel of the detector. Moreover, because of the specific structure of silicon sensors, their response might be different for different interaction locations within a single channel (pixel or strip). Given that ionization energy loss with electron-hole pair creation is the major mechanism of interaction of charged particles with the silicon bulk, to mimic the operation conditions one needs to create a certain set of charge carriers at a certain location in the sensor. Common experimental techniques for this purpose include:

- **Test beam** (discussed in detail in Chapter 7). The sensor is exposed to a particle beam delivered by an accelerator, and for each of the beam particles, the trajectory is reconstructed by an independent reference detector system. Test beams provide an environment which is

very similar to that of a particle physics experiment, since the detector is exposed to actual high-energy ionizing particles. Being rather complicated setups, test beams are of limited availability.

- **Radioactive sources.** Another technique to induce ionization is using radioactive sources. However, due to random nature of radioactive decays, they are not an optimal choice for detector testing, which often requires knowledge on the exact coordinates and timestamps of incident particles. Another limitation of using radioactive sources is the rate and energy of particles emitted.

- **Test pulsing.** Some detector front-ends are equipped with dedicated test circuits, which allow to inject electric pulses with known amplitudes in the readout chain. This allows to test the functionality of the readout electronics *independently*[1] of the silicon sensor.

- **Light-induced injection.** Electron-hole pairs can be created in a silicon bulk via photoelectric effect by exposing the detector to a light source. A common approach is the use of focused pulsed lasers, with its advantages and disadvantages being discussed in the following. A key aspect of this technique is that the exact injection position and time can be precisely controlled, so very localized probing can be performed. Additionally, this technique is of great use for studies of timing characteristics of sensors (as presented in Chapter 4). Also, the use of non-coherent diffuse light is possible [51].

2.1.2 Charge injection with pulsed lasers

Mechanisms of particle interaction with matter are significantly different for charged particles and photons. However, photons can be used to create an ionization pattern very similar to the one from ionizing particles. This might be achieved with the help of pulsed laser beams of specific wavelengths (see Figure 1.8) and focusing optics. In the following paragraphs, the key parameters and aspects of laser injection systems are discussed.

Spatial profile along the beam

The passage of minimally ionizing particles, that results in a uniform line pattern, can be mimicked by exposing the sensor to an infrared (IR) laser beam ($\lambda \simeq 1050\,\text{nm}$). For these wavelengths, the absorption depth in silicon is of the order of millimeters. For silicon sensors that are typically few hundred micrometer thick, the beam attenuation (Eq. 1.11) can be neglected. Thus, charge deposition along the beam can be considered uniform. Passage of a "slow" particle (i.e. $\beta\gamma \ll 1$), that would deposit all its energy over a short distance as soon as it enters the silicon bulk (as defined by Eq. 1.6) can be mimicked with a red laser ($\lambda \simeq 650\,\text{nm}$). With an absorption depth of a few µm, all the charge will be deposited in a surface layer in a similar way.

Spatial profile across the beam

To ensure that charges are deposited in the silicon bulk as a line or a spot, the laser should be focused to a narrow beam. The shape of the beam produced by focusing lenses can be approxi-

[1] Not entirely, as there are e.g. effects of the pixel capacitance.

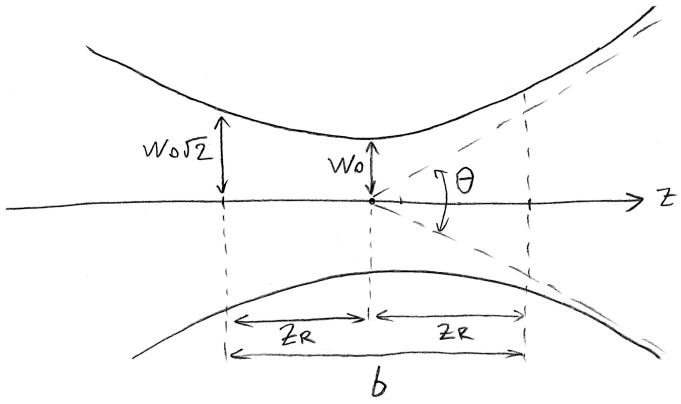


Figure 2.1: Longitudinal profile of a Gaussian beam.

mated with a *Gaussian beam* model [52]. The intensity distribution in the beam parallel to the z-axis with wavelength λ in a medium with refractive index n, as a function of coordinates, reads:

$$I(x,y,z) = I_0 \cdot \left(\frac{w_0}{w(z)}\right)^2 \cdot \exp\left(-2\,\frac{x^2+y^2}{w(z)^2}\right), \tag{2.1}$$

where:

- I_0 is the intensity at the center of the focal plane ($x = y = z = 0$);
- w_0, the *beam waist*, is a parameter characterizing transversal size of the beam in the focal plane; in practice, w_0 is constrained by the properties of the laser source and optics;
- w is the beam radius at a coordinate z, defined as:

$$w(z) = w_0\sqrt{1+\left(\frac{z}{z_R}\right)^2}; \tag{2.2}$$

  for every cross-section, $1-1/e^2 \simeq 86\%$ of the beam energy is contained within a circle $x^2 + y^2 \leqslant w^2$;
- $z_R = \pi n w_0^2/\lambda$ is the *Rayleigh length*; in a region $-z_R \leqslant z \leqslant z_R$ the beam radius does not exceed $w_0\sqrt{2}$.

The Gaussian beam with the intensity described by Eq. 2.1 has a converging-diverging shape, shown in Figure 2.1. To ensure line-like charge deposition, a part of the beam around the focal point is used, and the beam waist w_0 should be minimized. Decreasing w_0, however, decreases the Rayleigh length and increases the beam divergence $\theta \simeq 2\,w_0/z_R$, thus decreasing the length of the "usable" beam segment, called *confocal parameter* $b = 2\,z_R$. This effect is partially mitigated by the fact that the beam shape is longitudinally expanded as the beam enters the silicon bulk due to the change of the refractive index (for visible light $n_{\mathrm{Si}} \simeq 3.5 - 4$). For an IR laser setup (λ = 1064 nm, $w_0 \simeq$10 µm) the confocal parameter b is approximately 300 µm in air and 2 mm in silicon.

Temporal profile of the laser pulse

At the time scale of charge collection in silicon sensors ($\sim$ ns), a passage of a relativistic particle and the corresponding ionization can be considered of infinitely short duration, whereas common laser drivers produce pulses of a non-negligible length, that should be accounted for in signal formation studies. The measured pulse shape is given as a convolution of the induced current i_{sensor} with the laser pulse shape L and front-end transfer function T:

$$i_{\text{measured}}(t) = \int \mathrm{d}\tau_1 \int \mathrm{d}\tau_2 \cdot i_{\text{sensor}}(\tau_2) \cdot L(\tau_2 - \tau_1) \cdot T(\tau_1 - t). \tag{2.3}$$

Two types of laser drivers with various pulse lengths were used in this work (presented in detail in Section 2.2.1). One of them has a pulse duration below 50 ps, which is still considered negligible w.r.t. signal formation time in silicon sensors. The other type with 1–2 ns long pulses has a disadvantage of smearing of electric signals, picked up from the silicon sensor, due to its non-negligible pulse duration.

Access to the bulk

The main limitation of this technique is the fact that the sensor surface needs to be directly exposed to the laser beam. Unlike highly energetic charged particles, visible light photons are unable to penetrate metal layers and other auxiliary structures, which are commonly present on front- and backside of sensors.

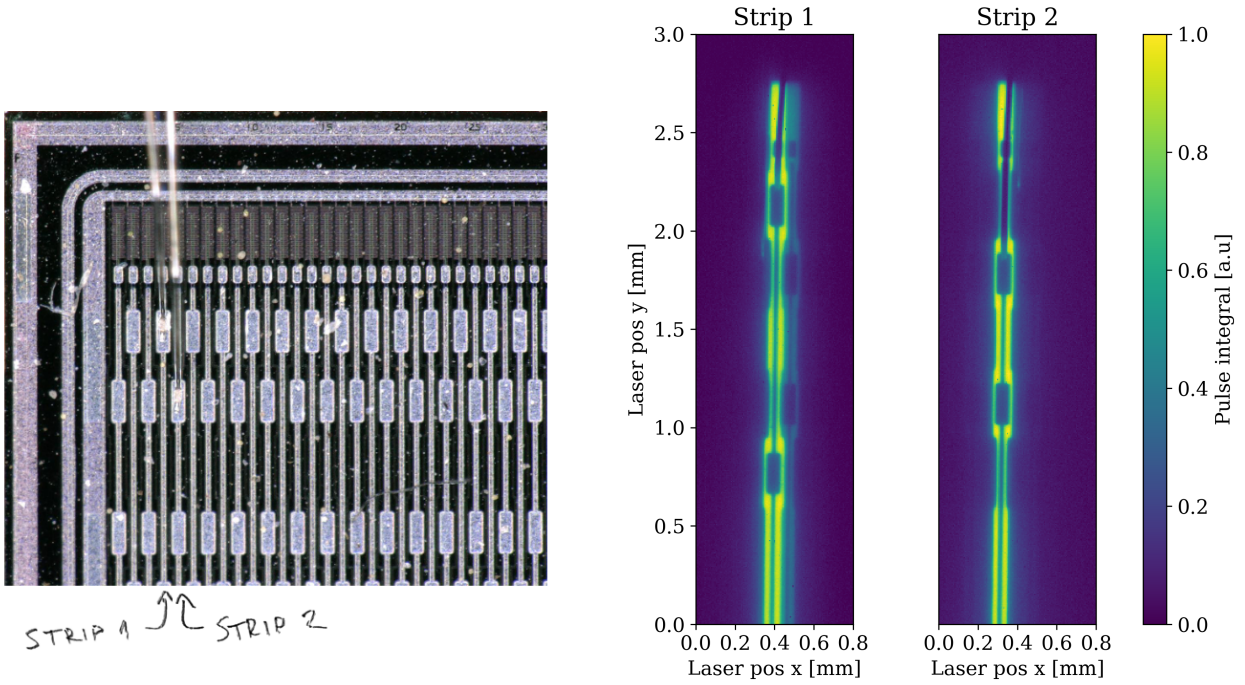
An example is shown in Figure 2.2. A test strip sensor (ATLAS12 mini-sensor [53], shown in Figure 2.2a) was illuminated with a focused IR laser in the Laserbox setup (introduced in Section 2.2), and signals in two adjacent strips were recorded. Figure 2.2b shows the integral of the electric pulse as a function of the xy-coordinate of the laser spot. Whereas the sensor itself is designed to be almost uniformly efficient across its surface, illuminating certain regions of it with the laser yields no signal as the beam is absorbed or reflected by metal traces and connecting wires.

Internal reflection

Upon the passage through the silicon bulk, the beam is subject to *internal reflection* on the silicon-air interface or on the backside metalization. The reflected part, passing through the bulk in opposite direction, also creates charge carries. With special focusing used, it is possible to use the internally reflected beam to deposit charges in regions inaccessible otherwise, e.g. right under opaque metal pads (not employed in this work).

Laser

To create m.i.p.-like charge deposition with an IR laser, the beam intensity should be tuned in such a way that the amount of charge created by the laser resembles the mean charge deposited by an ionizing particle passage ($\sim$ 100 pairs/µm). In order to numerically estimate the laser-induced ionization, the precise knowledge of the intensity and shape of the beam is required. These could be either derived from characteristics of the laser system, if those are known precisely, or measured with a dedicated photometric device. Additionally, one needs to account for transitivity and reflectivity of the air-silicon interface, as well as for light refraction. These effects significantly depend on the surface finish of the silicon sensor. Alternatively, the laser intensity can be tuned based on



(a) Micro-photograph of the ATLAS12 mini-sensor.

(b) Total induced signal in two adjacent strips as a function of the injection coordinates.

Figure 2.2: Metal traces blocking light access to the sensor bulk.

a prior knowledge of the amplitude of a m.i.p. signal in the sensor under test (or a similar one). This information can be obtained e.g. by using a radioactive source, from cosmic ray data, or at a test beam. In general, the laser intensity adjustment is very application-specific.

2.1.3 Spatial configurations of charge injection

By varying the laser wavelength and the spatial configuration of a TCT experiment, one can create different scenarios of charge transport in the sensor and thus study its specific properties.

Red-TCT

Use of red lasers allows to create initial charge carrier distributions such that only one type of charge carries contributes to signal formation. Due to its low penetration depth, red laser light only creates carriers at the surface, in an immediate proximity of the implants. The beam can be aimed at an arbitrary spot on the sensor surface, allowing to resolve bulk features at various locations. For an n^+-in-p sensor, the following stands:

- If the sensor is illuminated from the frontside (Figure 2.3a), electrons are absorbed by the readout implants almost instantaneously. Due to short drift paths, electrons only make a minor contribution to the induced pulse. Holes are repulsed from the readout implants and drift through the entire sensor bulk, going through a gradient of weighting potential and

thus forming a dominant fraction of the pulse, allowing to study specific properties of holes, e.g. mobility and lifetime;

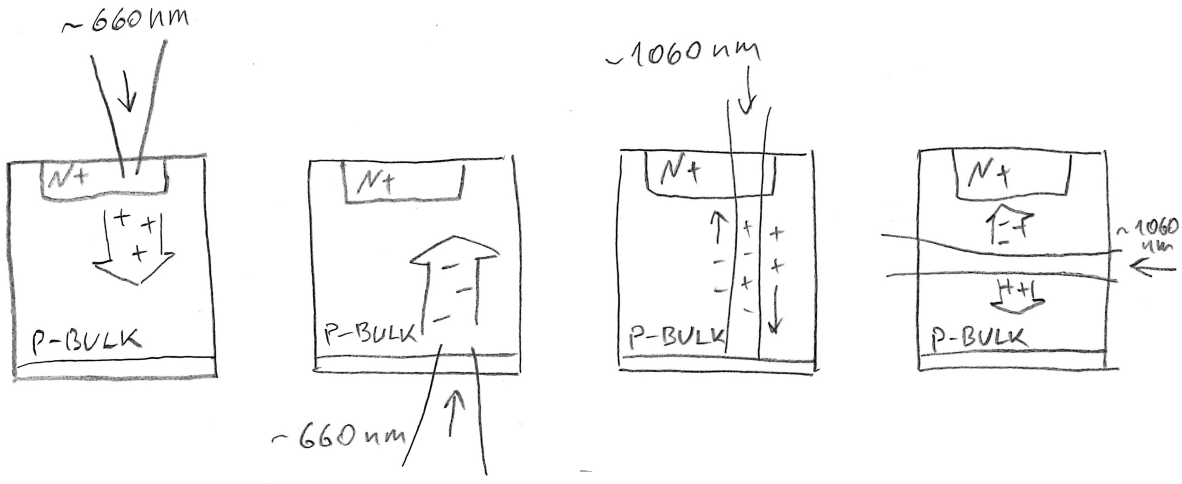
- Similarly, under the backside illumination (Figure 2.3b), holes are instantly absorbed by the backside implant and do not significantly contribute to the signal. The signal then is induced almost exclusively by electrons that drift through the whole bulk from the backside to the frontside.

Infrared-TCT

The deposition of charge by an infrared laser can be assumed uniform along the beam throughout the entire bulk thickness (Figure 2.3c). Thus, there is no significant difference between frontside and backside illumination, and either sensor side can be used. This shape of charge deposition is indispensable for qualification of the sensor response to m.i.p.s. Similarly to the red-TCT, the beam can be aimed at a specific spot on the sensor plane to study certain regions of the sensor.

Edge-TCT

An infrared laser can be placed such that the beam is parallel to the sensor plane and enters the sensor from the edge (Figure 2.3d). This requires special treatment of the sensor edge (i.e. polishing) to prevent light diffusion on the silicon-air interface. This configuration allows for a charge injection at a specific z-coordinate in the sensor, thus facilitating studies of certain features of the bulk as a function of depth [54]. The beam entering from the side, however, only provides resolution along one coordinate across the sensor, averaging along the other.



(a) Red – frontside. **(b)** Red – backside. **(c)** Infrared – top. **(d)** Infrared – edge.

Figure 2.3: Spatial configurations of TCT measurements. An n^+-in-p sensor is shown as an example.

Two-photon absorption

An advanced laser injection technique using a *two-photon absorption (2PA)* effect [55, 56] can be employed to constrain charge deposition in all three coordinates. If a laser with a wavelength such that $E_G/2 < hc/\lambda < E_G$ is used, where E_G is the band gap, absorption of photons via regular photoeffect is suppressed. However, photo-absorption by the lattice is still possible, if the combined energy of two photons is transferred *simultaneously* to an electron in the valence band. The probability of this effect scales as beam intensity squared, thus, in a focused Gaussian beam it would predominantly occur in a small region around the focal point. Hence, charge deposition can be localized in an area as small as a single µm transversely and a few µm longitudinally. The ability to inject charge in any given location in the silicon bulk allows to qualitatively study spatial distributions of e.g. electric field in it. Experimental implementations of such a technique require the use of lasers of high intensity and extremely capable optical systems, thus to date 2PA injection setups are very rare.

2.1.4 Extractable observables

Via TCT, the following parameters of charge collection and bulk properties in silicon sensors can be studied by analyzing the induced pulse shape $i(t)$:

Detector sensitivity

By using an infrared laser and adjusting its intensity one may create charge deposition very similar to that of an m.i.p. One can then test charge collection in the silicon device, basically verifying that a m.i.p.-like charge deposition creates pulses of a measurable amplitude on the readout electrodes. The *sensor efficiency* (i.e. the probability to detect a particle) can then be inferred by taking the properties of the front-end circuits into account.

This technique is particularly useful for segmented detectors to study the pulse shape and the resulting efficiency as a function the charge deposition coordinate w.r.t. readout implants.

Depletion voltage

The full depletion voltage V_{depl} can be measured via IR-TCT by analyzing the dependency of the total collected charge $Q = \int_0^T i(t)\mathrm{d}t$ as a function of bias voltage V_{bias}. Only the fraction of charge that was deposited in the depleted zone will contribute to the signal.[1] The total induced charge Q then increases $\propto \sqrt{V_{\text{bias}}}$ (as given by Eq. 1.15) and is saturated at $V_{\text{bias}} = V_{\text{depl}}$.

Profile of the electric field

The dependency of the electric field amplitude on the depth z in the sensor can be studied with red-TCT. The initial charge deposition is almost point-like, therefore the shape of the pulse, which is produced as the group of charge carriers drifts through the bulk, would resemble the

[1] Although charge that is deposited in the undepleted region can also contribute to the signal by diffusing to the depleted region, the impact of this effect can be minimized by adjusting the upper integration limit T when calculating the total charge.

$E(z)$ profile. Whereas the exact dependency of the instantaneous current includes the weighting potential and carrier mobility, qualitatively, the increasing pulse current implies that the electric field amplitude increases along the charge cloud trajectory and vice versa.

Such a technique is useful to test radiation-damaged p^+-in-n sensors for type inversion (introduced in Section 1.5.4). In a non-irradiated sensor, the amplitude of the electric field decreases from the frontside to the back, and a frontside injection creates a gradually decreasing pulse. In contrast, in a sensor with full type inversion, the p-n junction moves to the backside, the electric field then increases towards the backside and the current pulses have a positive slope [57]. Certain types of radiation damage can also result in a so-called *double junction* structure with higher field amplitude near implants and lower amplitude in the middle. In this case, pulses of a characteristic "saddle" shape would be observed [58].

Charge carrier lifetime

In radiation-damaged silicon sensors, the charge carrier lifetime τ can be comparable to the signal formation time due to the carrier trapping. This can be accounted for in the Shockley-Ramo theorem (Eq. 1.22) by introducing time-dependent decay of the free charge in the sensor: $q = q(t) = q_0\, e^{-t/\tau}$. With TCT, the lifetime τ can be measured with the following technique [59]:

- Red-TCT is used, so that only one type of charge carriers contributes to the signal;
- pulse shapes $i_1(t)$ and $i_2(t)$ are recorded at two different bias voltages V_1 and V_2 ($V_1 > V_2 > V_{\text{depl}}$).
- in the absence of carrier trapping (infinite lifetime), the total induced charge in both cases would be the same; in presence of carrier trapping $Q_1 > Q_2$ since at higher bias voltage charge is collected faster and smaller fraction of it is trapped or recombined;
- the parameter τ can be obtained as a solution to the equation:

$$\int_0^T i_1(t)\, e^{-t/\tau} \mathrm{d}t = \int_0^T i_2(t)\, e^{-t/\tau} \mathrm{d}t. \tag{2.4}$$

2.2 The Laserbox setup

An experimental setup to perform TCT measurements typically consists of a laser with an optical system, mechanical positioners that allow to adjust and aim the beam, and instruments to operate the device under test (DUT) and read it out. The *Laserbox* setup at DESY used in this work is based on a complete hardware solution, produced by Particulars [59]. The scheme of the setup with its main components is shown in Figure 2.4. The inner part of the setup is shown in Figure 2.5. An extensive overhaul of the setup was performed by the author, including the development of a new control software with capabilities of automation and remote control, integration of the Caribou readout system [60], and upgrade of multiple components.

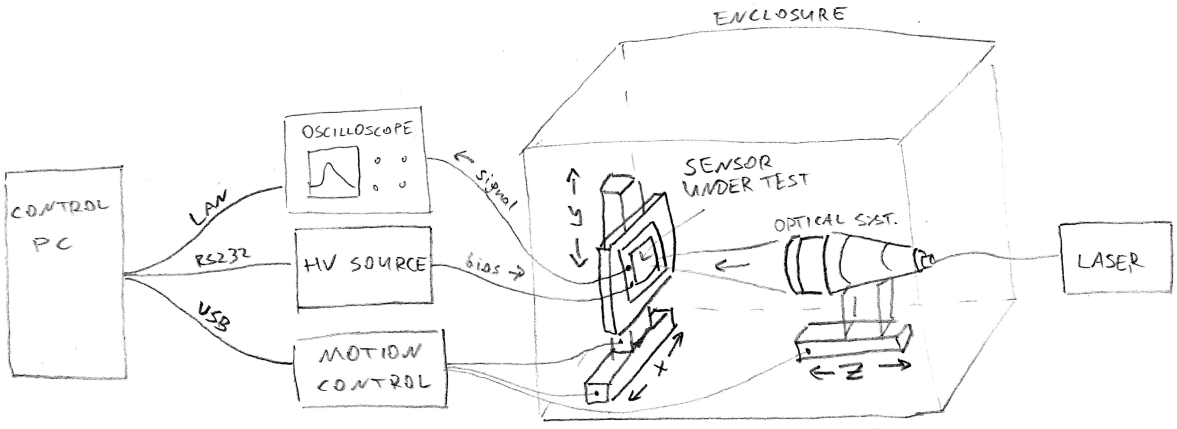


Figure 2.4: Scheme of the Laserbox, configured for measurements on passive silicon sensors. Auxiliary components are not shown.

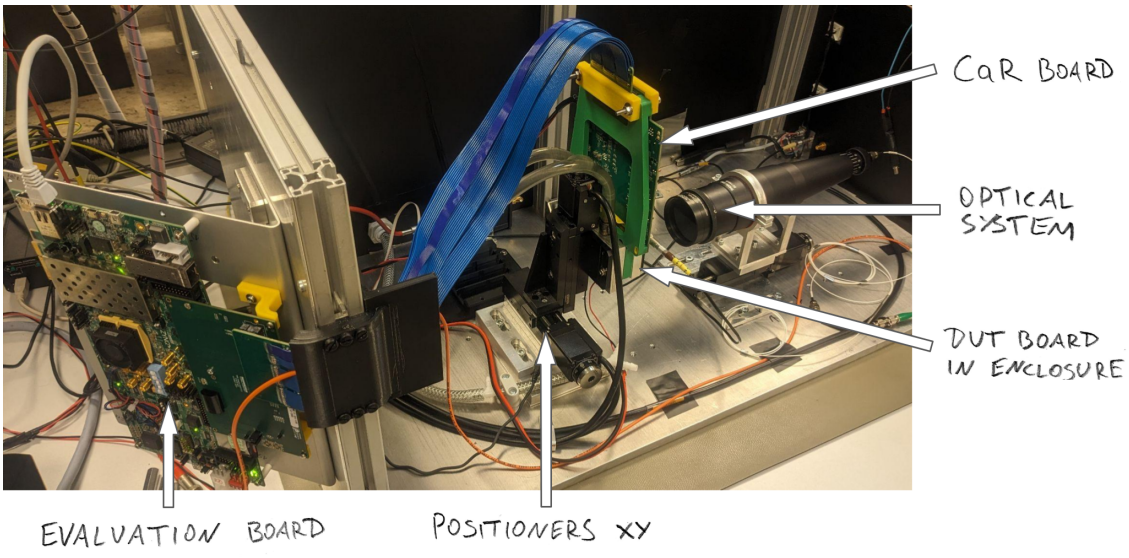


Figure 2.5: Inner part of the Laserbox, configured for characterization of a Caribou-based device (as presented in Section 2.2.1 and Chapter 4).

2.2.1 Setup components

Lasers and optics

Four different diode fiber-coupled lasers are available in the setup and are listed in Table 2.1. These lasers are specialized models, suited for charge injection into silicon devices. They operate in a pulse mode and can be triggered either by an external signal or at a constant frequency and can produce pulses at a rate of up to 1 MHz. The stock laser system by Particulars [61] generates pulses of adjustable duration, although at the lowest setting it still exceeds 1 ns, which is comparable with characteristic charge collection time in planar silicon sensors ($\mathcal{O}(1\,\text{ns})$). Another laser system, ALS PiLas [62], was installed, capable of providing significantly shorter pulses.

Table 2.1: Manufacturer specifications of laser units.

Name	Wavelength [nm]	Peak power [mW]	Pulse duration
Particulars LA-01 R FC	660	100	∼1.5 ns
Particulars LA-01 IR FC	1064	100	∼1.5 ns
ALS PIL067SM	672	900	45 ps
ALS PIL106SM	1063	700	40 ps

Lasers are focused with an optical system, shown schematically in Figure 2.6. A Gaussian beam is created by spreading the beam with an adjustable expander [63] and focusing it with a lens. The beam can be truncated with an adjustable iris and attenuated with a set of neutral density (ND) filters. The characterization of the beam obtained with this system is presented in Section 2.2.2.

The laser pulse can be monitored with dedicated reference photo-diodes [64, 65]. The fiber output of the laser driver can be split between the optical system and the reference diode, which is read out directly with an oscilloscope. The reference diodes provide sufficient speed to measure the pulse duration of Particulars lasers pulses, but not of the ALS system. Nevertheless, the sharp rising edge of the pulse from reference diode allows to quantify the *jitter*[1] of the ALS laser, which is crucial for time measurements. The jitter was found to match the claimed <5 ps specification.

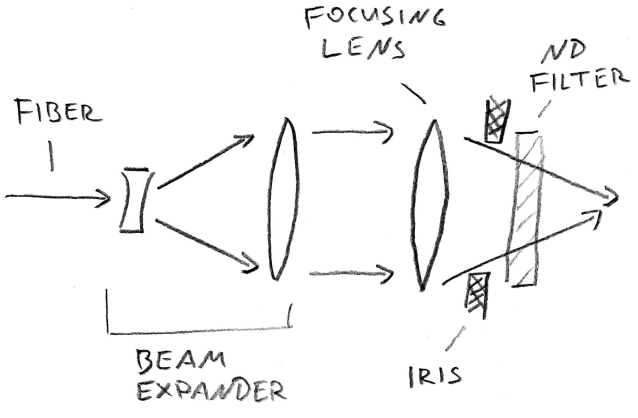


Figure 2.6: Scheme of the optical system in the Laserbox.

[1] Variance of the delay between the trigger signal for the laser and the start of the actual laser pulse.

Mechanics and utilities

Spatial adjustment of the laser beam is performed with motorized translation stages [66], that provide sub-µm positioning precision. The entire lens can be moved in z (along the beam) for focusing. The DUT is mounted on the positioners that move it in xy (across the beam).

The optical system, the positioners and the DUT are enclosed in a metal box, that offers protection from ambient light and EM-interference. The enclosure also provides laser safety for the operator; it is equipped with an interlock that prevents the laser[1] from switching on if the box is open. The DUT can be thermally stabilized with a two-stage cooling system, consisting of a Peltier element and a liquid chiller. The setup can also be flushed with conditioned dry air to prevent water condensation.

Passive DUT integration and readout

Passive sensors, i.e. not equipped with any front-end, are installed into the setup using dedicated DESY-developed circuit boards (see Figure 2.7). A DUT is clamped down or glued on a contact pad with conductive glue, which establishes contact with the sensor backside for grounding or biasing. Contact pads on the frontside are wirebonded to the traces on the board. External instruments can then be connected to the board.

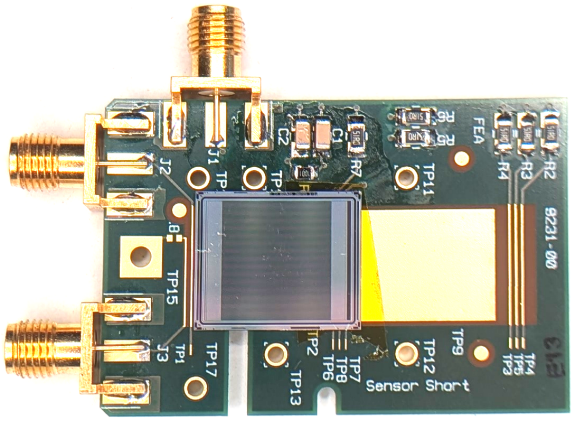


Figure 2.7: Circuit board for mounting a passive sensor in the Laserbox. Here, a baby-PS-s sensor (shown in Section 3.2.1) is attached.

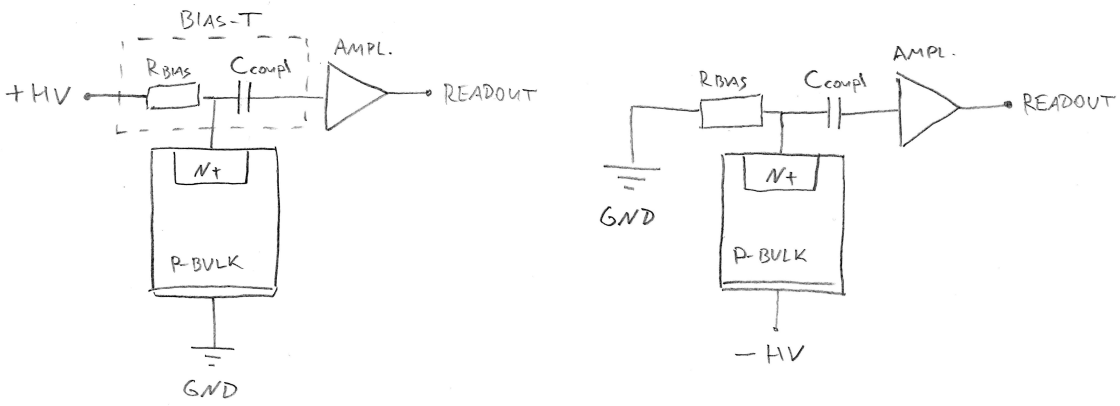
Schemes for connecting passive silicon sensors, used in this work, are as follows:

- The connection of an n^+-in-p diode (used for measurements, presented in Section 2.2.2) is shown in Figure 2.8a. The sensor backside is grounded. At the frontside, one should decouple the signal readout from the high voltage; this is done via a circuit called *Bias-T* [67]. The positive high voltage[2] is supplied through a biasing resistor R_{bias}, and the readout is connected through a coupling capacitor C_{coupl}, which isolates it from the high voltage.

[1] Only for Particulars lasers that have higher *sustained* power.
[2] Such that the n^+-in-p device is reverse-biased.

- The connection of an n^+-in-p AC-coupled strip sensor (used for measurements, presented in Chapter 3) is shown in Figure 2.8b. Negative bias voltage is supplied to the backside. A biasing resistor and a coupling capacitor for each strip are embedded directly into the sensor. The grounding and the readout are connected to dedicated contact pads on the sensor frontside.

The signals from the sensor are picked up with high-bandwidth amplifiers [68] and recorded with a digital storage oscilloscope [69]. Two independent amplification and readout channels are available, e.g. for the readout of two strips on a strip sensor.



(a) Diode with frontside biasing via Bias-T. **(b)** AC-coupled strip sensor with backside biasing.

Figure 2.8: Connection schemes for different passive silicon sensors.

Integration of Caribou readout electronics

Detectors with a dedicated front-end (either monolithic or with separate ASICs) and thus providing digital outputs require certain electronics to be communicated with. One of the solutions to provide data acquisition and control for such detectors is the Caribou platform [60]. Thanks to its versatility, Caribou can be adapted to operate with different detector prototypes. The Caribou system consists of three main hardware parts:

- The evaluation board with a system-on-chip and a field-programmable gate array (FPGA), based on a commercially available solution. The evaluation board runs the acquisition software *Peary* [70] under Linux;

- The control-and-readout (CaR) board hosts an arsenal of various electronic components to handle the detector chips, such as data links, voltage/current sources, analog-to-digital converters, and alike. The CaR board is connected to the evaluation board;

- The chipboard hosts the DUT and thus is usually designed uniquely for each DUT type. The chipboard is meant to have as little active electronic components as possible, since its only purpose is to connect the DUT to the services on the CaR board.[1]

[1] And optionally to some external instruments, such as high voltage supplies and oscilloscopes.

The Laserbox is equipped with a set of mounts to install the Caribou system in it (see Figure 2.5). The CaR board with the chipboard are installed on the xy-positioners, and the evaluation board is attached outside of the enclosure, being connected to the CaR board via 320-pin flexible bus.

During TCT or laser injection studies on such devices, the Caribou system would manage the DUT, including controls and data acquisition. Pulse shape measurements with an oscilloscope are still available if dedicated analog outputs are provided on the chipboard.

Control software

A new software application [71] was developed by the author in order to enable automated control and data acquisition with the Laserbox. This software, written in Python with Qt, issues commands to and receives monitoring data from the positioners and the high voltage source. It is also capable of retrieving and storing waveforms from the digital oscilloscope. The components can be controlled manually via graphical user interface (see Figure 2.9), or with built-in scanning routines, or remotely via a dedicated Python client. To enable remote control of Caribou-based devices, the user can design a Python script that issues a sequence of commands to both the Laserbox and Caribou.[1]

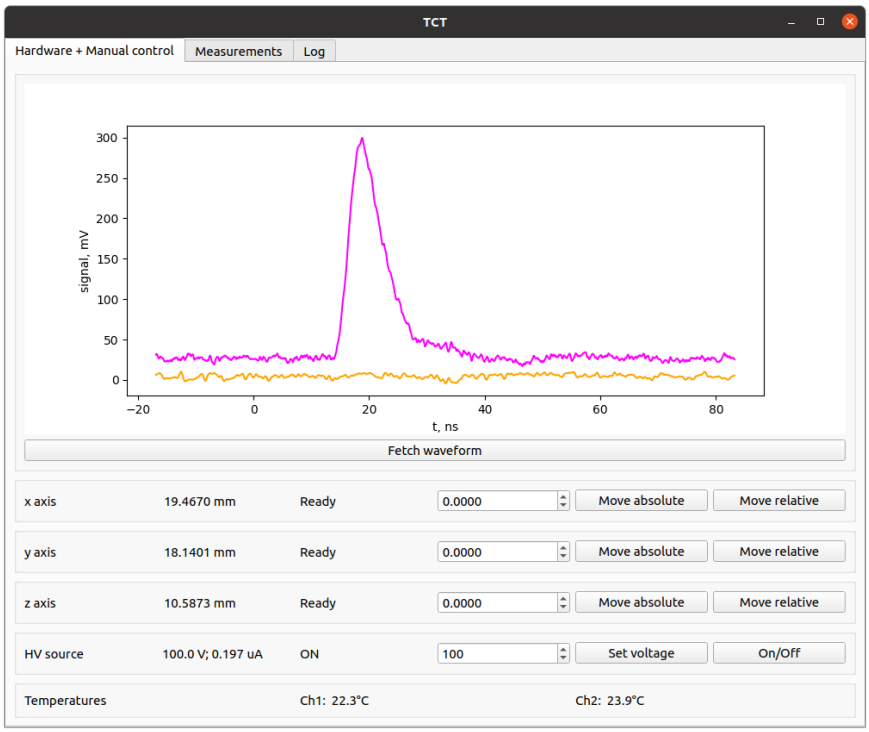


Figure 2.9: Control application for the Laserbox. The view with hardware monitoring and controls is presented. Signals, acquired from two adjacent strips of a sensor, are shown; the laser is set such that it deposits charge only in one strip (purple curve), so no signal is registered from the other strip (yellow curve).

[1] Peary provides similar functionality for remote control via Python scripts.

2.2.2 Beam spot characterization

In order to ensure localized charge deposition, the optical system should be tuned to produce the spot of a finite known size. To achieve that, the laser focus can be set and characterized with a *knife-edge* technique. Here, the Particulars IR laser was used. During such a procedure, the laser is aimed at a sensor, which has a sharp transition between the sensitive and non-sensitive regions. An n^+-in-p silicon diode was used. The surface of the diode is metalized, except for a circular opening of 1 mm in diameter (see Figure 2.10).

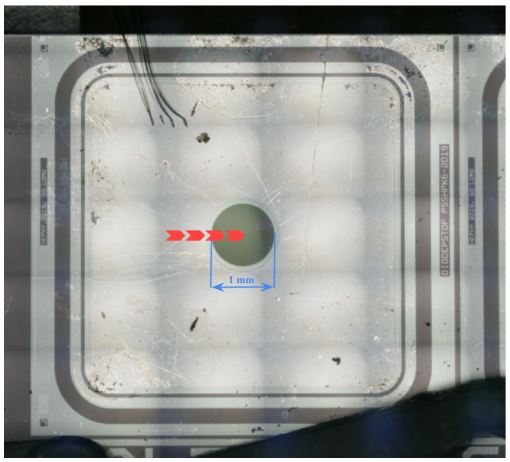


Figure 2.10: Microscope picture of the test diode. The scanning path is shown with red arrows.

The beam spot is then moved in x, across the edge of the opening over a distance of 1 mm[1]. For every location in x, the induced pulse amplitude is measured. The obtained dependency of the amplitude on x (see Figure 2.11a) is then fitted with an error function with four parameters: position x_0, magnitudes of left and right parts $y_{L,R}$, and width σ:

$$f(x) = y_L + (y_R - y_L) \cdot \mathrm{erf}\left(\frac{x - x_0}{\sigma\sqrt{2}}\right). \tag{2.5}$$

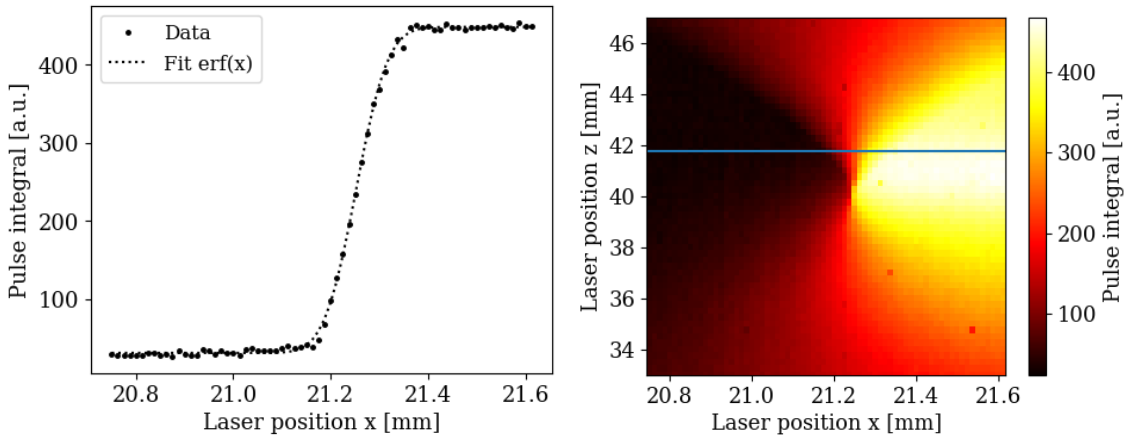
Given that the signal amplitude is proportional to the integral of the beam intensity over the sensitive diode area, from Eq. 2.1 follows:

$$w = 2\,\sigma. \tag{2.6}$$

By repeating this procedure for various z values (as shown in Figure 2.11b), the longitudinal beam profile $w(z)$ and its minimum w_0 can be obtained. This scanning procedure can be automated, as the Laserbox control software provides such capabilities.

Measurements of the longitudinal beam profile are shown in Figure 2.12 for various settings of the optical system and are used to optimize the optical system. Truncating the beam by closing the iris reduces the beam convergence, resulting in a nearly cylindrical beam, as indicated by its constant width. Adjusting the beam expander changes the beam waist and also shifts the focus in z. The optimal result is achieved with the expander set to the maximum magnification. The parameters of the beam in the optimal case can be obtained by fitting the measured profile using

[1]Since $w \ll R_{\text{opening}}$, the curvature of the opening border can be neglected.



(a) Pulse integral as a function of laser spot position x, as obtained in the scan. The data (solid markers) are shown together with the fit (dotted line), which yields $\sigma = 103\,\mu m$.

(b) Pulse integral as a function of laser spot position x and lens position z, shown with color. The blue line indicates the profile, shown in Fig. 2.11a

Figure 2.11: Principle of beam waist characterization with a knife-edge technique, performed with a test diode.

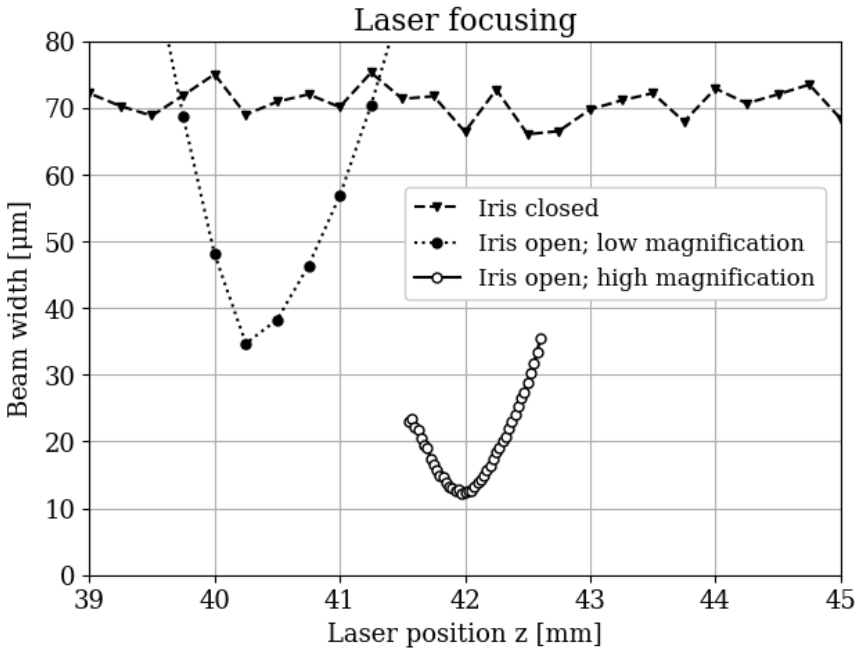


Figure 2.12: Beam profile for various settings of the optical system.

Eq. 2.2, which yields:

$$w_0 = 12.3\,\mu\text{m},$$
$$z_R = 251\,\mu\text{m}.$$

The obtained spot size of a few µm, together with the sub-µm positioning precision, enables testing of specific in-pixel features of segmented silicon detectors, which typically have the pitch of $\mathcal{O}(10\,\mu\text{m})$.

It should be mentioned, that the focusing procedure needs to be performed separately for different DUTs, e.g. for the measurements, presented in Chapter 4.

2.2.3 Applications

Use of Caribou allows for fast and seamless integration of Caribou-compatible DUTs in the Laserbox. Several projects on silicon detector R&D at DESY benefit from the Laserbox, including:

- The setup was used to acquire pulse shapes in a generic silicon strip sensor, induced via charge injection with different lasers. This data was used to validate a Monte Carlo simulation module for interaction of laser beams with semiconductor detectors, developed in the course of this doctoral project, discussed in Chapter 3.
- Characterization of the dSiPM [49], a CMOS SiPM with a digital readout, developed at DESY. The timing performance of the chip and various in-pixel effects were studied in great detail, as presented in Chapter 4.
- In-pixel efficiency studies were performed [72] on the H2M MAPS prototype. The infrared laser was tuned to produce an m.i.p.-like charge deposition, using the pulse amplitude data from radioactive source measurements and a test beam. The H2M chip was illuminated from the backside, since the chip frontside has metalization. Data from the laser setup were used to investigate inefficiencies of certain regions within every pixel of the matrix, which were correlated to the layout of the on-chip electronics.
- Pulse shapes acquisition was performed [73] on a prototype of a pixelated CMOS sensor – the Analog Pixel Test Structure [74].

Chapter 3

Monte Carlo simulations of laser injection

This chapter is devoted to modeling of silicon detector operation via MC methods, used to simulate charge injection with pulsed lasers. In Section 3.1 the concept of simulations for R&D into silicon detectors is explained. In Section 3.1.2, the Allpix^2 detector simulation framework is presented. Section 3.1.3 covers implementation and details of the `DepositionLaser` module, an Allpix^2 module developed in the course of this thesis for modeling of interaction of laser beams with silicon sensors. In Section 3.2, the results of simulating the sensor response to the laser pulse are compared to the experimental data, acquired in the Laserbox setup.

3.1 Monte Carlo simulations for detector development

With recent advances in detector technologies and increasing complexity of silicon devices, i.e. adoption of CMOS monolithic sensors, computer modeling techniques, particularly Monte Carlo (MC) simulations [75], are widely employed for the detector R&D [76]. By modeling physical processes in a detector, simulations allow for evaluation of detector performance and can be used to optimize the detector design before the production, thus potentially reducing the amount of prototyping iterations. Furthermore, MC simulations are often performed to reproduce experimental characterizations of sensors, e.g. test beam campaigns. This approach brings benefits to both experiment and simulation. Comparisons between data and MC simulations can be used for interpretation of the experimental results and for understanding of the tested sensor in greater detail, as well as for validation or improvement of the simulation algorithms.

The simulation module developed in this work enables simulations for sensor testing scenarios that use laser charge injection into the bulk, i.e. TCT experiments. First applications of the module can be found in Refs. [77, 78].

3.1.1 General concepts of modeling the detector response

Detector performance is estimated by modeling its average response to various particle hits. This is achieved by simulating the entire chain of physical processes in the silicon detector, as described below.

Depletion and electric field in the sensor

Microscopic properties of the sensor, required to simulate its response, are the profile of the electric field, the shape of the depleted zone at a given bias voltage, and the distribution of the weighting potential. In a trivial case of a planar sensor, these quantities can be assessed analytically, e.g. the electric field can be assumed perpendicular to the sensor surface with its amplitude being linearly dependent on the depth in the sensor. For more complicated devices, e.g. for segmented detectors with complex arrangements of implants, analytical estimates are not feasible in practice. In this case, the needed quantities should be computed numerically with specialized *technology computer-aided design (TCAD)* software solutions. This technique is based on numerically solving two equations on the grid:

- The Poisson equation for the electric potential, using the sensor design with the bias voltage applied as boundary conditions;
- The transport-diffusion equation for the charge carriers, taking the electric field into account.

Since charge carriers themselves affect the electric field in the sensor, the solution for the required bias voltage is obtained iteratively: the bias voltage is increased to the target value in steps. At each step, a stable solution for the spatial distribution of the intrinsic charge carriers is found; the electric field is then recalculated taking the new charge carrier distribution into consideration. As a result, 2D- or 3D-maps are created, containing values of electric field, electric potential etc, segmented according to the grid. It should be noted that this technique is computationally heavy, i.e. typically requires up to hundreds of CPU-hours.

Particle interaction with the sensor

For interaction of particles with the sensor bulk, the amount and the positions of deposited charge carriers should be determined. In most cases, a simple model suffices: for a passing m.i.p., charge carriers are deposited in a thin line along the particle trajectory, and the amount of charge can be estimated using the Bethe-Bloch formula (Eq. 1.6).

A more precise estimate can be made using an MC simulation, e.g. with *Geant4* [79], one of the most widely used frameworks for modeling particle interactions with matter. Geant4 models various physical processes (ionization, multiple scattering, creation of δ-electrons etc), and thus is able to generate realistic distributions of the deposited energy for various impinging particles, which are in turn converted into distributions of the deposited charge.

Sensor response

The drift of charge carriers in the sensor bulk is then simulated using the previously obtained electric field and the distribution of the deposited charge carriers. The current pulse induced in the sensor electrodes can then be calculated with the help of the Shockley-Ramo theorem (Eq. 1.20), using the weighting field distribution.

One approach to simulate the charge carrier drift is to use the TCAD diffusion equation solver. While it yields accurate results, this approach is inefficient from the computational power perspective, since one typically needs to simulate the passage of many particles ($\gtrsim 10^4$) to properly estimate the detector performance.

The alternative approach is to assume that the total amount of charge carriers deposited by a particle is too small to significantly alter the electric field in the sensor, and to neglect any collective effects. Then, one can model individual charge carriers (or small groups thereof) one by one, which can be done very efficiently with the MC method. This way, large simulation samples can be produced at a low computational cost. This approach is being actively explored, particularly using the Allpix^2 software framework, presented in the following section.

Front-end electronics response

Signal processing in the detector front-end usually can be precisely modeled with simple numerical algorithms. For a more elaborate modeling, dedicated circuit simulators (e.g. SPICE [80]) can be employed, simulating the front-end response to the sensor signal based on the front-end design.

3.1.2 The Allpix^2 framework

The Allpix^2 (*Allpix Squared*) software framework [81, 82] is developed mainly at DESY to implement the simulation chains as described above, employing the MC method. It is written in C++ with ROOT [83] and uses a modular structure. Each module contains one simulation algorithm modeling one of the individual steps of the detector operation. For each of these steps, there are multiple implementations of modules accomplishing the same task, but using different algorithms and models. This makes the simulation highly configurable. A list of modules and parameters for them, according to the user preference, is specified in a configuration file which is passed to the framework as an input (an example is given in Appendix A). Besides that, a file with a description of geometry and type of detectors in the simulated setup should be provided. The framework operates on an event-by-event basis, with each event typically describing one instance of charge deposition in the setup (e.g. passage of one particle). Once run, the framework will execute the modules in the specified order for every event and store outputs. Typical simulation flow with main modules and configuration commonly used for them is shown in Figure 3.1.

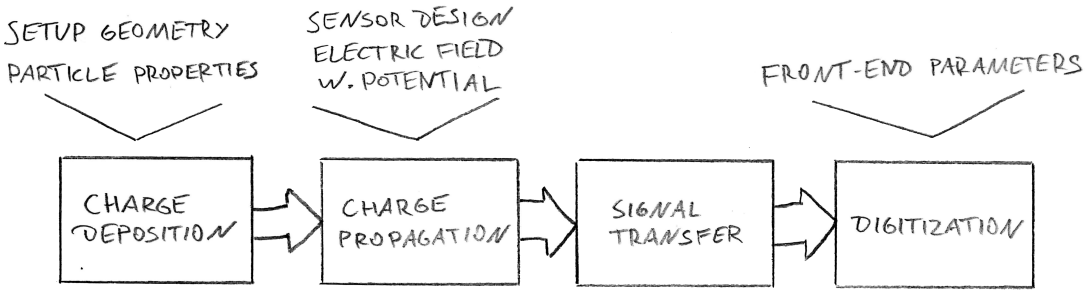


Figure 3.1: Simulation flow of the Allpix^2 framework [81].

Below, the module types with their variants are listed in an order, in which they are used.

Charge deposition

First, the initial distribution of charge carriers in the sensor should be produced. Simple analytical models or arbitrary user-defined distributions can be created with `DepositionPointCharge` and `DepositionReader` respectively. Alternatively, the passage of particles through the detector can be modeled in great detail with Geant4 MC algorithms that can be interfaced by Allpix^2. The properties of the incoming particles can be defined by the user (with `DepositionGeant4`), read from external generators (with `DepositionGenerator`), or modeled with the CRY generator [84] to mimic cosmic rays (`DepositionCosmics`). One can also model charge deposition with a pulsed laser with the `DepositionLaser` module, described in Section 3.1.3. As a result, individual charge carriers or groups of them are stored as framework objects for further processing.

Bulk properties and fields definition

To model the drift of charge carriers in the sensor and the signal formation, one should define distributions of several of the following quantities in each sensor: electric field, weighting potential, doping concentration. The exact selection of quantities depends on the charge propagation and signal formation models used in the following steps. Also, a magnetic field can be introduced. Each quantity is defined with a corresponding module, e.g. `ElectricFieldReader`. Each value can be either described analytically, or imported from TCAD as a grid.

Charge propagation

After the initial charge deposition and the fields in the bulk were defined, the motion (in the framework referred to as *propagation*) is simulated. Each charge carrier is tracked step-by-step, taking electric and magnetic fields, diffusion, and carrier mobility into account. While the `GenericPropagation` module just constructs the carrier trajectories, the `TransientPropagation` also calculates the current, induced on the electrodes, allowing for pulse shape simulations. Alternatively, a simplistic one-step estimate is available via `ProjectionPropagation`, offering significant savings of computing time at a cost of precision.

Signal transfer

The basic transfer modules count the total induced charge on the readout electrodes of the sensor: `SimpleTransfer` counts carriers, arriving to the readout implant; `InducedTransfer` takes into account the weighting potential at the origin of each of the carriers; `CapacitiveTransfer` allows one to simulate cross-coupled pixels/strips. The `PulseTransfer` module can also construct the full current pulse, if the temporal information on carrier drift is available from the previously executed `TransientPropagation`; otherwise, the pulse is constructed from the times of arrival of individual carriers.

Digitization

Various commonly used front-end electronics can be modeled with Allpix^2. Generic schemes with charge-to-digital converters and time-to-digital converters can be described with the

`DefaultDigitizer` module, and charge-sensitive amplifiers – with `CSADigitizer`. External circuit simulators can be employed as well. An interface for that is provided with the `NetlistWriter` module.

Auxiliary modules

Several auxiliary modules do not simulate any physical processes, instead, they provide data input, output, visualization, and processing.

- *GeometryBuilderGeant4* transforms the Allpix^2 geometry to a Geant4-compatible one, if Geant4-based deposition generators are used.
- `DetectorHistogrammer` calculates and visualizes performance metrics of the detector, based on a set of simulated events.
- Internal objects of Allpix^2 with information on intermediate simulation steps can be stored to and read from files with `ROOTObjectWriter` and `ROOTObjectReader` modules. This enables running simulation chains in parts, facilitating more flexible use of the framework and allowing one to perform a thorough analysis of simulation data.
- Data can be formatted in a specific way and stored with various `Writer` modules for further analysis with other frameworks. This is often used to simulate test beam experiments and then to analyze data with test-beam-specific frameworks, such as Corryvreckan (discussed in Chapter 7.3 and [85]) or EUTelescope [86].
- `VisualizationGeant4` allows to render an image of the experimental setup with a GUI utility, included in Geant4.

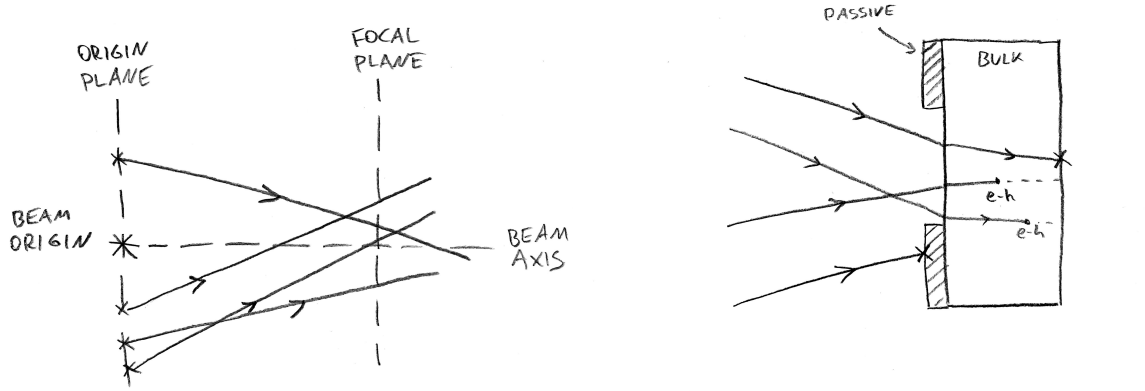
3.1.3 DepositionLaser module

For the purpose of simulating interaction of laser beams with semiconductor sensors, a dedicated module, the `DepositionLaser` [87, 88], was developed in the course of this doctoral project. As a deposition module, `DepositionLaser` is intended to be the first module in the simulation chain. As an input it takes the Allpix^2 setup geometry and user-defined parameters of the laser beam, specified in the configuration file.

Ray tracing model

The `DepositionLaser` module simulates the propagation of a gaussian laser beam through the experimental setup. While Geant4 offers a way to simulate optical photons [89], the exact implementation of that makes its integration to Allpix^2 unreasonable due to its architecture. Therefore, a standalone algorithm for modeling interaction of laser beams with semiconductor sensors was implemented in the module. An MC approach with the laser beam being modeled as a set of individual straight-line rays is employed. The exact algorithm is as follows (illustrated by Figure 3.2):

1. The user specifies geometrical properties of the beam, namely: coordinates of the origin, direction of the beam axis, origin-focus distance, opening angle, and beam waist. Also, parameters of the laser pulse should be set, including laser wavelength, pulse duration, and number of photons.



(a) Generation of the initial ray distribution. **(b)** Interaction of rays with detectors and passive objects.

Figure 3.2: Ray tracing principle in the `DepositionLaser` module.

2. A set of rays is produced, as shown in Figure 3.2a. For each ray, its origin in the origin plane and direction are generated randomly. The random distributions are chosen in such a way that in the focal plane of the beam the transversal intensity distribution is gaussian with the required beam waist (see Figure 3.3a), and the beam is has a specified divergence angle. Also, the time of origin is randomly generated with a Gaussian distribution for each ray to account for the nonzero duration of the laser pulse (see Figure 3.3b).

3. Each ray is traced through the setup. For determining intersections with the setup components, the Liang-Barsky clipping algorithm [90] is used. Interactions of rays with setup components are shown in Figure 3.2b.

4. If a ray encounters an object declared passive in the Allpix^2 geometry, it is terminated (see Figure 3.2b, bottom). This feature allows one to simulate the beam being partially blocked with e.g. metal strips on the sensor surface.

5. If a ray encounters a sensor, the interaction with the semiconductor bulk is simulated. If the sensor material is silicon, its optical properties, i.e. refraction index and absorption depth, are obtained from a built-in lookup table [38], available for laser wavelengths in the range of 250–1400 nm. Otherwise, these values should be set by user.

   5.1. The refraction angle of the ray on the air-semiconductor interface is calculated. Reflection on the surface is not modeled, as it is irrelevant for the simulation and requires precise knowledge of optical properties of the sensor surface. All rays are assumed to enter the bulk.

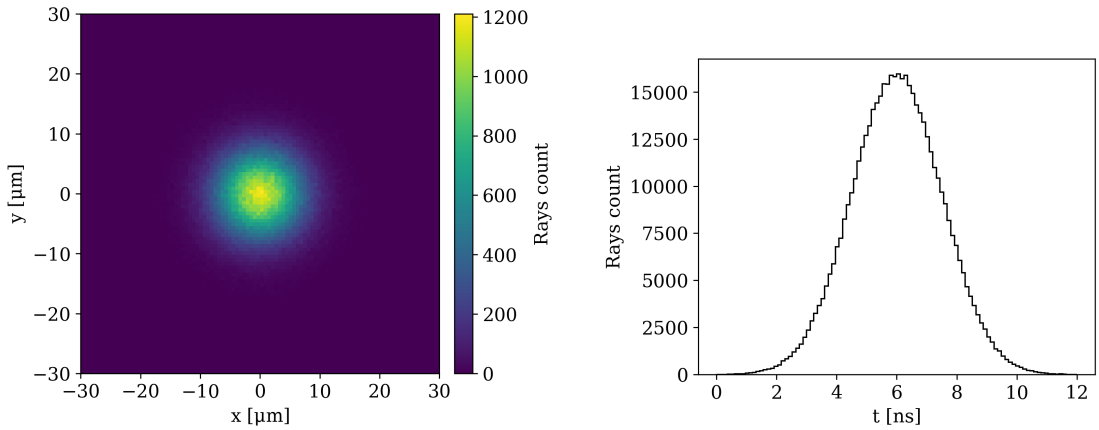
   5.2. A straight-line intersection segment of the ray with the bulk is determined.

   5.3. The penetration depth is generated randomly, following the exponential distribution with the specific absorption depth (as given by Beer law, Eq. 1.11).

   5.4. If the generated depth exceeds the length of the intersection segment, this ray is terminated (see Figure 3.2b, top). Internal reflection and leaving the bulk are not modeled.

5.5. Otherwise, this ray is considered absorbed via photoeffect. One electron-hole pair is created at the generated depth along the intersection segment. The origin time is calculated and assigned to these charge carriers, based on the ray origin time and travel time, for further use in the signal formation simulation logic.

6. The set of deposited charge carriers is accumulated and stored in Allpix^2 in order to be picked by following modules in the simulation chain (typically one of the propagation simulators).

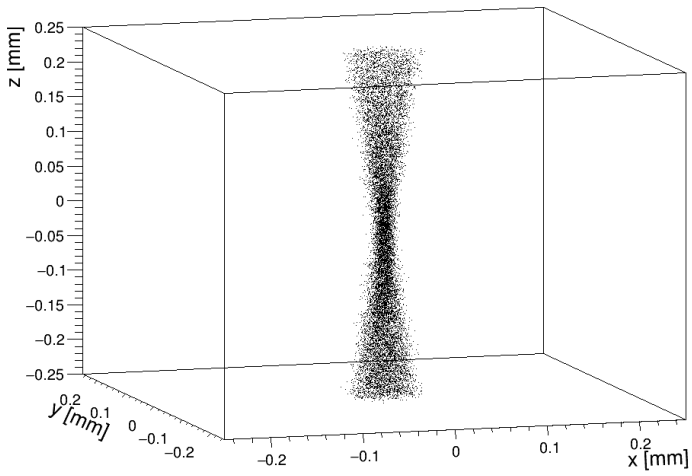


(a) Transversal beam profile at the focal plane.

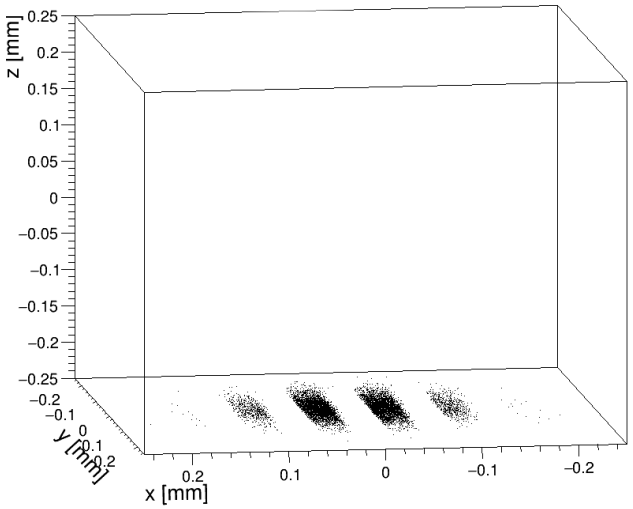
(b) Temporal structure of the laser pulse.

Figure 3.3: Simulated properties of the initial rays distribution. Model case with a converging beam with a waist of 10 µm and a Gaussian temporal profile with $\sigma_t = 1.5$ ns

Examples of spatial distributions of deposited charges for various laser beam configurations are shown in Figure 3.4. Illumination of the silicon bulk with a focused infrared beam is presented in Figure 3.4a. Uniform distribution of charge across the entire sensor thickness and a converging-diverging shape of a gaussian beam is clearly seen. Figure 3.4b shows the charge distribution for illumination of a strip sensor with a red laser of a large beam size. The charge is deposited in a thin surface layer, and a gaussian transversal beam profile is visible. Positions of metal traces on the sensor, defined in the simulation geometry, can be seen as regions without deposited charge.



(a) Illumination with an infrared laser beam with a waist of 10 µm.



(b) Frontside red illumination with a defocused beam. The sensor frontside is facing down on the figure and is covered with metal strips of 50 µm width and 50 µm spacing between them.

Figure 3.4: Examples of simulated deposited charge distributions. A generic 500 µm thick silicon strip sensor is used as a target. Each electron-hole pair is shown with a black dot.

3.2 Comparison of simulation and experiment

A TCT experiment was performed and then modeled with Allpix^2 in order to validate the simulation and to prove its prediction capabilities. A series of waveform measurements was performed on a strip sensor called *baby-PS-s*, described in the following, using the Laserbox setup (introduced in Chapter 2). The experimental setup was modeled using Allpix^2 in order to assess, whether the simulation chain is able to accurately reproduce the pulse shapes, induced by laser charge injection.

3.2.1 Experimental dataset

The baby-PS-s sensor, used in this study, is a test structure, produced alongside with PS-s sensors for CMS Phase-2 Outer Tracker modules (presented in Section 5.3.2) on the same silicon wafers. The sensor is comprised of strips with a pitch of 100 µm and uses a strip layout, identical to the regular PS-s sensors. However, the sensor has a smaller size of $18 \times 15\,\mathrm{mm}^2$ with a total of 128 strips. The sensor under test, mounted on a circuit board, is shown in Figures 3.5 and 2.7.

Depletion voltage of the sensor, required for tuning the TCAD simulation, was extracted from electrical characteristics of the sensor, measured in a probe station [91]. The obtained capacitance-voltage (C-V) and and current-voltage (I-V) curves are shown in Figure 3.6. The distinct kink of the C-V curve (as explained in Section 1.5.1) indicates the depletion voltage of approximately 280 V.

Waveforms were obtained for various bias voltages (up to 400 V), using red and infrared lasers with various pulse lengths (50 ps with the ALS system and ∼1.5 ns with Particulars lasers, introduced in Section 2.2.1). A defocused laser spot (waist $\simeq$ 100 µm) was used to uniformly illuminate a region between two strips, which effectively produces pulse shapes, averaged over the in-strip hit coordinate. In order to reduce the effect of the readout noise, an average of 100 consecutively acquired pulses was calculated.

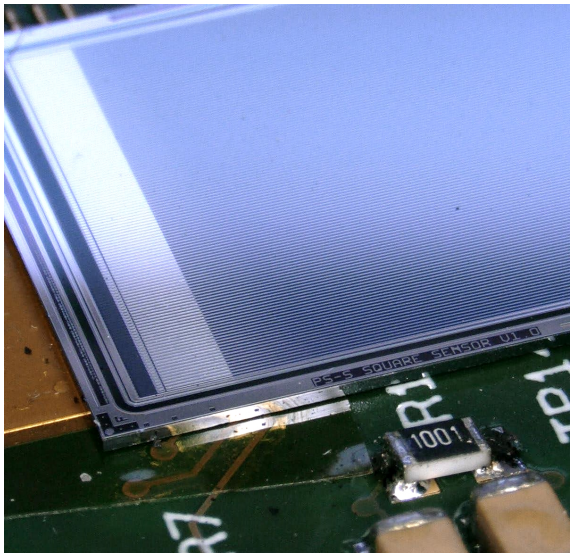


Figure 3.5: Sensor under test, glued onto a circuit board.

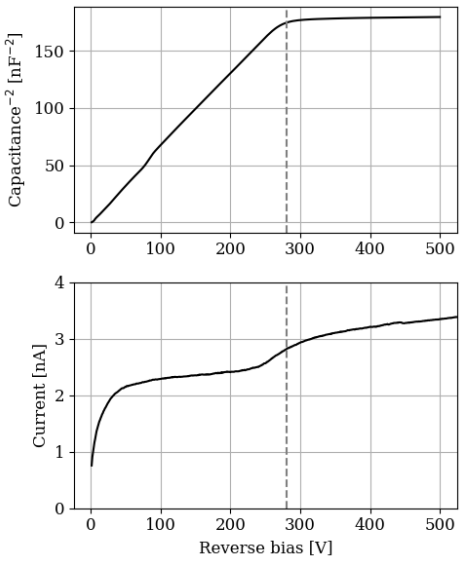


Figure 3.6: Measured C-V and I-V curves of the sensor. The full depletion voltage is shown with a dashed line.

3.2.2 TCAD modelling of the sensor

The electric properties of the studied sensor were simulated [92] using the Sentaurus TCAD software [93]. The geometrical configuration of the sensor and characteristic doping concentrations are given as inputs. Since exact values of doping concentration are not provided by the sensor manufacturer, the approach with generic doping profiles [94] is used. The layout of the sensor is shown in Figure 3.7 and includes the following parts:

- The p-bulk with the thickness of 290 µm. The bulk doping concentration needs to be manually adjusted in such a way that the simulated full depletion voltage matches the value, derived from the C-V measurement;
- The silicon dioxide passivation layer;
- The n^+ readout implant. The implant is 25 µm wide and 2 µm thick. A sigmoid transition of doping concentration between the implant and the bulk is added to ensure the continuity of the doping and to imitate diffusion of dopants in the real sensor;
- The p^+ backside implant;
- The metal electrodes. The upper section of the metal strip is 37 µm wide.

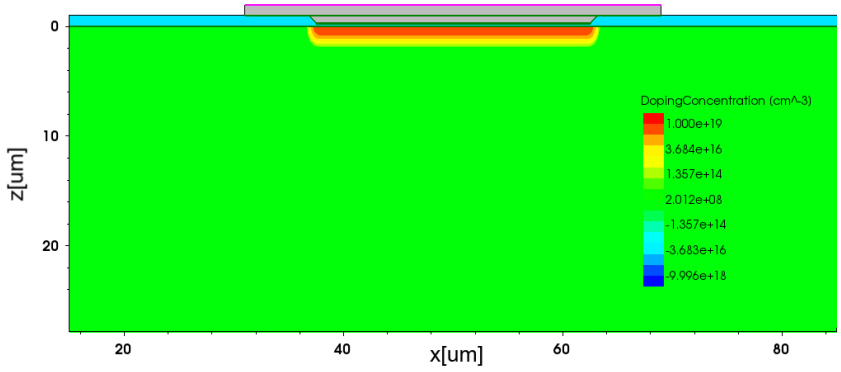


Figure 3.7: Cross-section of one strip of the PS-s sensor, zoom to a region around the readout implant. The bulk is shown in green, the passivation layer in blue, the readout implant in shades of orange, and the metalization in gray [92].

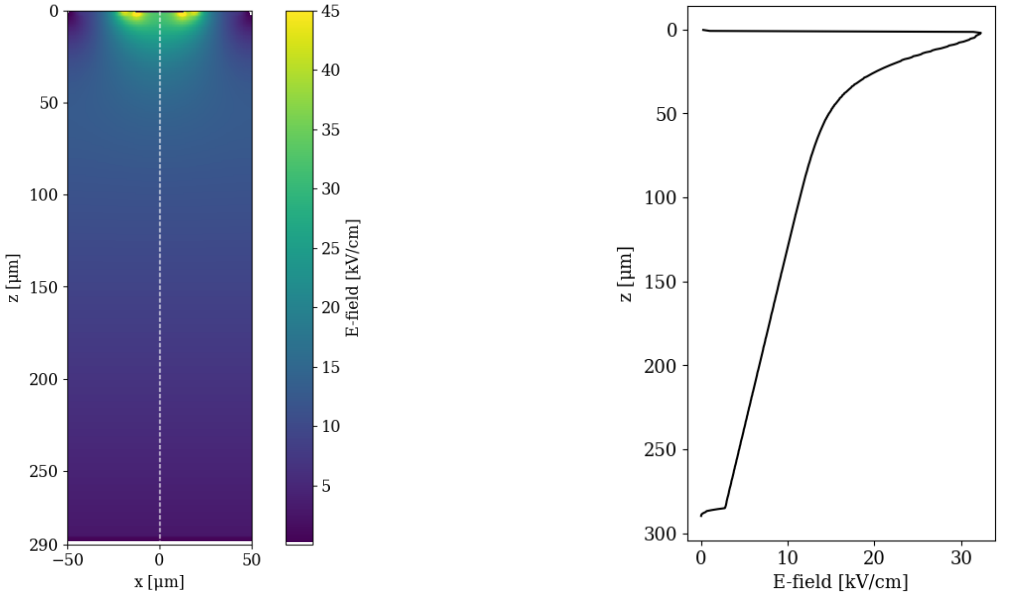
Since the strip sensor exhibits a translational symmetry along strips, a 2D TCAD simulation (in the xz-plane, transversal to the strips) is sufficient to determine the field maps. The simulation is performed in a region of 1 strip width (2 strips for weighting potential), and the obtained field maps are then replicated in x and extended in y to cover the entire sensor volume. This is valid as E_y is expected to be zero.

The electric field was modeled for a range of bias voltages; for every simulation scenario, an appropriate electric field map is used. An example of the calculated electric field magnitude map is presented in Figure 3.8. A characteristic linear dependence, typical for planar sensors,[1] is observed in the bulk ($z \gtrsim 50$ µm), whereas in the proximity of electrodes the field shape is more complex, i.e.

[1] $\vec{E} \parallel \vec{z};\ E_z \simeq a - bz$.

the dependence on the depth is not linear, and also the lateral component of the field is present (i.e. $E_x \neq 0$), indicated by inhomogeneity of the field near the lateral edges of the n^+ implant.

The weighting potential map is shown in Figure 3.9. As in other sensors with fine segmentation, signal in the (baby-)PS-s is mostly induced by charge carries drifting near the readout implant due to a higher gradient of the weighting potential in this region. It is important to mention that in this study the weighting potential was modeled assuming full depletion of the sensor, and the obtained map was used for all simulation scenarios. However, the distribution of the weighting field is expected to change in under-depleted sensors, which may lead to not entirely accurate description of the sensor at lower bias voltages.



(a) Magnitude of the electric field. Profile along the dashed line is shown in Figure 3.8b.

(b) Magnitude of the electric field at $x = 0$.

Figure 3.8: Distribution of the electric field in the sensor. A bias voltage of 300 V is applied.

3.2.3 Allpix^2 simulation of the laser injection

The Allpix^2 modeling of the laser injection experiment is organized as follows. The contents of the main configuration file, including typical values of simulation parameters, are listed in Appendix A.

1. The sensor geometry is defined: the sensor thickness is set to 290 µm. A sensor segment of a transversal size of $2\times2\,\text{mm}^2$ is simulated, consisting of 20 adjacent strips. For each strip, a passive object is added onto the sensor surface to simulate the blockage of the beam with opaque metal traces (see Figure 3.10).

2. The `DepositionLaser` module is used as the deposition generator to simulate the light interacton with the sensor. Simulated parameters of the beam, i.e. the wavelength and the pulse duration, are set to match those of the used laser driver.

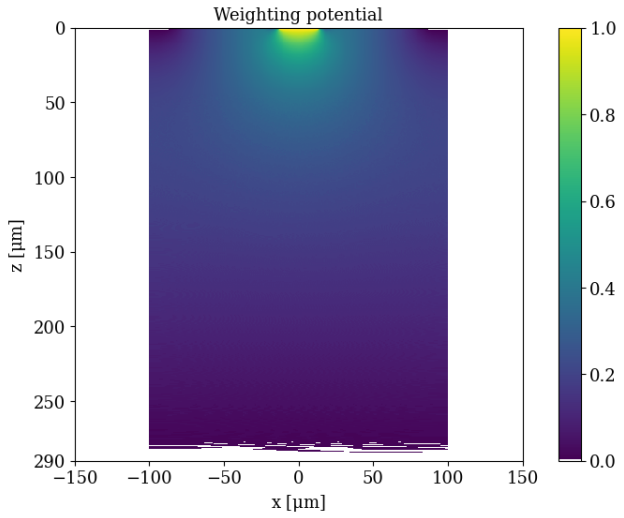


Figure 3.9: Weighting potential distribution in the sensor. One readout implant is shown.

3. The electric field and weighting potential maps, previously obtained with TCAD, are loaded with `ElectricFieldReader` and `WeightingPotentialReader`.

4. `TransientPropagation` is used to model the drift of charge carriers. Showcases of modeled trajectories are presented in Figures 3.11a (IR laser) and 3.11b (red laser). In either case, electrons would drift across the bulk towards the readout implants, and near the frontside they are laterally dragged towards implants, as governed by the electric field. Holes have similar trajectories, but in the opposite direction: near the frontside they are repulsed from the implants, and after that they drift towards the backside. Parameters for numerical integration of particle trajectories are set as follows: the total integration time is set to 20 ns so that all charge carriers are fully propagated; the integration step is set to 0.01 ns to achieve sufficient granularity.

5. Pulses shapes are constructed and stored with the `PulseTransfer` module.

6. Smearing of the pulse shape in the readout chain is modeled with an external Python script. An asymmetric CR-RC filter [95] with constants $\tau_{CR} = 1\,\mu s$ and $\tau_{RC} = 0.5\,ns$ is applied to the signals to account for the effects of the AC-coupled readout and for the finite bandwidth of the amplifiers respectively. Although these parameters could be derived from properties of the front-end circuits, in this study they are chosen retrospectively to provide the best data-simulation agreement.

7. A normalization is applied to the simulated pulses to match amplitudes to the experimental ones, since it is not possible to estimate the total amount of charge in the experimental pulse analytically. The scaling coefficient is the same for all pulses within one series (i.e. acquired with one laser setting).

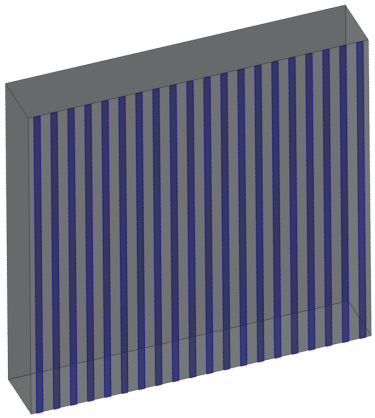
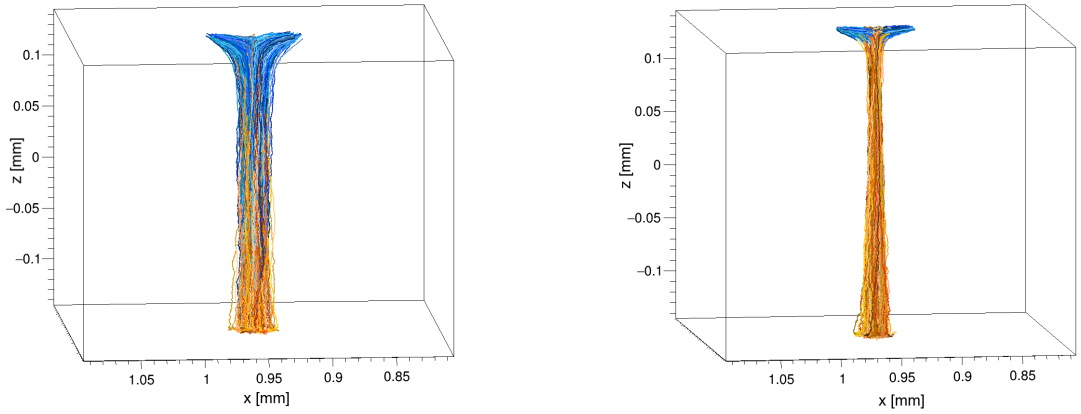


Figure 3.10: View of the sensor segment in the Allpix^2 simulation, produced via the `VisualizationGeant4` module. The sensor bulk is shown in gray, and the metal strips in blue.



(a) Illumination with an IR laser.

(b) Frontside illumination with a red laser.

Figure 3.11: Simulated drift trajectories of charge carriers in the sensor bulk. The sensor frontside is facing up. Electrons trajectories are shown in blue and holes' in orange. Two adjacent strips are located at $x = 0.9\,\text{mm}$ and $x = 1\,\text{mm}$. The laser beam is set to have a waist of $10\,\mu\text{m}$ for demonstration purposes, so that characteristic carrier trajectories can be clearly seen. The beam is pointed at $x = 0.95\,\text{mm}$, in the middle between two strips. Allpix^2 coordinate definitions are used: $z = 0$ denotes the centerplane of the sensor, thus the frontside is at $z = 0.145\,\text{mm}$.

3.2.4 Results

Choice of the laser pulse duration

An example of a pulse shape produced via illumination with the Particulars infrared laser (1064 nm wavelength, 1.5 ns pulse duration) together with the simulated pulse shape is shown in Figure 3.12. In this particular case, the waveform shape is dominated by the temporal structure of the laser pulse (see Figure 3.3b and Eq. 2.3). Therefore, the specific features of the pulse shape, that would have been produced in the sensor from instantaneous charge deposition, are smeared. Similar smeared pulses are observed with the Particulars red laser as well.

Such current pulses can nevertheless be reproduced with the simulation, since the `DepositionLaser` implements adjustable laser pulse duration.

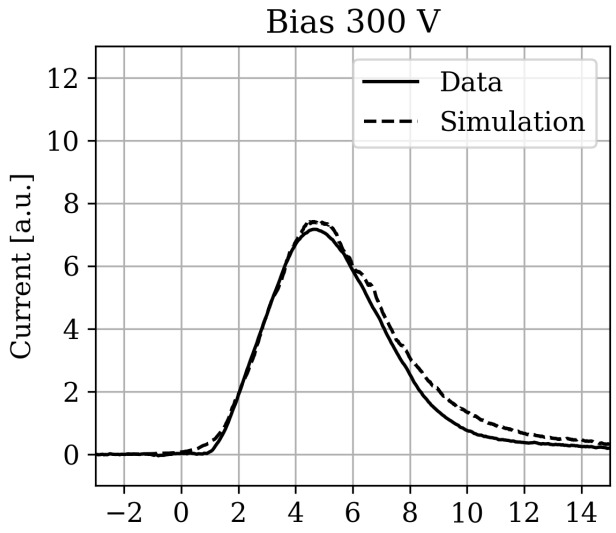


Figure 3.12: Example of a smeared pulse, obtained by illuminating the sensor with infrared Particulars laser.

Use of picosecond ALS lasers, in contrast, allows to resolve the fine structure of the pulse, as the impact of the laser pulse shape is negligible (as explained below). It should be noted that ALS lasers deposit a significantly smaller amount of charge compared to Particular lasers. Thus, the amplitudes of the produced pulses are also smaller, which makes them subject to noise and electromagnetic interference from the laser driver.

Infrared-TCT

A series of waveforms for different bias voltages for the illumination with the infrared ALS laser is presented in Figure 3.13, as well as the comparison with simulations. Integrals of these waveforms, i.e. the total collected charge, are shown in Figure 3.14.

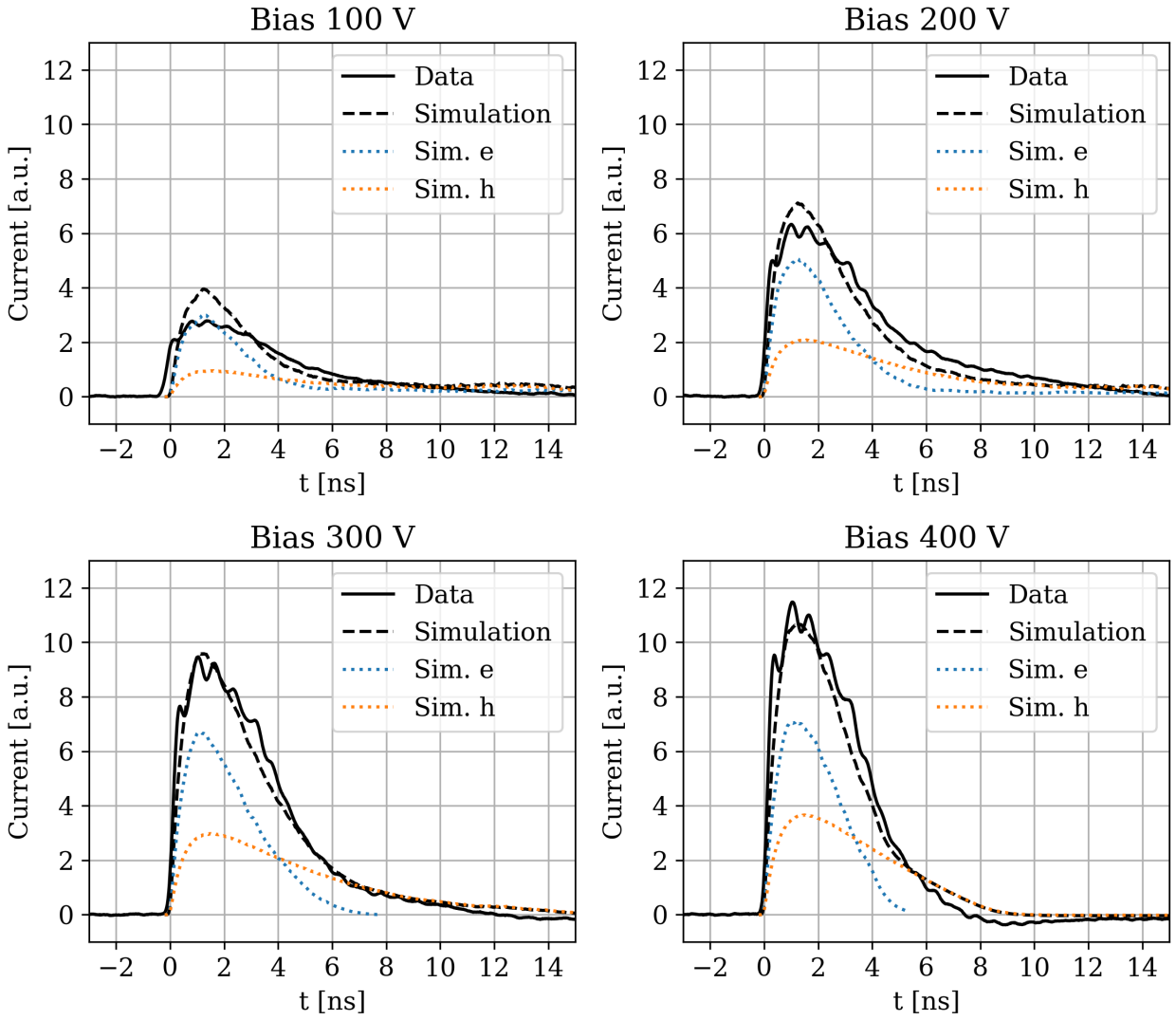


Figure 3.13: Comparison between measured and simulated pulse shapes. Illumination with infrared ALS laser (1063 nm wavelength, 40 ps pulse duration). Contributions to the simulated current from electrons and holes are shown with the dotted lines.

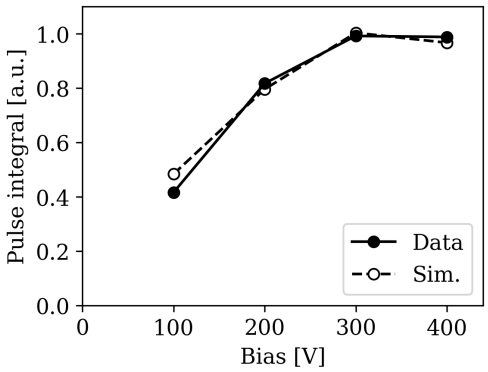


Figure 3.14: Integral of the induced pulse for IR illumination as a function of bias voltage.

The following features are observed:

- The integral of the pulse increases with the bias voltage, until the full depletion voltage ($V_{\text{bias}} \simeq 280\,\text{V}$) is reached (see Figure 3.6). This is due to the fact that the thickness of the depleted region grows, and thus the amount of collected charge carriers. This effect is matched by the simulation.
- The amplitude of the pulse increases even after the sensor is fully depleted, as higher bias voltage creates higher electric field, and the charge is collected faster, even if its amount is unchanged. This effect is reproduced by the simulation as well.
- The pulse has a sharp rising edge, since charge is deposited nearly instantaneously along the entire sensor thickness. The shape of the rising edge is affected by the amplifier, which smears it to a certain extent.
- The shape of the pulse tail is partially reproduced by the simulation. The discrepancy can be explained by the unknown doping concentration, which governs carrier mobility and diffusion.
- The shape of the tail is defined by the gradual decrease of the amount of the two types of charge carriers in the bulk over time, as they reach the implants and recombine. The simulation is able to independently model pulses from electrons and holes. Having previously obtained knowledge on characteristic features of electron and hole pulses (e.g. from the said simulation), one can then infer some of their properties from the measured pulse, which is the sum of the two. The electron component decays approximately three times faster due to the ratio of the carrier mobilities in silicon: $\mu_e/\mu_h \simeq 3$. The electron component also has a larger integral, since the PS-s sensor is electron-collecting: on average, a deposited electron drifts through a larger difference of weighting potential than a hole (assuming m.i.p.-like charge deposition from an IR laser).
- For the under-depleted sensor ($V_{\text{bias}} = 100\,\text{V}$), the pulse shape is not modeled entirely correctly because of imprecise modeling of the weighting potential in TCAD, which is often complicated for under-depleted sensors [92].

Red-TCT

Measured and simulated waveforms for frontside illumination with the red ALS laser are shown in Figure 3.15. Integrals of these pulses are shown in Figure 3.16.

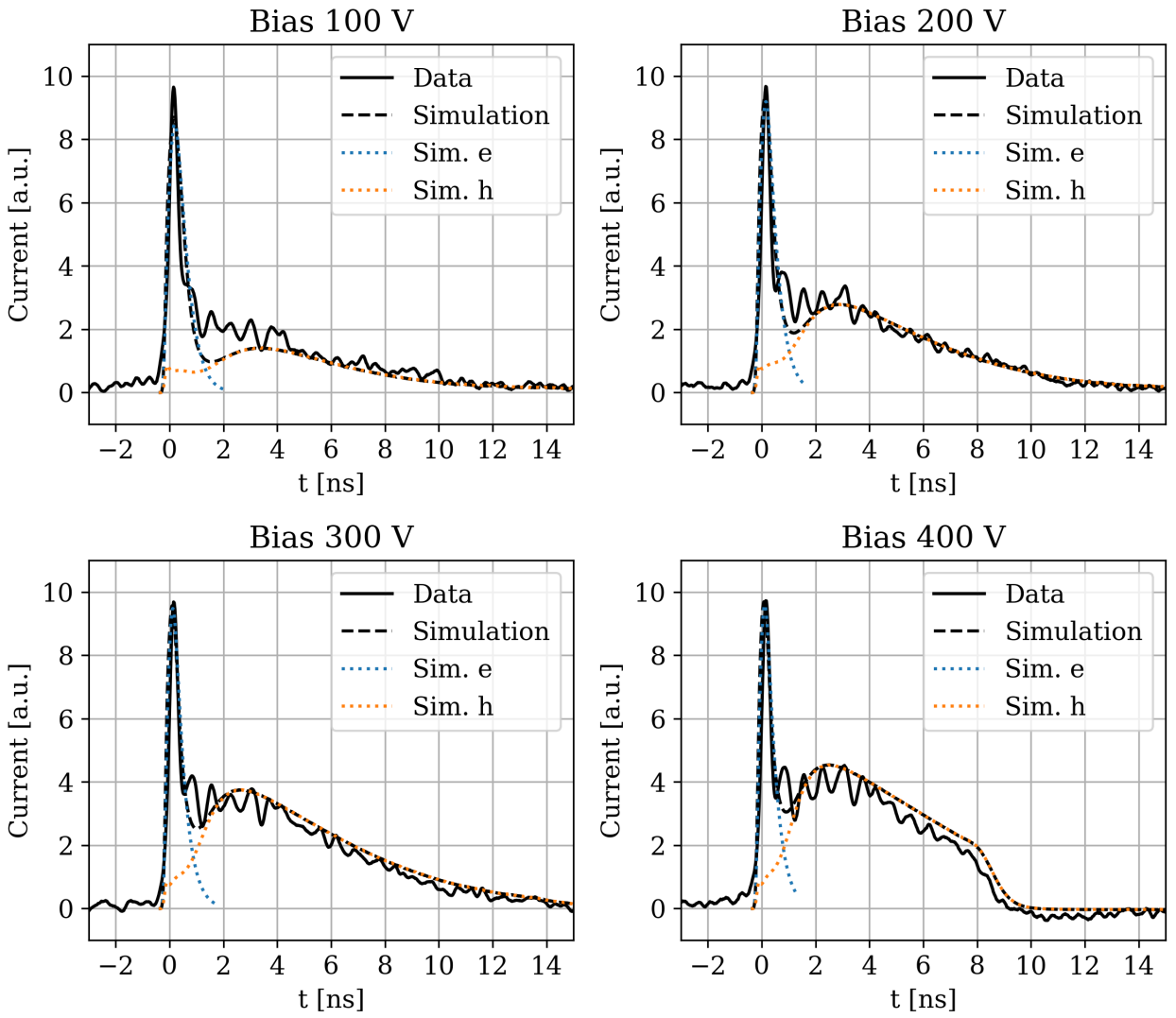


Figure 3.15: Comparison between measured and simulated pulse shapes. Illumination with red ALS laser (672 nm wavelength, 45 ps pulse duration). Contributions to the simulated current from electrons and holes are shown with the dotted lines.

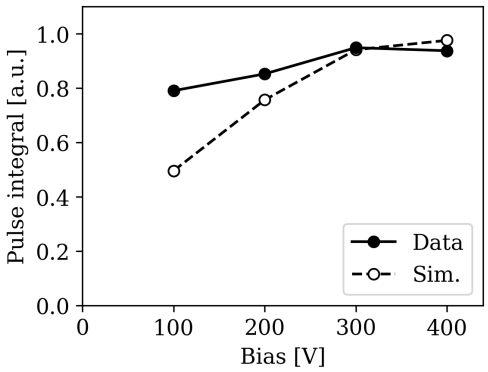


Figure 3.16: Integral of the induced pulse for red illumination as a function of bias voltage.

Their characteristic features are:

- Each pulse starts with a short peak from electrons, which is well reproduced by the simulation. Unlike the model case of red-TCT, where electrons immediately recombine in the readout implant, here they are deposited at a certain lateral distance from the readout implant, which is defined by the shape of the metal traces on the sensor. Therefore, a short drift phase is present (see Figure 3.11b), which produces a current spike, as both electric field and weighting potential gradient are high close to the readout implant. The integral of the electron peak depends on the width of the metal strips, which affects the initial spatial distribution of the deposited charge. The shape of the peak is mostly defined by the front-end circuitry.

- For the red laser, the pulse integral dependency on the bias voltage is less pronounced than that of the infrared illumination (see Figure 3.16). This is due to the fact that all the charge is deposited in a thin surface layer, which is already depleted even at rather low voltages. Therefore, a major fraction of the charge will be collected regardless of the depletion depth. This, however, is not matched by the simulation as closely, as the simulation underestimates the current right after the peak (i.e. at $t \simeq 1\,\mathrm{ns}$) at lower bias voltages. Particularly, for the under-depleted sensor ($V_{\mathrm{bias}} = 100\,\mathrm{V}$) the integral is underestimated by approximately 30%. A possible reason for that is the previously mentioned issue with estimating weighting potential for under-depleted sensors.

- A decaying component of the signal originates from the propagation of holes. Its shape is determined by the shape of the electric field and the weighting potential along holes' trajectories, as well as by the initial distribution of holes (i.e. value of weighting potential at the origin). Despite the differences between the baby-PS-s sensor and a model case for red-TCT (i.e. planar sensor without metalization on top), characteristic behavior of the hole current resembling the field profile is still observed. E.g. for the sensor at $V_{\mathrm{bias}} = 400\,\mathrm{V} > V_{\mathrm{depl}}$, a step-like decrease to zero is observed at the end of the falling edge, in contrast with the gradual decrease of the other pulses [59].

The demonstrated data–MC agreement serves as a validation of the `DepositionLaser` module and shows that Allpix^2 is sufficiently capable of modeling the experiments with laser charge injection into silicon detectors.

Chapter 4

Timing performance of a digital silicon photomultiplier prototype

This chapter is devoted to silicon detectors designed for precise timing measurements. Sections 4.1 – 4.2 provide a brief overview of timing performance of silicon sensors and technologies used to improve it. Section 4.3 introduces a DESY-designed prototype of a *digital silicon photomultiplier (dSiPM)*, which is a novel monolithic detector based on single-photon avalanche diodes with integrated readout, produced in CMOS technology.

In the course of the doctoral project, the laser injection setup, described in Chapter 2, was adapted to study single-photon-sensitive detectors, such as the dSiPM. A characterization of the dSiPM with a focus on its timing properties was performed for the first time using the laser injection, which is presented in Section 4.4.

4.1 Timing with silicon sensors

Besides passage coordinates, another property of a particle, which can be measured with a silicon tracking detector, is the precise time of the hit. In particle physics this is typically employed in time-of-flight systems (mentioned in Section 1.2). In future experiments with high *pile-up* (number of particle collisions per bunch crossing), reconstruction of multiple tracks at once can be improved if some of the individual detectors provide precise timing information. This concept is called 4D-tracking [96]. Other typical applications for precise timing detectors require detection of photons, not charged particles. Notable examples are readout of bunches of scintillating fibers [97], industrial *light detection and ranging (LiDAR)* systems [98], and medical imaging with the *time-of-flight positron emission tomography (ToF-PET)* technique [99].

For each impinging particle, detectors with timing capabilities are expected to produce a *timestamp* – a measured time of the particle passage with respect to a certain reference time. The key characteristic of a detector in this regard is its *time* resolution. Whereas a constant offset in the measured timestamps in most cases can be identified and accounted for, the measured timestamp would have a characteristic uncertainty originating from different physical phenomena that occur due to the functioning principles of the detector. The standard deviation σ of this uncertainty is the time resolution.

The time resolution of a silicon detector σ^2_{Total} is affected by several factors [100], and can be estimated[1] as:

$$\sigma^2_{\text{Total}} = \sigma^2_{\text{Timewalk}} + \sigma^2_{\text{Jitter}} + \sigma^2_{\text{Landau}} + \sigma^2_{\text{Formation}} + \sigma^2_{\text{TDC}}, \tag{4.1}$$

where:

- σ_{Timewalk}: in most readout systems the timestamp is measured as the moment when the current of the induced pulse crosses a pre-defined threshold. As the rise time of the pulse is nonzero, two pulses of identical shape, but with different amplitudes, would cross the threshold at different time. This discrepancy is called *timewalk* and can be estimated as:

  $$\sigma_{\text{Timewalk}} = t_{\text{rise}} \cdot \frac{\text{Threshold}}{\langle \text{Amplitude} \rangle}, \tag{4.2}$$

  approximating the rising edge of the pulse with a linear function. Timewalk can be explicitly compensated for by using a calibration curve if information on the pulse amplitude is provided by the detector front-end. Alternatively, the effect can be eliminated by the use of a *constant fraction discriminator* instead of a fixed threshold.

- σ_{Jitter}: the current pulse is subject to random noise that adds up to the signal and might affect the moment when the threshold is crossed. The timing uncertainty caused by this effect is called *jitter*:

  $$\sigma_{\text{Jitter}} = \frac{\text{RMS}_{\text{noise}}}{\partial I / \partial t} \sim \frac{\text{RMS}_{\text{noise}} \cdot t_{\text{rise}}}{\langle \text{Amplitude} \rangle}. \tag{4.3}$$

  The jitter can be dealt with by the use of low-noise sensors, e.g. SiGe BiCMOS [102].

- σ_{Landau}: the distribution of ionization-induced charge carriers along the particle track in the sensitive volume is not uniform and subject to significant fluctuations. For passage of two identical particles, different initial carrier distributions can be formed, thus pulse shapes and timestamp readings would be also different.

- $\sigma_{\text{Formation}}$: the pulse shape might as well be affected by the process of the pulse formation itself, determined by the characteristics of the electric field and the weighting potential in the sensor. For instance, hits at different transversal coordinates w.r.t. a collection implant in the detector may result in different pulse shapes.

- σ_{TDC}: the threshold crossing time is measured and digitized by a dedicated circuit, a *time-to-digital converter (TDC)*. This digitization has a finite precision limited by the TDC properties. A TDC determines in which of consecutive pre-determined time intervals (or *TDC bins*) the signal arrives. Assuming a constant bin width, the contribution to the time resolution reads:

  $$\sigma_{\text{TDC}} = \frac{\text{Bin width}}{\sqrt{12}}. \tag{4.4}$$

The uncertainties in the time measurement, intrinsic to the sensor (i.e. timewalk and jitter), decrease if the sensor produces pulses of higher amplitude and with shorter rise time. It should be noted that the increase in signal must come from the sensor itself, not from an amplifier: whereas amplifiers increase the signal level, they also increase the noise and extend the rise time. A common solution to produce higher signals is the use of silicon sensors with *internal gain*.

[1] Additionally, cross terms of sensor- and electronic-related contributions can be present, as explained in Ref. [101].

4.2 Silicon sensors with gain

4.2.1 Ways to increase signals

Potential possibilities to increase signal levels in conventional silicon sensors in order to improve their time resolution are:

- Higher sensor thickness: the average amount of carriers per unit travel distance in silicon is fixed (as given by the Bethe-Bloch formula, Eq. 1.6). The total amount of charge can be nevertheless increased by the use of thicker sensors. However, it can be derived from the Shockley-Ramo theorem (Eq. 1.20), that for a planar sensor the increase in thickness does not affect the pulse amplitude. The higher initial amount of charge carriers is compensated by the increase of the distance to travel and by the decrease of the weighting field gradient;
- Faster carrier drift: applying higher bias voltage creates higher electric field in the sensor, causing a faster carrier drift. The voltage, however, cannot be increased indefinitely, as it would eventually result in a breakdown. Another option to improve the drift time is a special arrangement of implants that decreases the drift distance, while keeping the sensor thickness (and the amount of charge) the same, which is employed in so-called *3D sensors* [103].

Another conceptually different approach can be employed for a significant improvement of time resolution: fast-rising signals can be induced in silicon sensors by exploiting the phenomenon of the *impact ionization* [37, 104]. Free charge carriers, drifting in the silicon bulk, may scatter on bound electrons in the valence band, transferring a portion of energy to them. This would move a bound electron to the conduction band, creating an additional electron-hole pair. The probabilities for such process per unit distance that an electron or a hole travels are given by the impact ionization coefficients[1] α_e and α_h. These coefficients strongly increase with the electric field:

$$\alpha_{e,h} \propto \exp\left(-\frac{E^{\text{crit}}_{e,h}}{E}\right). \tag{4.5}$$

If the charge carriers drift in an electric field above a certain critical value E^{crit}, they undergo multiple impact ionization processes, and the amount thereof rapidly increases. This process is also referred to as an *avalanche*.

In practice, electric fields required to maintain the charge multiplication are created locally in the silicon bulk by introducing highly-doped *gain implants* under the readout implants, which would create a so-called *amplification region* with a large electric field (of the order of $10^5\,\text{V/cm}$). The development of an avalanche in the amplification region induces a fast-rising pulse with a high amplitude in the detector. The total multiplication factor after a passage of a set of carriers through a high-field region is called *gain*. The functioning and operation mode of such sensors significantly depends on its their gain and the exact properties of the gain implants.

Sensors with a gain of the order of 10 are classified as *low-gain avalanche detectors (LGADs)*. Gain implants in LGADs create a *proportional* amplification: the amount of charge carriers after multiplication is directly proportional to their initial amount. In their form and function LGADs are similar to conventional silicon sensors without gain, but offer a factor 10 improvement in the timing resolution: from $\mathcal{O}(\text{ns})$ to hundreds of ps [106], while maintaining other characteristic properties of silicon sensors, such as good spatial resolution.

[1] Sometimes referred to as Townsend coefficients in an analogy to the theory [105], originally developed for gases.

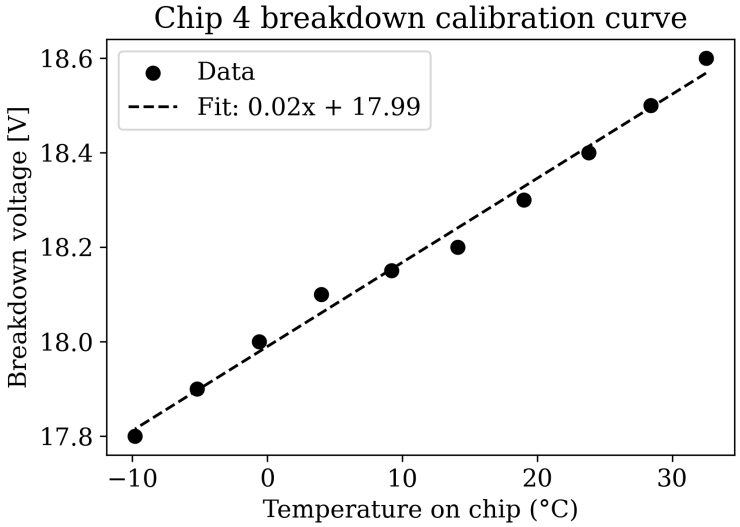


Figure 4.1: Breakdown voltage of CMOS SPADs used in this work. A linear fit is used to estimate V_{bd} at a given temperature. Adapted from [49].

In sensors with a much higher gain (10^4 and higher), the amplification is not proportional anymore [107]. An avalanche might grow infinitely,[1] retaining no information on the initial charge, and even a single charge carrier is enough to trigger such an avalanche. Devices with such gain are capable of detecting energy depositions as low as ones induced by individual optical photons, and thus are called *single-photon avalanche diodes (SPADs)*. The minimal bias voltage that is enough to create the electric field sufficient for the avalanche development is called the *breakdown voltage* V_{bd}. The breakdown voltage as a function of temperature for the SPADs used in this work (introduced in Section 4.3) is shown in Figure 4.1.

SPADs are operated at a bias voltage exceeding V_{bd}, so that avalanches can be created; the voltage excess is referred to as *overvoltage*. In order to limit the avalanche development, SPADs are biased through a resistor. The development of an avalanche creates a current through the SPAD, resulting in a voltage drop on the resistor and a consequent decrement of the bias voltage supplied to the SPAD, cutting the avalanche off. This effect is called *quenching*. The downside of high-gain detectors is their proneness to develop avalanches spontaneously from thermally generated charge carrier pairs. Such events, the *dark counts*, cannot be distinguished from particle-induced signals. The dark count rate increases with overvoltage and temperature. Due to extremely fast rising pulses, SPADs offer excellent intrinsic timing resolution of tens of ps.

4.2.2 Silicon Photomultipliers

While being sensitive to any type of energy deposition, e.g. from ionizing particles [108], SPADs are conventionally used to detect optical photons, e.g. for picking up scintillation or Cherenkov light. Large-sized SPADs are unusable in practical applications, as registration of a single photon or a single dark count[2] renders the whole device inactive until the avalanche is fully quenched. This can be overcome by implementing a device with multiple smaller SPADs arranged

[1] It is typically limited by external factors, such as maximal possible current, sustained in the detector biasing circuit.
[2] The dark count rate scales proportionally with the SPAD area.

in a dense grid. Such arrays are called *silicon photomultipliers (SiPMs)*. A comprehensive overview of SiPMs can be found in [109]. In silicon photomultipliers, SPADs are electrically connected in parallel, and each cell is equipped with an independent quenching resistor, so if one cell triggers, the rest of the matrix remains active. Modern SiPMs are manufactured via a lithography process, and individual SPADs are designed and isolated from each other, such that a developing avalanche stays contained in one SPAD. Nevertheless, a so-called *crosstalk* event may happen: a neighboring SPAD can be triggered by an optical photon, emitted by the avalanche. The probability of such an event is relatively low (typically less than 1%).

By measuring the total current flowing through a SiPM, one can infer the exact number of SPADs currently firing. This makes SiPMs a well-suited instrument for photon counting applications, i.e. precise measurements of intensity of light pulses. For optimal results the light should be diffused over the matrix, and the matrix should not be saturated. E.g. if a major fraction of SPADs is firing, the probability of absorption of several photons in a single SPAD is increased. However, the resulting signal would not differ from that originating from absorption of exactly one photon per cell. One should also consider two additional factors for photon counting:

- Quantum efficiency: the probability of a SPAD to trigger on an incoming photon is typically 30–50% and depends on the photon wavelength and SPAD design;
- *Fill factor*: not the entire SiPM area is active, as a certain area on the chip is required to isolate SPADs with e.g. *trenches* [110]; the active area is typically 80–90%.

Typically, SiPMs are produced via specialized lithography processes and are only available as off-the-shelf designs, i.e. no modifications to the design are possible. With their cells being connected in parallel, SiPMs provide no spatial information on the incident particles. Thus, for employing SPAD-based devices in applications requiring high spatial and time resolution simultaneously, modifications to the readout scheme should be made.

4.3 The DESY Digital SiPM Prototype

Developments on the readout circuitry for SiPMs are carried out at DESY Hamburg. The research goal is to develop a readout with a dedicated channel for each cell, akin to pixelated tracking detectors. Such a design allows to determine which exact SPADs produced the signal, thus providing information on hit position. Additionally, it allows to disable certain cells (e.g., noisy or faulty) on demand, and also allows to implement a simplistic hit pattern recognition mechanism.

The previous generation of a postion-sensitive SiPM featured a separate digital readout chip, produced in a 130 nm Global Foundry CMOS process [111, 112]. The readout ASIC was attached to a SPAD matrix via flip-chip bump-bonding, forming a *hybrid SiPM*. The prototype matched the design goals, however, due to its hybrid design, the sensitive surface of the SPAD matrix was inaccessible for optical photons.

With recent advances in semiconductor manufacturing technologies, SPADs are now also offered in *process development kits (PDKs)* at several specialized CMOS foundries [113, 114], thus facilitating custom designs of application-specific integrated circuits incorporating SPADs. A cross-section of a CMOS SPAD, similar to the one used in this work,[1] is schematically shown in Figure 4.2. This new iteration of the design is implemented following a *monolithic* approach, with

[1] The exact design is not disclosed by the manufacturer.

the sensitive cells including the in-cell electronics and the readout circuit produced on one substrate. This prototype is referred to as the *digital SiPM*, or *dSiPM*. Its design and specifications are discussed below, and a detailed reference can be found in [49].

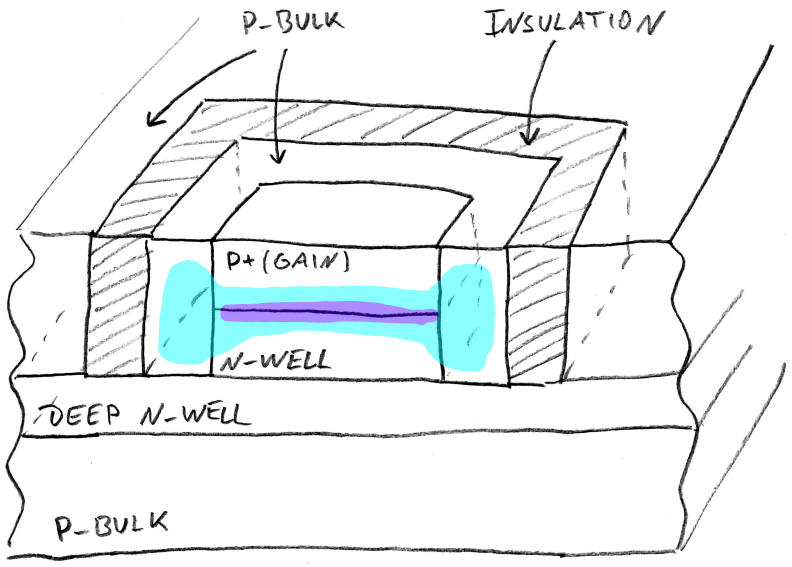


Figure 4.2: Schematic cut section of the p+/n-well CMOS SPAD. In the cut plane, the amplification region is shown with dark blue shading, and the depleted region in light blue. Adapted from [114].

As the dSiPM provides spatial and temporal information simultaneously, it is an interesting detector candidate for applications requiring high-granularity timing detectors for either ionizing particles or photons.

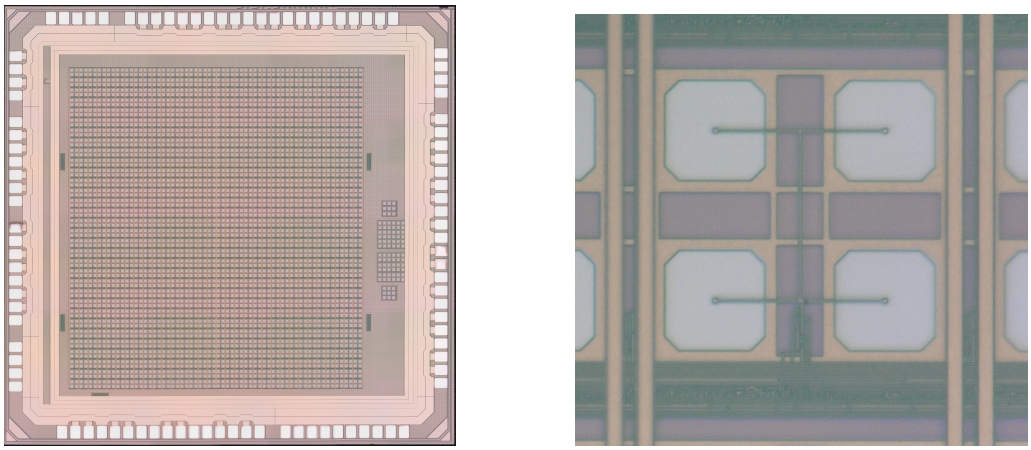
4.3.1 dSiPM specifications

The dSiPM prototype is manufactured using the LFoundry 150 nm CMOS technology. It consists of a sensitive SPAD matrix of approximately $2.2 \times 2.4\,\mathrm{mm}^2$ in size and several peripheral blocks, including time-to-digital converters, encoders, a data transmitter and a data receiver, as well as test structures.

Matrix design

The sensitive matrix has 32×32 independent pixels with a pitch of $69.6 \times 76\,\mu\mathrm{m}^2$. The matrix is subdivided to four quadrants of 16×16 pixels. In each pixel there are four 20×20 $\mu\mathrm{m}^2$ SPADs, connected in parallel and sharing a common quenching circuit. This results in a fill factor of 30%. The front-end is also shared between the four SPADs. Thus, within a pixel it is impossible to distinguish which exact SPADs of the four were triggered. This layout with multiple SPADs in a pixel was chosen deliberately in order to optimize the fill factor, while staying compliant to

the rules of the manufacturing process, e.g. limits on a minimal distance between SPADs and other elements. Although an arrangement with less SPADs (or smaller SPADs) per pixel could be produced as well, which would increase the granularity and spatial resolution, the fill factor and thus the detection efficiency for such a scheme would be significantly lower.



(a) Full chip ($\sim 3 \times 3\,\mathrm{mm}^2$). **(b)** One pixel ($69.6 \times 76\,\mu\mathrm{m}^2$).

Figure 4.3: Micro-photographs of the dSiPM chip [73].

Data acquisition scheme

The readout of the matrix is performed in a frame mode. The frame frequency is set by an external clock of 3 MHz (period of 333.3 ns). Additionally, the chip receives a 408 MHz system clock (period of 2.45 ns). One cycle of the frame clock comprises 136 cycles of the system clock, during 128 of which the matrix stays active, thus resulting in acquisition frames with a duration of 313.7 ns, followed by an inactive period of 19.6 ns, which is used to reset and configure the matrix. For each frame the full *hitmap* of 1-bit depth is stored (i.e. list of cells triggered during this frame). The device also allows for operation in a so-called *2-bit mode*, acquiring frames of doubled duration and storing up to 3 hits per cell.[1] In this work the 1-bit mode is used. The chip is equipped with a set of programmable logic gates that are able to perform row-wise operations and generate a *validation bit* for each frame, which functions as an arbitrarily configurable flag, which can be used e.g. to implement simplistic on-chip cluster recognition.

The chip is capable of a sustained outbound transmission of the acquired frames at the frame clock speed. The high-level control of the chip and the data readout is performed with the Caribou system [60].

Timestamping scheme

To complement the intrinsically high timing resolution of the SPAD matrix, the chip is equipped with four high granularity time-to-digital converters, one per matrix quadrant. All cells of a quadrant are connected to the quadrant's TDC via the so-called *wired-OR* network. This circuit

[1] In order for multiple hits to be counted in a single cell within a frame, the cell needs to be fully quenched between the arrivals of photons. Thus, this offers no additional dynamic resolution for short light pulses.

is unable to provide an exact timestamp for every cell triggered, instead it measures when *the first* cell triggers within a quadrant in that frame. Hence, each frame comes with up to four timestamps, one per quadrant. A timestamp consists of three values, as explained below.

- A *bunch counter* value `bc` representing the frame number. Frame numbers are stored in a register with a depth of 40 bits, allowing for more than 100 hours of continuous data taking. Frame counting is performed by both the chip and the external control system.
- A *coarse timestamp* value `ct` representing a number of counts of the system clock cycle, at which the hit comes. The coarse timestamp is calculated with a 7-bit counter in each TDC and resets at the beginning of every frame. The granularity of the coarse timestamp is $1/(408\,\text{MHz}) = 2.45\,\text{ns}$.
- A *fine timestamp* value `ft`, which is the most precise measurement within a system clock cycle. Fine timestamps are measured in each TDC with a *delay-locked loop (DLL)*. The DLL consists of 32 sequentially connected analog delay elements, and the first one is triggered synchronously with the system clock. The clock signal propagates through the delays, and a *phase comparator* ensures that the integrated delay along the full delay pipeline equals the duration of the system clock cycle. If necessary, the phase comparator adjusts the delays. At the moment the hit occurs, the number of the delay element to which the signal has propagated, encodes the time interval between the rising edge of the system clock and the hit arrival. This number constitutes the 5-bit value of the fine timestamp. The design granularity (later referred to as a *fine TDC bin size*) of this measurement is $2.45\,\text{ns}/\,32 = 76\,\text{ps}$. Due to manufacturing variances, the delay elements on the prototype do not meet the specification, effectively resulting in a lesser amount of usable bins and a deterioration of the granularity of the fine TDC to approx. 95 ps. A possible workaround is discussed in Section 4.4.2.

The full timestamp can be reconstructed from the values `bc`, `ct`, and `ft` as:

$$\text{Timestamp} = \frac{\texttt{bc}}{3\,\text{MHz}} + \frac{\texttt{ct}}{408\,\text{MHz}} + (\texttt{ft} + \frac{1}{2}) \times \text{DLL bin width}. \tag{4.6}$$

4.4 Characterizations of the timing performance with a pulsed laser

Given the inhomogeneous structure of the chip, characterization of its hit detection efficiency and timing properties as functions of the in-chip and in-pixel coordinates is of particular importance, and it can aid future design choices. Previous studies at the DESY II test beam facility [115, 116, 117] have shown a variance of the time resolution as a function of the position within a SPAD.

In this thesis, a study of the chip characteristics was for the first time performed with a pulsed laser setup, introduced in Section 2.2. The ALS red laser [62] is used, which is suitable for precise timing measurements thanks to the sharp rising edge of the pulse and a low triggering uncertainty of less than 5 ps. The ability to precisely adjust the laser beam and to explicitly trigger the laser pulse allows for direct measurements of the parameters in question. Due to the digital design of the chip and dedicated test structures, impacts of different phenomena affecting the dSiPM time resolution can be quantified independently.

Two prototypes of identical design were tested in this work, referred to by their serial numbers as chip 1 and chip 4.

4.4.1 Adjustment of the laser system

Various testing scenarios can be implemented by adjusting the laser spot. Namely, changing the spot size may include or exclude certain effects, affecting measurements of the SPAD matrix properties. Two possible options are:

- For obtaining best-case estimates of SPADs characteristics, discussed in Sections 4.4.3 and 4.4.4, the laser should be de-focused and run at high intensity, which ensures illumination of the entire pixel area.
- For resolving in-pixel features (as presented in Section 4.4.5) the laser system has to be adjusted to ensure localized charge deposition. This procedure is discussed below.

The laser spot should be tuned such that the deposition is point-like, and a SPAD can be triggered only if the laser is pointing directly at it. Assuming a Gaussian shape of the beam, this requires achieving an $\mathcal{O}(\mu\mathrm{m})$ beam waist, and positioning the lens such that the beam reaches the device surface at the beam waist.

Additional complications are caused by the single-photon sensitivity of the matrix. Even if a beam is focused to a narrow gaussian spot, its intensity in off-center regions can still be enough to trigger SPADs with a non-negligible probability. This effect can be mitigated to a major extent by precisely adjusting the intensity of the laser. By use of neutral density (ND) filters, the beam can be optically attenuated, and the sensitivity of the SPAD matrix itself can be adjusted by changing the bias voltage. By varying these parameters, the injection scheme can be tuned such that intensity in the off-center regions is low enough, while in the center it is still sufficiently high.

Laser parameters and coarse focus adjustment

The process of the coarse adjustment of the injection parameters is presented in the Figure 4.4. The laser is aimed at the center of the SPAD matrix, triggered once per readout frame, and full hitmaps are stored. Since the dSiPM matrix has no dynamic depth, the intensity of the laser in a given pixel is estimated as a probability of this pixel to trigger. Multiple frames are acquired for each setting, and the fraction of frames, where a given pixel was triggered, is calculated.

The top row in Figure 4.4 shows the dependence of the spot size on the position z of the lens along the optical axis. At $z \simeq 27.3\,\mathrm{mm}$ an approximate focus is achieved, and moving the lens in z in both directions results in spreading of the observed spot. It should be mentioned that due to the use of a circular aperture the beam shape is distorted [118]. Two noticeable distortion effects are:

- non-Gaussian transversal beam profile closer to the lens; a distinct ring shape is observed instead (Figure 4.4, upper row, rightmost frame);
- asymmetrical behavior with changes in z: quick spreading of the Gaussian distribution if the lens is moved further from the matrix (z decreases), and slower spreading with emergence of the ring profile if the lens is moved closer to the matrix (z increases).

Both of these beam distortion effects can be observed during the consequent fine z adjustment as well. The middle row in Figure 4.4 shows the impact of changing the laser intensity with neutral density[1] filters on the size of laser spot perceived by the dSiPM. Increasing the attenuation decreases the area of the spot in which cells are triggered. The bottom row in Figure 4.4 showcases

[1] Neural density (ND) is a base-10 logarithm of the intensity attenuation factor.

the effect of changing the bias voltage. Increasing the voltage makes the matrix more sensitive, thus more pixels are triggered by the same laser pulse. Parameters, ensuring triggering of one cell with a high probability, whereas in the adjacent cells this probability is low, are chosen as the working point.

Fine adjustment

After the coarse z value is found, a fine adjustment is performed. An adaptation of the knife-edge method is used (explained in detail Section 2.2.2). The procedure is as follows:

- The beam spot position within a single pixel is varied along a line path, shown in Figure 4.5a. This requires preliminary aiming in x and y to ensure a proper positioning of the scan path within a pixel. For the optimal z position a "double-box" structure is expected for the hit probability dependency on the in-pixel x laser spot coordinate.
- The previous step is repeated for various z values around the previously found coarse z position.

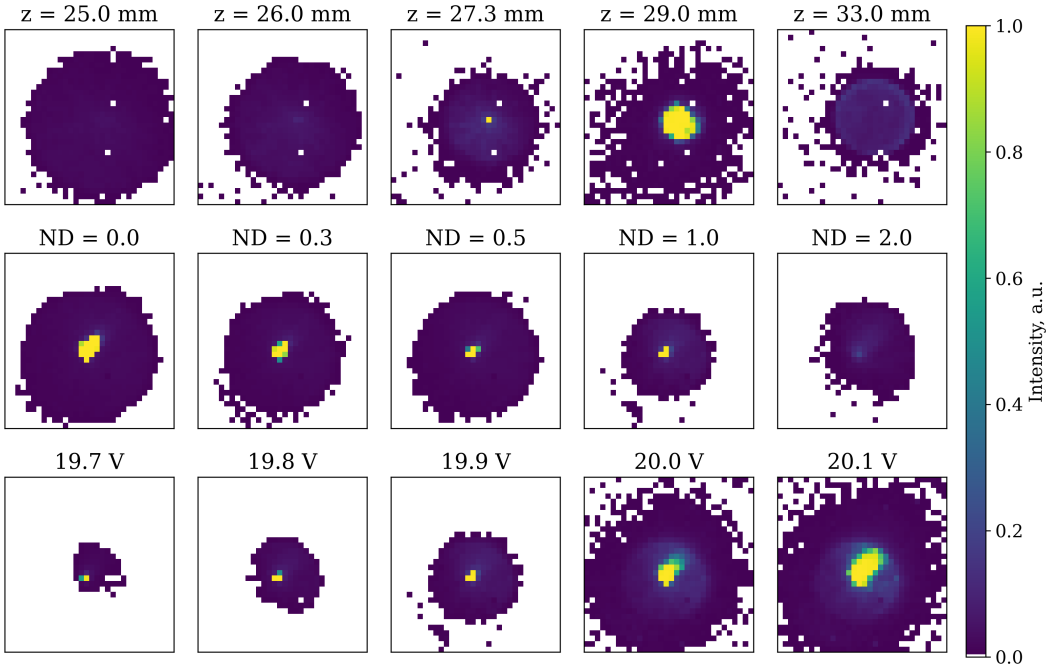
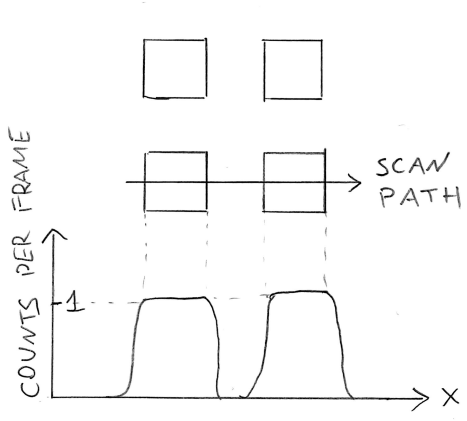
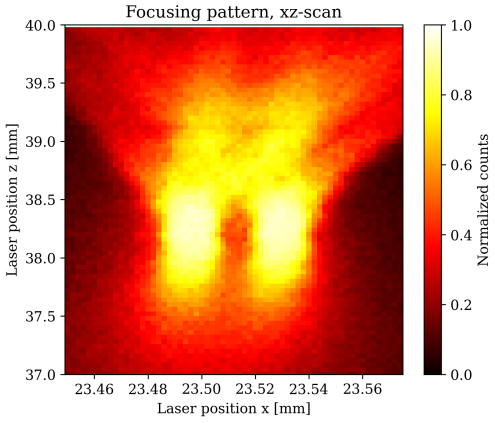


Figure 4.4: Impact of the laser system parameters and the bias voltage on the laser spot size measured by the dSiPM matrix, chip 1. 1000 frames were acquired for each set of parameters; the color scale shows the probability of a given pixel to register a photon.
Top: adjusting the focal distance. V_{bias} is fixed at 19.8 V, and ND = 1.5.
Middle: adjusting the intensity with filters. z is fixed at 28.0 mm, and V_{bias} is at 19.9 V.
Bottom: adjusting the bias voltage. z is fixed at 28.0 mm, and ND = 1.0.



(a) Expected hit probability as a function of laser position x. The laser spot is moved across two SPADs within a single pixel, schematically shown on top.



(b) Results of the knife-edge scan performed with chip 4. Hit probability as a function of laser spot position x and lens position z is shown with color.

Figure 4.5: Focusing the laser for dSiPM studies with the knife-edge technique. The optimal z value does not match one shown in Figure 4.4, since a different laser mount was used.

An example of the fine focusing is shown in Figure 4.5b. Each row on the graph shows one scan along x, while the vertical axis represents the focal distance. The hit probability at a given (x, z) is shown with color. At $z \simeq 38.2\,\text{mm}$ the sharpest transition and thus the optimal focus is achieved. The sharpness of the transition between the efficient and inefficient regions is determined by two convolved factors: the finite laser spot size and also by the structure of SPADs themselves, as there is no optical window on the surface that would fully block the light in the peripheral regions of the SPAD ensuring a sharp transition in efficiency between regions. The characteristic asymmetry of the beam in z can as well be observed in this measurement. If the lens is moved further away from the dSiPM (z decreases), the image of each individual SPAD smoothens out because of the Gaussian-like profile of the beam. However, if the lens is moved closer (z increases), a diverging pattern can be seen in place of every SPAD, which hints on the ring-like intensity profile.

4.4.2 TDC characterization

The delay-locked loop in fine TDCs of the dSiPM shows certain deviations from the nominal specifications. By design, all delays are adjustable. A dedicated feedback circuit, a *phase comparator*, tunes all delays at once to ensure that the total delay of all 32 elements equals one cycle of the system clock. On the dSiPM prototype, however, the delay elements are out of specifications with longer delay values than specified, and the possible adjustment range is not enough to ensure a proper functioning of the phase locking. Additionally, there is a variance in width between individual bins.

The actual widths of individual bins, however, can be characterized and accounted for. If the dSiPM measures timestamps for hits that have a uniform random distribution over time, the same occupancy is expected in each fine TDC bin, assuming that their widths are equal to 1/32nd of the clock cycle. If the bin widths are different and larger than expected, the following stands:

- the occupancies of the few last bins are 0, as the signal propagation never reaches them within one clock cycle;
- for all non-zero bins, the observed occupancy is proportional to their effective width;
- the total duration of all non-zero bins adds up to one system clock cycle.

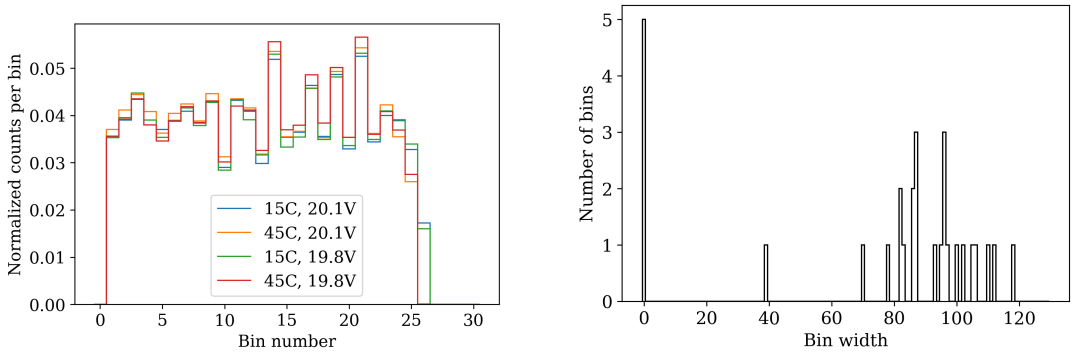
This allows to derive an effective width of each individual TDC bin. With the non-uniformity of the DLL bins taken into account, the Eq. 4.6 transforms to:

$$\text{Timestamp} = \frac{\texttt{bc}}{3\ \text{MHz}} + \frac{\texttt{ct}}{408\ \text{MHz}} + \sum_{i=0}^{\texttt{ft}\ -\ 1} (\text{DLL bin width})_i + \frac{(\text{DLL bin width})_{\texttt{ft}}}{2}. \tag{4.7}$$

Experimentally this can be performed by acquiring frames with noise events. Technically, for *random noise* the measured timestamp distribution reflects the *expectancy* of the time-to-next-event and thus is exponential, given the statistical thermal origin of the noise. At the time scale of a single system clock cycle, which is relevant for fine TDCs, the exponential dependency can be nevertheless neglected and the distribution of the noise events can be assumed uniform. This technique was previously presented in Ref. [49] (with dark noise counts), and also in Ref. [117] (with ionizing particles at the DESY II test beam, which in this case can be considered uniformly and randomly distributed).

Such a qualification is carried out for every TDC in every chip. Below, a qualification of TDC of quadrant 2 of chip 1 is shown as an example. Figure 4.6a presents the noise occupancies of individual fine TDC bins, measured at different voltages and different temperatures on the chip, thus at various noise rates. It is clearly observed that similar occupancy trends persist in all series. The calculated widths for all 32 bins are shown in Figure 4.6b. Five bins, namely, bin 0 (not used by design) and bins 28–31 are not usable and therefore have calculated widths of 0. Bin 27, at which the clock cycle ends, is not used fully and has an effective width of ∼40 ps. For the remaining 26 bins the widths are distributed in the range of 70–120 ps with a mean of 96 ps and a variance of 14 ps.

While the individual bin widths vary from chip to chip and from quadrant to quadrant, the total number of usable bins and their mean calculated width are similar. The digitization contribution to the dSiPM time resolution can be then estimated as (Eq. 4.4): $\sigma_{TDC} \simeq 28$ ps.



(a) Noise occupancies of all bins, measured at 4 different conditions.

(b) Calculated bin widths. Average over all series, presented on Figure 4.6a.

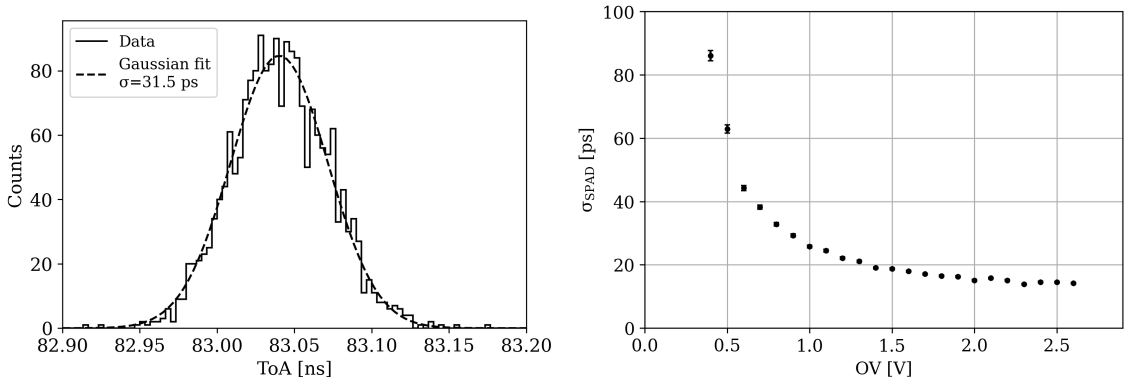
Figure 4.6: Characterization bin widths of fine TDC in quadrant 2 on chip 1.

4.4.3 Intrinsic SPAD time resolution

Intrinsic timing characteristics of the SPADs can be measured on the dedicated test SPAD arrays, located at the dSiPM periphery. Analog outputs of these test structures are read out directly with external instruments, therefore the contribution of the on-chip digitization electronics is thus excluded. For this measurement, the test SPADs are uniformly illuminated with a de-focused laser beam. The laser is triggered by an external clock generator. Analog outputs of the test array are picked up with a digital oscilloscope [69] with a bandwidth of 4 GHz and a sampling rate of 40 Gsa/s. The time-of-arrival (ToA) for each signal is measured by constant fraction discrimination at 50% of the amplitude, using the laser trigger signal as the timing reference of the oscilloscope. The variance of the ToA distribution then represents the intrinsic time resolution of a SPAD.

The ToA distribution for a SPAD, biased at OV = 0.8 V (as used in measurements in Section 4.4.5), is shown in Figure 4.7a. The timing resolution of as a function of overvoltage is shown in Figure 4.7b. As the overvoltage increases, the timing resolution improves and reaches ∼15 ps. This can be explained by the fact that higher voltage results in higher gain and higher signals, therefore improving sensor-dependent components of the time resolution, such as σ_{Jitter} and σ_{Timewalk}. This measurement was previously presented in Ref. [119].

It should be noted that this measurement demonstrates the best-case scenario for SPADs, as the amount of incoming photons is high, and they are distributed across the entire SPAD area. Hence, the photon that causes the fastest signal within the most optimal region in a SPAD would be taken into account for estimating the time resolution. For smaller and more localized charge depositions, such as from single photons (as presented Section 4.4.5) and minimum-ionizing particles [115, 116, 117], a significantly slower response and overall worse time resolution may be observed in peripheral regions of SPADs.



(a) Timestamp distribution for OV = 0.8 V.

(b) Time resolution as a function of overvoltage.

Figure 4.7: Intrinsic timing resolution of SPADs. Adapted from Ref. [119].

4.4.4 Matrix delay measurements

In order for a timestamp to be recorded when a hit occurs, a signal from a pixel needs to reach the TDC on the periphery through the *wired-OR* circuit. Therefore, timestamps measurement by the TDCs is affected by the signal propagation across the chip matrix. The physical length of the connecting traces is different, depending on a location of a given pixel within a quadrant. This results in an additional delay for pixels located closer to the matrix center and thus non-uniform timing characteristics of the matrix as a whole.

This non-uniformity of delays should be quantified, which can be done by analyzing the chip response to laser pulses, injected at a defined moment. The measurement was performed on chip 4 as follows:

- All pixels in the matrix except for one are masked. This ensures that the timestamp is defined by the only active pixel;
- The laser beam is shot at the unmasked pixel. The beam is defocused and kept at a high intensity to minimize in-SPAD variations;
- The laser is triggered synchronously with the frame clock, such that the pulse always is expected to come with the same delay w.r.t. the beginning of the frame, thus the same time-of-arrival is expected;
- Multiple frames are acquired with the dSiPM;
- Mean and variance of the time-of-arrival distribution are extracted;

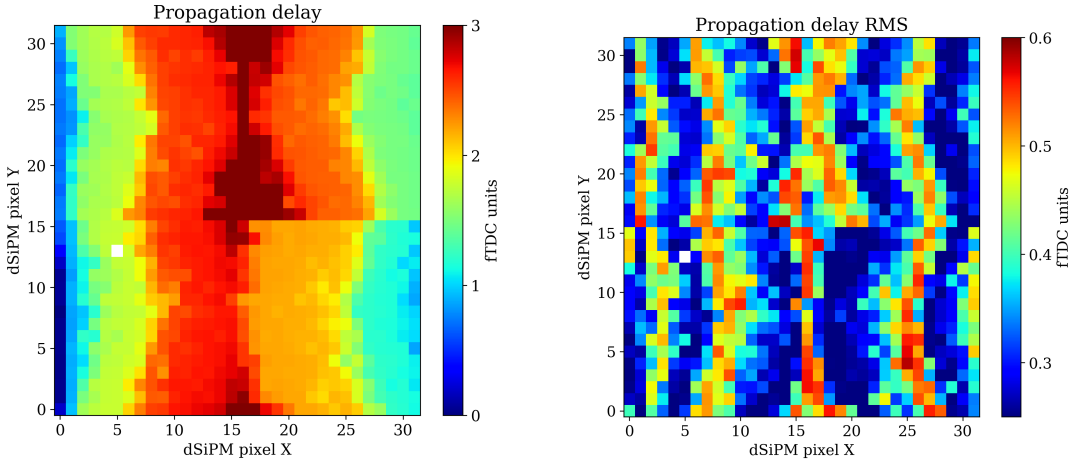
- The procedure is repeated for every pixel in the matrix.

The results of this measurement are used to produce an in-chip map of the propagation delay, shown in Figure 4.8a. A distinct structure is observed in the delay map. The delay is correlated with the physical distance to the TDC location; for the furthermost pixels, i.e. located in the matrix center, this delay reaches 3 counts of the fine TDC. Given that most of the events are typically contained in just one fine TDC bin, discrete regions with additional delays of 0, 1, 2, and 3 fine TDC counts are seen on the matrix.

Figure 4.8b shows the variance of the ToA distribution for every pixel. For distributions with a majority of counts in a single bin, the variance is in line with the mathematical estimate of $1/\sqrt{12} \simeq 0.3$. For pixels, where the delay is such that the ToA of the signal is close to a border between two bins and the distribution thus consists of two adjacent approximately equally populated bins, the variance increases to $2/\sqrt{12} \simeq 0.6$.

This data on the propagation delay can be used to correct the timestamps by post-processing. This, however, requires a certain knowledge on which exact pixel triggered the first and defined the timestamp in this frame. This is possible only if a single pixel (or a cluster of closely located pixels) was hit in this frame, or if additional information on the hit position is available. Such a correction was performed e.g. during test beam studies, presented in Refs. [115, 116, 117], where the spatial information on every particle passage is provided by the beam telescope.

Measurements presented in this section were previously published in Ref. [49].



(a) Mean of the ToA distribution in units of fine TDC bins. A baseline value is subtracted from every point to only show the difference in delays.

(b) Variance of the ToA distribution.

Figure 4.8: Signal propagation delay for every pixel in the dSiPM matrix (chip 4).

4.4.5 In-pixel timing resolution

In the measurements presented in the previous sections, the entire pixel of the dSiPM was uniformly illuminated, which gives no insights on variations of the chip performance within a single pixel or SPAD. Though, the design of the pixel cells implies a certain non-uniformity of the characteristics. These µm-scale features can be resolved thanks to the ability of the laser system

to produce very localized charge depositions. However, the obtained maps would have a certain smearing, defined by the finite size of the laser spot ($\mathcal{O}(\mu m)$).

Laser aim and device sensitivity

To study such variations of the detection performance within a single SPAD, the laser beam needs to be focused to a small spot and pointed at a single pixel within the matrix. Similar to the procedure presented in Section 4.4.4, all other pixels are disabled, and the measured quadrant timestamp is thus set by this exact pixel. For this measurement the dSiPM is operated at a bias voltage approximately corresponding to an overvoltage of 0.8 V in order to deliberately decrease the matrix sensitivity and to facilitate the ability to resolve in-pixel features (detailed in Section 4.4.1). This is a sub-optimal working mode for SPADs, as at this voltage they yield an intrinsic time resolution of $\sim$30 ps, whereas sub-20 ps can be achieved with higher bias (see Figure 4.7). Nevertheless, operating SPADs at OV = 0.8 V still allows one to study the characteristic trends.

In order to aim the laser, to benchmark the laser spot, and to match coordinates of the laboratory frame of reference to in-pixel coordinates, several iterations of two-dimensional scans are performed. A set of frames is recorded for each laser position and the chip sensitivity is defined as probability for the pixel under study to trigger:

$$\text{Probability} = \frac{\text{Number of frames } \textit{with} \text{ the triggered pixel}}{\text{Number of frames}}. \tag{4.8}$$

After finding the optimal parameters of the laser, a scan with fine steps of 1 µm and 7000 frames per position is performed for chip 4. This number of frames was found to be optimal from the point of view of the data buffering in the readout system. The result is shown in Figure 4.9. The four efficient regions are clearly observed, corresponding to the four SPADs within a pixel (as compared to Figure 4.3, right). The chip registers hits with a high probability within SPADs. The measured sensitivity in regions adjacent to SPADs is however nonzero due to a finite laser spot size and some intrinsic sensitivity right next to SPADs.

A minor shearing (few µm) of the left side of the map is related to a temperature change during the long measurement (see Figure 4.10), resulting in a thermal expansion of support structures. The scan was performed left-to-right, thus column number (i.e. position x) corresponds to measurement time. The temperature started to increase at the start of the data taking, but stabilized while the non-sensitive part ($x < -30$ µm) was scanned.

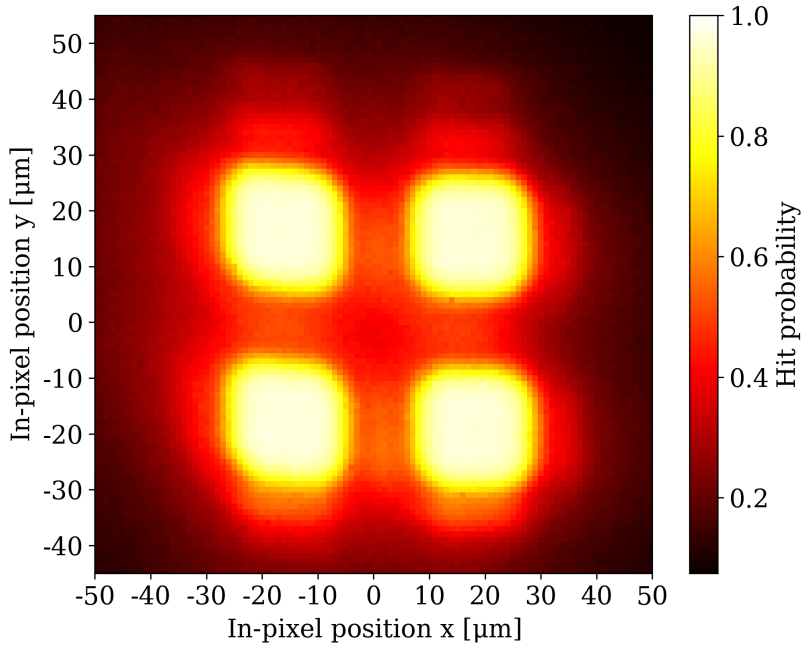


Figure 4.9: In-pixel sensitivity map of a single dSiPM pixel. Area of $100 \times 100\,\mu m^2$ is scanned with a step size of 1 µm.

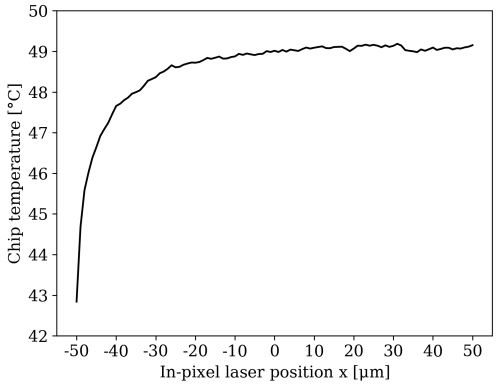


Figure 4.10: Temperature on the chip during the scan.

Time residuals and fitting strategy

For a timing detector, a time residual is generally defined as a difference between the true time of the hit arrival and the measured timestamp. Since in this measurement scheme the laser is triggered synchronously with the dSiPM clock, the measured ToA is equivalent to the residual with a certain offset, so here these two concepts are used interchangeably. The typical structure of the dSiPM time residual is shown in Figure 4.11a. The dataset shown in this figure represents the accumulation of all timestamp distributions recorded during the scan shown in Figure 4.9. Characteristic features of the time residual are:

- the peak (here at 222 ns) that corresponds to frames with detected photons;
- uniform background left to the main peak from frames, in which a dark noise hit occurred *before* the arrival of the laser pulse; the noise is uncorrelated to the chip clock;
- uniform background right to the main peak from frames, in which no photon was detected and a dark noise hit occurred *after* the laser pulse; thus probability density is lower than for noise events left to the peak.

The magnified view of the peak is shown in Figure 4.11b. Three separate components of the peak are:

- the central part (the narrow peak with a width of few dozen ps) arises from events with an avalanche instantly created upon an absorption of a photon; this is the desired working mode of SPADs and referred to as *fast* events; for quantitative estimates of the dSiPM timing resolution the fast part is fitted with a Gaussian;
- the shoulder right to the peak originates from events where an avalanche is preceded by a comparably slow drift of charge carriers or diffusion (1–2 ns), as initial the charge deposition happens at a distance from the avalanche region; the response is thus slower than in the optimal case; this contribution can be approximated with an exponential function; such events are referred to as *slow*;
- the origin of the shoulder left to the peak requires further investigation; empirically it can also be approximated with an exponential function.

The fit function f with six parameters $p_0 - p_5$ to quantify the different components of the time residual then reads:

$$f(x, \vec{p}) = p_2 \cdot \exp\left(\frac{-(x-p_0)^2}{2p_1^2}\right) + p_3 \left[\exp\left(\frac{-x}{p_4}\right)\theta(x-p_0) + \exp\left(\frac{x}{p_5}\right)\theta(p_0-x)\right], \quad (4.9)$$

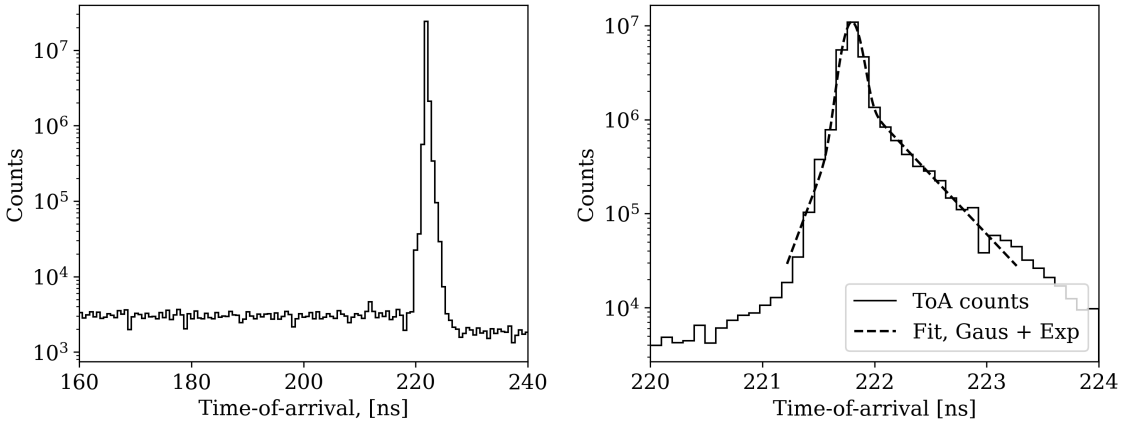
where:

- p_0, p_1, and p_2 define the Gaussian part and stand for its mean, variance and magnitude correspondingly;
- p_3, p_4, and p_5 are the magnitude (one for both) and the decay constants of the left and right exponents;
- θ is the Heaviside step function.

The timestamp distributions come with a binning strictly defined by the fine TDC bins, and the fit needs to be applied to a small number of bins ($\sim 10 - 15$). This limits the applicability of conventional fitting methods, such as least squares or χ^2, since in this use case they are not robust enough, i.e. prone to over-fitting to bin centers. The fit technique used is a Poissonian likelihood method, as it provides optimal results for sparsely binned datasets [120]. The optimal parameters are obtained by minimizing the negative logarithmic likelihood function

$$\mathrm{NLL}(\vec{p}) = -\sum_i (y_i \cdot \ln f(x_i, \vec{p}) - f(x_i, \vec{p})), \tag{4.10}$$

where x_i and y_i are the centers and occupancies of each bin.



(a) General structure of the residual.

(b) Main peak, fitted with function f (Eq. 4.9); the bin width corresponds to fine TDC bins.

Figure 4.11: Time residuals of the dSiPM accumulated over a full pixel area.

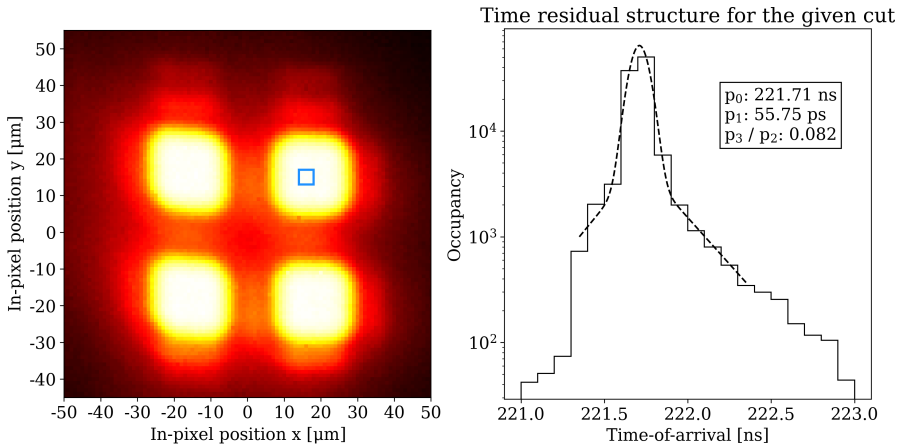
Results

By extracting the parameters of the time residual independently for every in-pixel position, one may study the effects of the SPAD design on the time resolution of the dSiPM. The general trend is shown in Figure 4.12. While the structure of the residual stays the same, its properties change as the laser spot is moved further away from SPAD centers: the main peak slightly shifts to the right (i.e. to larger ToAs) and its width increases, and the contribution of the slow component raises.

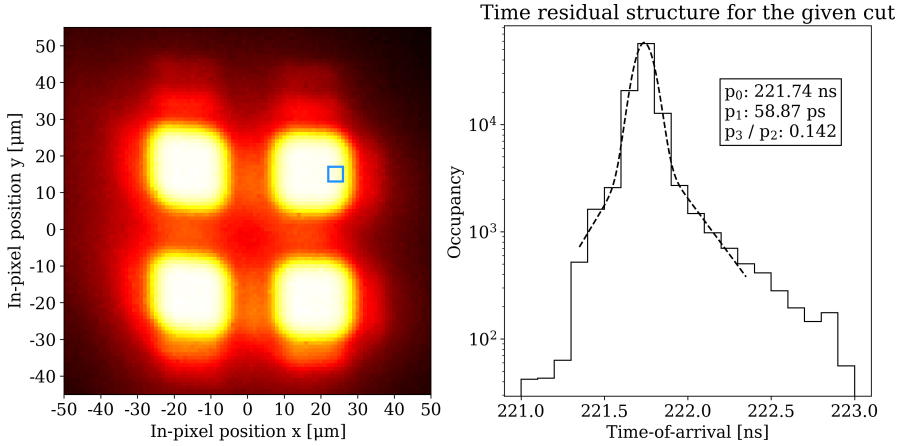
To quantify these findings, the fit parameters are evaluated as a function of the in-pixel position. The detailed in-pixel maps of the these parameters are shown in Figure 4.13. It should be noted that the data in the off-SPAD areas with low sensitivities are highly affected by statistical fluctuations due to the low number of events.

Figure 4.13a shows the mean ToA of the fast component p_0. The values observed are 221.70 ns in the SPAD centers with a uniform increase by 50 ps closer to SPAD edges. In off-SPAD regions the mean ToA additionally increases by $\sim$100 ps. A left-right asymmetry of the data in the off-SPAD area is possibly related to the temperature change on the chip.

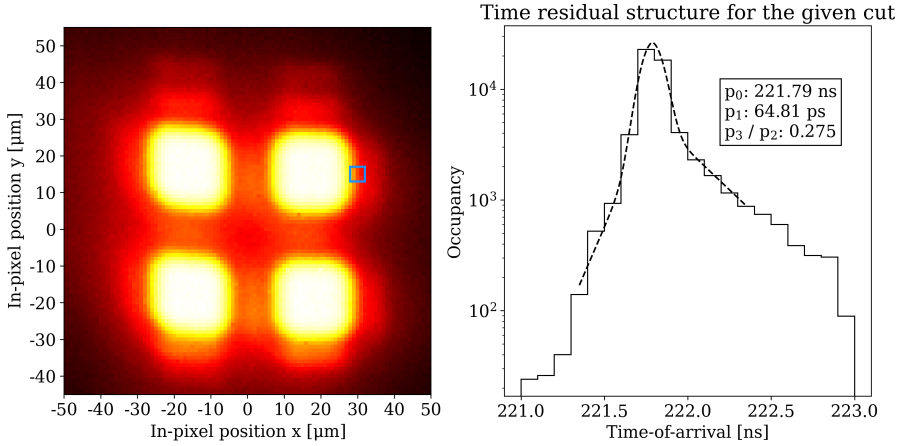
Figure 4.13b shows the Gaussian width of the fast component p_1, i.e. the optimal-case time resolution. Within the SPADs, the resolution ranges between approximately 50 ps and 57 ps. Since



(a) SPAD center.



(b) SPAD off-center region.



(c) SPAD edge.

Figure 4.12: Timing residuals for the data integrated over a selected in-pixel region. Selection regions are shown with blue rectangles on the left; timing residuals together with fits are shown on the right. Fit parameters p_0 and p_1 for the fast component are shown, as well as the relative significance of the slow component p_3 / p_2.

the used likelihood fit method does not intrinsically provide uncertainties, this variance of the parameter p_1 across adjacent data points can be used to estimate the uncertainty of the time resolution measurement as ± 4 ps. In the SPAD periphery, distinct regions with a wider (70–75 ps) peak are seen. For the two SPADs on the left, the optimal time resolution is slightly better than for the two SPADs on the right. This can be explained by the minor temperature change (as shown in Figure 4.10); slightly lower temperature during data taking for the left side of the plot results in slightly higher overvoltage and better intrinsic time resolution.

Figure 4.13c demonstrates the ratio of the count of the slow events to the fast events. The fraction of slow events is estimated with the given fit function as a ratio between the integral of the exponential parts only and the integral of the whole function:

$$\text{Slow ratio} = \frac{I_{\text{slow}}}{I_{\text{slow}} + I_{\text{fast}}}, \tag{4.11}$$

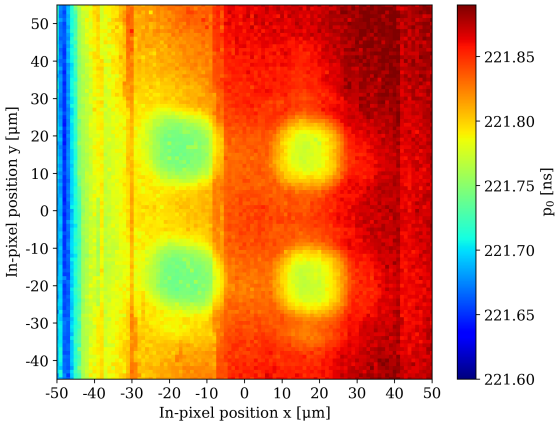
where

$$I_{\text{slow}} = \int_{p_0}^{+\infty} \mathrm{d}x \cdot p_3 \exp\left(-\frac{x - p_0}{p_4}\right) + \int_{-\infty}^{p_0} \mathrm{d}x \cdot p_3 \exp\left(\frac{x - p_0}{p_5}\right) = p_3(p_4 + p_5), \tag{4.12}$$

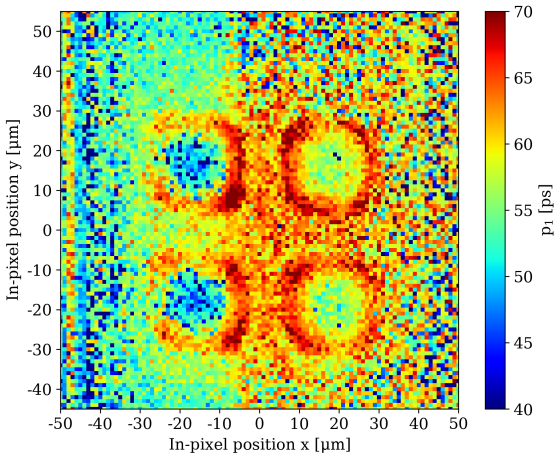
and

$$I_{\text{fast}} = \int_{-\infty}^{+\infty} \mathrm{d}x \cdot p_2 \exp\left(-\frac{(x - p_0)^2}{2p_1^2}\right) = p_1 p_2 \sqrt{2\pi}. \tag{4.13}$$

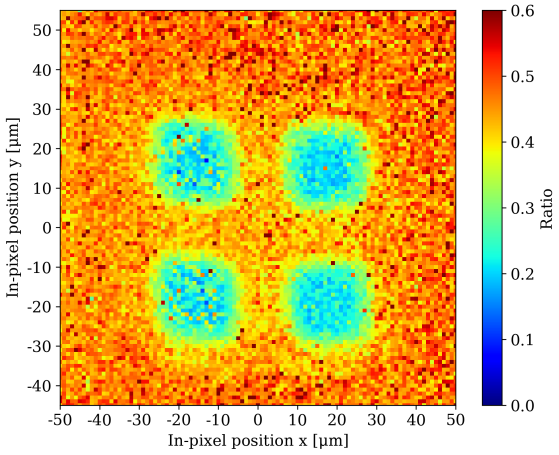
Within SPAD regions the calculated ratio stays below 20%, whereas in off-SPAD regions it typically 50% or higher. It should be mentioned that the classification of slow events provided by the fit function treats events as slow just by the fact of obeying exponential distribution regardless of their delay w.r.t. the main peak. Thus, several events with a relatively small delay can be considered slow. The occurrence map for delayed events selected by an explicit time cut is shown e.g. in Figure 4.14.



(a) Mean time-of-arrival of the fast component p_0.



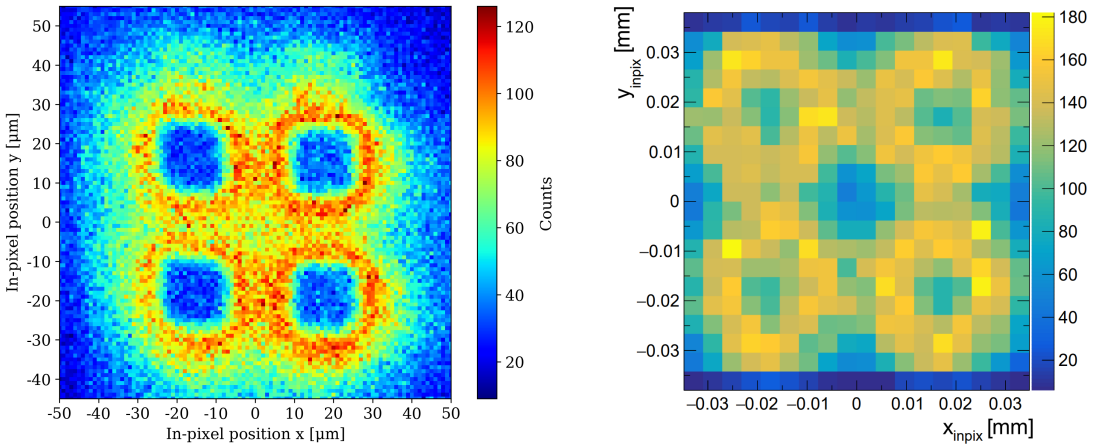
(b) Width of the fast component p_1.



(c) Fraction of slow events $I_{\text{slow}}/(I_{\text{slow}}+I_{\text{fast}})$.

Figure 4.13: In-pixel timing characteristics of the dSiPM, obtaining via fitting the time residual by function (Eq. 4.9) for every injection position.

Another metric of the presence of slow events, namely, an amount of hits with a delay of >1 ns w.r.t. the ToA of the laser pulse, is shown in Figure 4.14. Given the uniform distribution of hits across the pixel, delayed events happen almost exclusively in ring-shaped regions around each SPAD. The possible mechanism for that stems from the structure of the electric field in the sensor. The work [113] suggests the existence of a low-field depleted region that extends around the amplification region (see Figure 4.2). If the initial energy deposition happens within this region, the avalanche is preceded by a drift of charge carriers, leading to delays of the hit registration. Additionally, for depositions happening right next to the depletion region, charge carriers might still trigger an avalanche if they reach the depleted region via diffusion and then consecutively reach the amplification region.



(a) Scanning with a laser (this work).

(b) Test beam study [117].

Figure 4.14: In-pixel distribution of delayed events ($1\,\text{ns} < (\text{ToA} - p_0) < 5\,\text{ns}$).

Summary of Part I

This part of the thesis focused on development of the laser injection technique, which is a powerful tool for testing novel detectors. For employing this technique for studies of silicon sensors, a dedicated experimental setup, the *Laserbox*, was commissioned and improved by the author of this work. Throughout this thesis, the Laserbox was used in two independent projects: a development of a Monte Carlo simulation approach for the laser injection technique and characterizations of the novel DESY-developed digital SiPM prototype.

The Laserbox setup

A flexible multi-purpose setup for performing laser injection studies was commissioned and improved for this work. It has successfully demonstrated capabilities to perform pulse-shape measurements, to resolve µm-scale features of silicon sensors, and to characterize time resolution of detectors at a level of $\mathcal{O}(10\,\mathrm{ps})$. The setup has become a valuable asset for the detector development community at DESY Hamburg, and it was already employed in studies beyond the scope of this thesis, e.g. in Ref. [72].

Monte Carlo simulation of laser injection

An approach for MC simulations of the laser injection technique was established, employing the Allpix^2 framework. For this purpose, a software module for the Allpix^2, the `DepositionLaser` module, was developed in the course of this thesis. It enables modeling of interaction of pulsed laser beams with semiconductor detectors. This module is publicly available in the official framework release since version 2.4.0 [121], and several use cases of the module by detector R&D community were already reported in Refs. [77, 78].

The developed simulation approach allows one to simulate response of semiconductor detectors with good accuracy, which was shown by comparison with experimental data, obtained in the Laserbox setup. In the presented case, pulse shape measurements were performed on a baby-PS-s sensor. The Allpix^2 framework together with TCAD was used to model the experiment and to reproduce shapes of the induced electric signals. Typical scenarios with red and IR lasers were simulated. The simulation results show good agreement to the experimental data, which implies that all the major physical effects are properly accounted for. It should be mentioned that the simulation achieves this result independently and generally does not require a retrospective tuning, only taking the sensor design and basic characteristics of the experiment, such as bias voltage and laser wavelength, as inputs. Nevertheless, several *post hoc* adjustments were made in this particular case to compensate for certain parameters being unknown. First, a scaling was applied to the simulated pulses as the exact amount of photons delivered to the bulk by the laser system is

unknown; the found scaling coefficient, however, can be further employed for an estimate of this number of photons under given experimental circumstances. Second, parameters for the front-end circuit modeling were tuned using the experimental data.

This study provides a validation of the simulation approach. In perspective, the developed technique enables performing a cross-correlation between precise sensor tests by laser injection and MC simulations, allowing one to improve understanding of signal formation processes in silicon sensors.

Characterization of digital SiPM

The Laserbox setup was further employed for an extensive characterization of the novel digital single-photon pixel detector, the dSiPM. The main focus of this study was on timing characteristics of the device. New measurement techniques were implemented, allowing for studying the in-pixel features of a single-photon-sensitive matrix by depositing small amounts of charge in a localized area. Different contributions to the time resolution of the device were studied. The key findings are listed in the following.

1. The intrinsic time resolution of a CMOS SPAD (comprising all the sensor-dependent terms of Eq. 4.1) depends on the bias voltage and the location of the deposited charge w.r.t. the sensitive area of the SPAD. In the optimal case, i.e. sufficient voltage (OV $\gtrsim$ 1.5 V) and hits located in the center of the sensitive region, SPADs show time resolution of <20 ps.

2. The performance of the on-chip time-to-digital converters is subject to manufacturing variations. Due to an increase of the effective bin width, the measured time resolution is 28 ps.

3. An additional delay dependent on the position of a pixel within the matrix is caused by variations in lengths of traces, connecting pixels to the TDCs. This delay can be quantified and compensated for.

4. The total time resolution of a dSiPM pixel is found to be 53$\pm$4 ps for the optimal regions (within SPADs). This is generally in line with expectations based on the performance of individual components. These results are also in agreement with the previous studies, performed at the electron test beam [115, 116, 117].

5. For charge depositions close to SPAD borders, a significantly slower (1–2 ns) response may occur. This is caused by the structure of the electric field near the avalanche region in SPADs. This can be a limiting factor for the performance in applications requiring precise timing for detection of single photons or m.i.p.s)

The dSiPM prototype is a successful proof-of-principle for a CMOS-based SiPM with an individual cell readout. This work confirms the potential of the DESY digital SiPM for applications requiring precise timing and high spatial resolution in a single detector, such as particle physics experiments or time-resolved imaging. Additionally, the study of the sensor features at a micrometer level can provide valuable feedback for the design of detectors, based on CMOS SPADs.

Part II

Building and testing PS modules for the CMS Phase-2 Outer Tracker

Chapter 5

CMS experiment at the LHC

In this chapter, the Compact Muon Solenoid (CMS) experiment at the LHC are presented with a focus on the tracking detector. This work contributes to the upcoming upgrade of the CMS experiment, particularly, to the development of novel silicon sensor modules for the new Outer Tracker.

5.1 The Large Hadron Collider

Machine overview

The Large Hadron Collider (LHC) at CERN [14] is the most advanced and capable accelerator to date. It is designed to put the Standard Model (SM) of particle physics to an ultimate test. The discovery for the Higgs boson in 2012 [122, 123] has marked a new era in precision tests of the SM and searches for its extensions. The major points of the LHC proton physics program [124] involve:

- studies of the Higgs boson and its properties;
- precise measurements of the SM parameters;
- studies of known anomalies of the SM;
- searches for *new physics*.

The LHC provides unprecedented physics reach at the *energy frontier* by providing particle collisions at center-of-mass energy $\sqrt{s}$ up to 13.6 TeV. The LHC also extends the *precision frontier*, i.e. studying rare processes by providing instantaneous luminosity[1] $\mathcal{L}$ of the order of 10^{34} $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ and accumulating a high integrated luminosity of $\mathcal{O}(100\,\mathrm{fb})$.[2]

The LHC is a synchrotron with a circumference of 27 km, consisting of two independent beam pipes. The primary operation mode is proton-proton collisions with two proton beams circulating in opposite directions. Proton bunches are injected into the main accelerator at a kinetic energy of 450 GeV from the pre-accelerator complex, consisting of LINAC, Booster, Proton Synchrotron,

[1] Luminosity $\mathcal{L}$ [$\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$] or [$\mathrm{fb}^{-1}\,\mathrm{s}^{-2}$] is a measure of the density and the rate of the bunch crossings in the collider. By definition, if particles interact with a cross-section σ, the rate of such interactions is given as $\sigma\mathcal{L}$.

[2] Integrated luminosity L [fb^{-1}] is the integral of the instantaneous luminosity $\mathcal{L}$ over the operation time.

and Super Proton Synchrotron. The protons are then gradually accelerated to 6.8 TeV in the main LHC ring, with the magnetic field of the bending dipoles being simultaneously *ramped up* to 8 T. The beams are then focused at four interaction points (IPs) and crossed with each other to bring bunches to collisions. Once declared stable, the beams can be maintained for several hours. The layout of the LHC is shown in Figure 5.1.

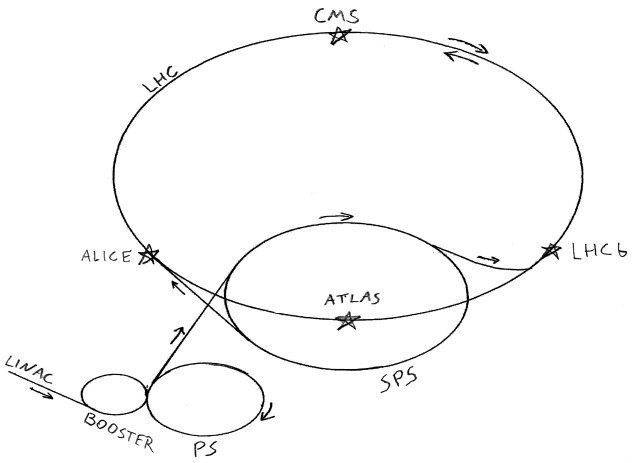


Figure 5.1: The layout of the LHC and its pre-accelerators used for proton injection (not to scale). The four main experiments at the interaction points are also shown. Adapted from [125].

The LHC can contain up to 2808 proton bunches (of $\sim 10^{11}$ protons each) in each ring, with a minimal spacing of 25 ns, resulting in a bunch crossing rate of up to 40 MHz. At every IP, focusing can be independently controlled to adjust the average number of proton-proton collisions per bunch crossing, called *pile-up* $\langle\mu\rangle$.

The LHC is also able to accelerate and collide ions (the primary type used is Pb^{82+}) with energies up to 6.8 TeV *per nucleon*. For the ion physics mode, the filling of the accelerator is limited to several hundred bunches of $\sim 10^8$ ions.

Experiments at the LHC

Four IPs are equipped with large-scale detectors. The detector designs differ to respect the particular physics goals of each experiment. The *general purpose* experiments **CMS** (Compact Muon Solenoid) [13] and **ATLAS** (A Toroidal LHC ApparatuS) [126] are located at opposing points of the main LHC ring. Both detectors are universal instruments with broad physics programs and are similar in design with one of the main differences being the configuration of the tracking detector magnets.[1] The sensitive layers surround the IP and aim to provide optimal resolution across almost a 4π solid angle, except for regions at a small angle w.r.t. the beam. Individual detectors and readout systems in ATLAS and CMS are optimized for data acquisition at a high rate of (40 MHz) and high pile-up $\langle\mu\rangle \simeq 30$. The **LHCb** (LHC beauty) experiment [127] is primarily aimed at physics of *b*-quark. The LHCb detector is built to provide good coverage and optimal resolution in the forward region for tracks of smaller density ($\langle\mu\rangle \simeq 1$), compared to ATLAS and CMS.

[1] Which is reflected in their names.

ALICE (A Large Ion Collider Experiment) [128] is optimized to study ion-ion collisions. Similarly to general purpose experiments, it offers high solid angle coverage. The ALICE detector is capable of registering tracks of extremely high multiplicity, characteristic for ion collisions, but at a lower bunch crossing rate.

High-Luminosity LHC upgrade

The timeline of the LHC operation is presented in Figure 5.2. The accelerator operates in continuous *Runs*, separated with *long shutdowns (LS)*. Runs are also interrupted with *(extended) year-end technical shutdowns ((E)YETS)*. During these, the experiments and the accelerator itself usually receive various upgrades.

The luminosity and the collision energy of the LHC have been being increased throughout its operation (see Figure 5.2). After the Long Shutdown 3, planned for 2026–2030, the LHC will enter the so-called *High-Luminosity* phase (HL-LHC) [129]. The HL-LHC will use beams of higher intensity and improved focusing methods. This will result in an increase of the instantaneous luminosity and the pile-up by a factor 5–7.5 of the nominal value. The upgrade of the accelerator will be complemented by significant upgrades of the experiments. Two main challenges that detectors will face at the HL-LHC are:

- Increased instantaneous luminosity. Detectors will have to deal with a higher particle flux therefore requiring higher granularity and better timing capabilities to resolve more tracks simultaneously. Additionally, improvements to triggering, readout, and acquisition systems are needed to handle the increased rate of the collected data.
- Increased integrated luminosity. During the High-Luminosity period, detectors will have to withstand more radiation exposure due to higher integrated particle flux. Some of the detector systems are already reaching the end of their operating life and require a complete replacement. The upgrade of these detectors demands a tenfold increase in radiation hardness.

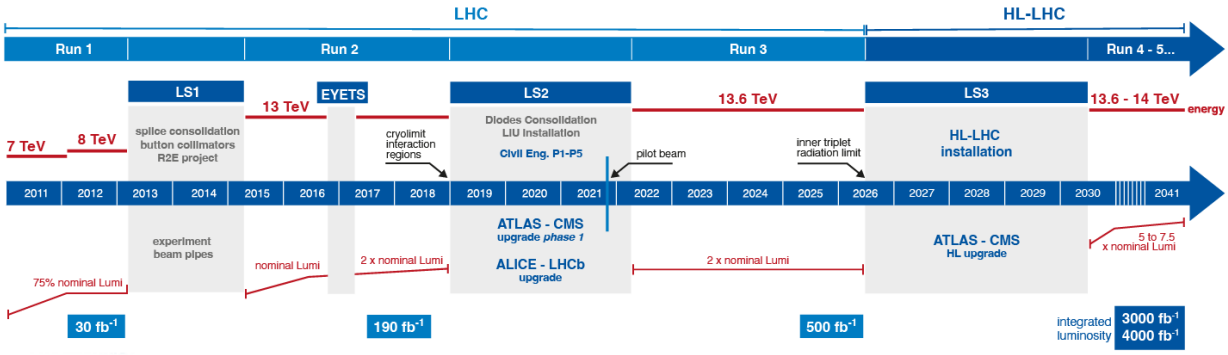


Figure 5.2: Projected timeline of the LHC operation as of early 2025. From [130].

5.2 Layout of the CMS detector

In this section, the current design of the Compact Muon Solenoid experiment is presented. The upgrades planned for the HL-LHC era (in the CMS framework commonly referred to as the *Phase-2* upgrade [131, 132]), are also discussed.

The layout of CMS is shown in Figure 5.3. A cylindrical coordinate system $\{R, \phi, z\}$ with the origin at the IP and the z-axis along the beam is used. Information on the polar angle θ is provided by the *pseudorapidity* $\eta \equiv -\ln\tan(\theta/2)$.

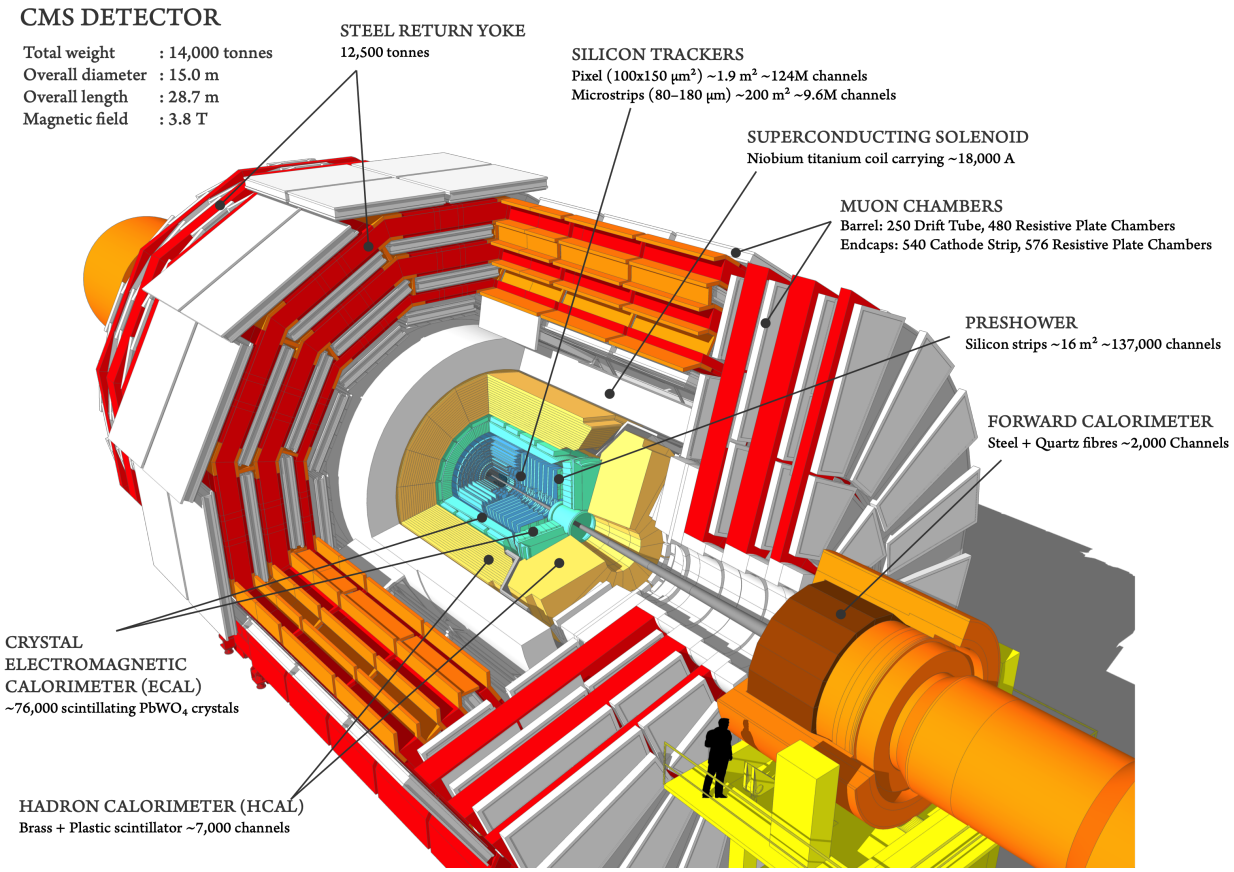


Figure 5.3: The current layout of the CMS detector. From [133].

5.2.1 Design philosophy

One of the main factors driving the design of the CMS is the performance of muon detection [134]. A major fraction of particles, created in proton-proton collisions undergo leptonic decays. In many cases, detecting leptonic decay modes yields optimal resolution for reconstructed parameters of the underlying process. Two processes of particular importance for the LHC physics program are the fully leptonic decays of the Higgs boson:

$$H \to ZZ^* \to l\bar{l}l\bar{l} \quad \text{and} \quad H \to W^+W^- \to \bar{l}\nu l\bar{\nu}.$$

Of leptonic final states, the muon channel has a cleaner signature and yields the superior measurement. This induces requirements on muon momentum resolution of $\Delta p/p \simeq 1\%$ at 100 GeV/c, which define the design of the CMS magnet. The magnet, being the key structural component of the setup, in turn defines the arrangement of the detector subsystems.

5.2.2 Magnet

The CMS detector uses a liquid-helium-cooled superconducting solenoid [135], providing a magnetic field of 3.8 T. The field with the solenoidal configuration bends trajectories of particles in the $R\phi$-plane (transversal to the beam). The inner bore of the magnet measures 5.9 m in diameter and 12.9 m in length. The tracking detector and the calorimeters need to be accommodated inside it, as the tracking detector is reliant on strong magnetic field, and having additional material (i.e. magnet coils) between these two systems and the IP is undesirable.

The magnet is surrounded by the iron yoke, that guides the return flux of the magnetic field (see Figure 5.3). The yoke consists of three layers with total thickness of 1.5 m, interposed with tracking planes of the muon system.

5.2.3 Muon system

The muon system of the CMS detector [136, 137] is comprised of gas detectors of different types, since gas detectors allow for a large area coverage at a moderate cost. A quadrant of the CMS detector with its muon system including parts planned for installation for the HL-LHC upgrade is shown in Figure 5.4.

The barrel part of the muon system consists of four layers, longitudinally divided into five wheels. Each wheel consists of 12 flat overlapping sections at an angle of 30° to each other (see Figure 5.3). Two types of gas detectors are employed in the muon barrel: *drift tubes (DT)* and *resistive plate chambers (RPCs)*. Each DT block has 12 layers of drift tubes, 8 with wires along the beam direction, measuring track curvatures in the $R\phi$ plane, and four layers with wires perpendicular to the beam, measuring tracks in the Rz plane. Drift tubes are chosen because they provide sufficient spatial resolution with a moderate number of readout channels by measuring the drift time of the charge carriers. Thus, a single layer of drift tubes with a pitch of 42 mm provides a spatial resolution of 250 µm, and the full muon barrel unit has spatial resolution of 100 µm. Each DT module is complemented by one or two RPC detectors. The RPC planes are primarily used for triggering and as a time reference because of their relatively good intrinsic temporal resolution[1] of 1.5 ns. The barrel RPCs have strips along the beam and offer spatial resolution of ∼1 cm.

[1] Compared to other types of gas detectors.

The muon endcaps currently use two types of detectors. Detectors in the endcap have a trapezoid shape, built into rings. The main tracking units in the endcaps are *cathode strip chambers (CSCs)*. Radial cathode strips provide resolution in ϕ, with wires determining the R coordinate acting as a multi-wire proportional chamber (MWPC). Every CSC block is comprised of six independent layers of gas chambers, which altogether provide a spatial resolution of 100 µm. Similarly to the barrel, the endcap also features RPCs of the same design serving the same purpose.

The longevity of all the gas chambers is estimated to be enough to stay functional during the entire HL-LHC operation. Also, no increase in granularity is required, so it generally stays unchanged. The front-end and readout electronics, however, will be replaced during the Phase-2 upgrade. Additionally, to increase the redundancy in the *forward* (= high η) endcap region, new detectors will be installed there: improved RPCs and *gas electron multipliers (GEMs)* which will contribute to both tracking and triggering.

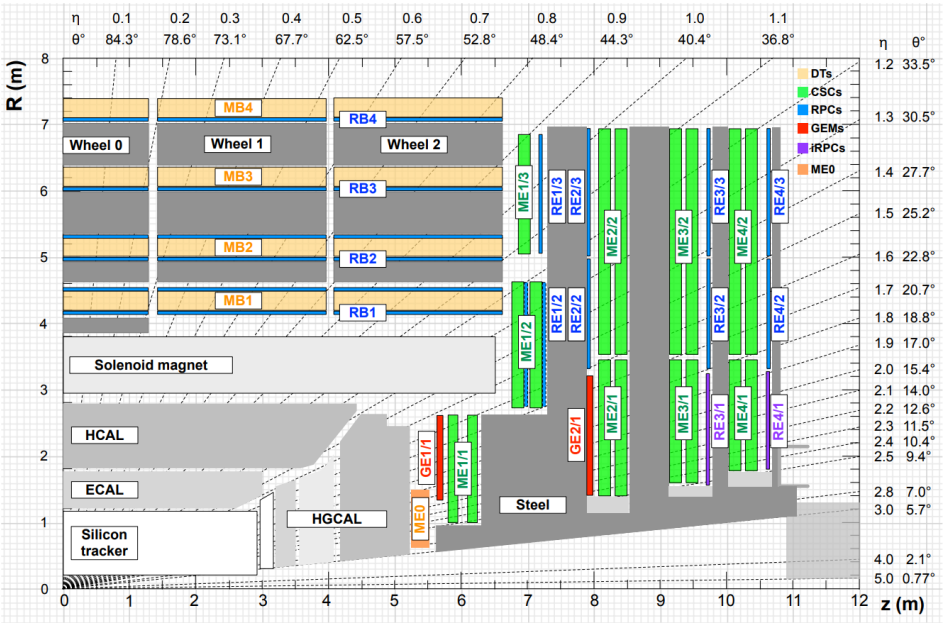


Figure 5.4: One quadrant of the CMS detector with the muon system highlighted, as planned after Phase-2 upgrades. From [137].

5.2.4 Calorimeter

The calorimeter of the CMS detector (see Figure 5.5) is located inside of the solenoid and consists of two subsystems: electromagnetic calorimeter (ECAL) and hadronic calorimeter (HCAL).

The **ECAL** [138, 139] is a *homogeneous*[1] calorimeter, consisting of ECAL barrel (EB) and ECAL endcaps (EE). Lead tungstate ($PbWO_4$) scintillator crystals are used as the sensitive medium. The crystals measure approximately $22 \times 22 \times 230$ mm^3 in the barrel and $28 \times 28 \times 230$ mm^3 in the endcaps. The crystals are arranged in a single layer and pointed at the IP. The scintillation light is read out with avalanche photodiodes in the barrel and with vacuum triodes in the endcaps. For

[1] All the absorbing material of the calorimeter is the sensitive scintillating medium.

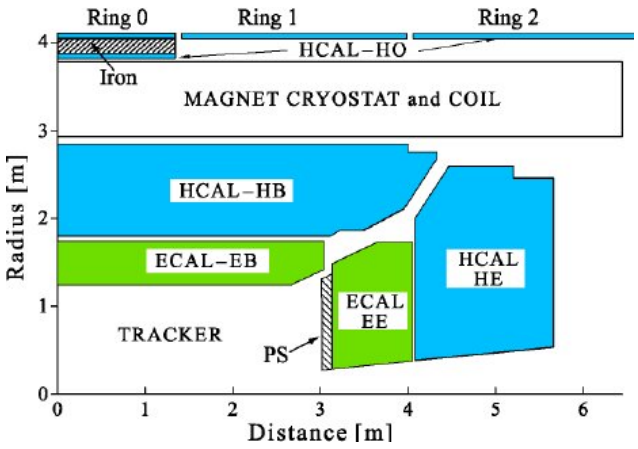


Figure 5.5: Current layout of the CMS calorimeters. Forward hadron calorimeter is not shown. From [138].

both electrons and photons at energies[1] of the order of 50 GeV, the ECAL typically provides an energy resolution $\Delta E/E \lesssim 3\%$. ECAL endcaps are supplemented with so-called *preshower* detectors ("PS" in Figure 5.5), composed of thin lead absorbers and two layers of silicon strip sensors with a pitch of $\sim$2 mm. The aim of the preshower detector is to distinguish between electromagnetic showers, originating from a single high-energy photon and photon pairs, produced in decays of neutral pions ($\pi^0 \to \gamma\gamma$).

The **HCAL** [140, 141] employs *heterogeneous* design, i.e. it consists of stacked passive absorber layers, interposed with sensitive medium. The sensitive layers are plastic scintillators with a thickness of 3.7 mm. The scintillation light is transported via wavelength-shifting fibers and picked up by SiPMs. The HCAL barrel (HB) consists of 17 sensitive layers, interposed with $\sim$5 cm thick brass absorbers while the HCAL endcaps (HE) have 19 scintillator layers with $\sim$8 cm absorbers. The transversal segmentation of the HCAL is $\Delta\eta \times \Delta\phi = 0.087 \times 5^\circ$ for the entire barrel and outer endcap regions, and is more coarse in the forward part of the endcap. Besides the barrel and the endcaps, the HCAL also includes two auxiliary parts, the outer barrel with two layers (HO) located outside of the solenoid and the forward calorimeter, located around the beam pipe behind the endcap return yokes (not shown in Figure 5.5).

The forward region is the most challenging for calorimetry because of a higher flux and higher energy of particles. During the LHC operation, the performance of the calorimeter in the forward region has deteriorated due to sustained radiation damage. Therefore, for the Phase-2 upgrade, the entire endcap calorimeter (preshower, EE, and HE) will be replaced with an entirely new high-granularity calorimeter (HGCAL) [142]. The HGCAL is a heterogeneous calorimeter with a total of 52 layers, using radiation-hard silicon sensors in the upstream and forward region (segmented into hexagonal pads of 0.52 cm^2 or 1.18 cm^2). The rest of the volume uses scintillating tiles (5×5 cm^2) read out by SiPMs. Both barrels of the calorimeter, EB and HB, are retained for the HL-LHC operation with front-end electronics on the EB being replaced.

[1] Typical energy of the products of $Z \to e^+e^-$ and $H \to \gamma\gamma$.

5.2.5 Tracker

The CMS was the first major experiment to implement an entirely silicon central tracking detector [143, 144]. The innermost part of the whole CMS detector, the *pixel* tracker, is located directly around the beam pipe. The pixel tracker is surrounded by the *strip* tracker.

Phase-0 pixel tracker

The region directly around the IP requires the most performing detectors in terms of both spatial resolution and capacity to measure particles at a high flux. Therefore, silicon pixel sensors were chosen for the inner tracker of the CMS detector.

The original pixel tracker (sometimes referred to as *Phase-0*) used n^+-in-n sensors of 285 µm thickness. For all the sensors, the standard pixel pitch is $100\times150\,\mu m^2$. This pitch is chosen deliberately to force charge sharing and improve the spatial resolution in the momentum-resolving plane. Thanks to charge weighting technique, the sensors are able to reach sub-20 µm spatial resolution, overcoming the binary resolution limit of $1/\sqrt{12}\times$ pitch. In the pixel barrel (BPIX), the detectors exploit the magnetic field of the solenoid, that causes charge carriers in the silicon bulk to drift at a *Lorentz angle* w.r.t. the electric field in sensors ($\vec{B}\perp\vec{E}_{\text{bulk}}$). To force charge sharing in the endcaps, the sensors in the endcaps are placed at an angle in a so-called "turbine" geometry. The readout chips (ROCs) are bump-bonded to the sensors, with each ROC having $80\times52 = 4160$ channels. A standard sensor is $18.6\times66.6\,mm^2$ in size and uses 2×8 ROCs. On the ROC borders, sensors have two rows or columns of pixels which are twice the normal size to ensure proper ROC connection.

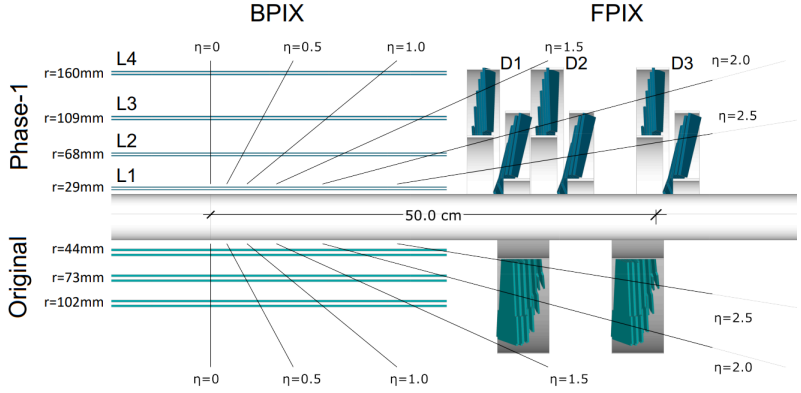


Figure 5.6: The layout of the CMS pixel tracker. Top: Phase-1. Bottom: Phase-0. From [145].

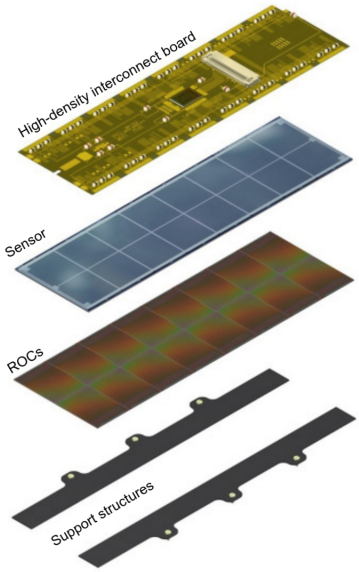


Figure 5.7: Design of a Phase-1 BPIX L2-L4 module [146].

The Phase-0 pixel tracker (shown in Figure 5.6, bottom) consisted of three barrel layers and two endcap layers. For the acceptance region $|\eta| < 2.5$, each particle originating in the IP would cross at least three silicon layers. The barrel consisted of 672 standard sensors (2×8 ROCs) and 96

half sensors (1×8). Each of the endcap discs consisted of 24 *"blades"*, with 7 sensors of various size (from 1×2 to 2×5 ROCs) per blade. The full pixel tracker consisted of 66 million pixels, covering a total area of 1.06 m^2.

Phase-1 pixel tracker

With the gradual luminosity increase of the LHC, the original pixel tracker was already reaching its limit on the particle flux. To allow further increase, in the middle of Run 2 the entire pixel tracker has been replaced with a new one, referred to as the *Phase-1* pixel detector [145] (shown in Figure 5.6, upper part). The Phase-1 detector uses sensors of a similar layout, but with improved ROCs, that can sustain detector operation and data readout at the increased particle flux. The pixel tracker now uses exclusively sensors with 2×8 ROCs, built into unified modules, consisting of a sensor with ROCs, a circuit board with auxiliary electronics, and support structures (see Figure 5.7). There exist three types of modules with slightly different designs: BPIX L1 (for the innermost barrel layer), BPIX L2-L4 (for other barrel layers), and FPIX (for endcaps). One additional detector layer was added to both the barrel and the endcaps, thus ensuring that every charged particle within the acceptance is now registered by at least four sensors. The innermost barrel layer BPIX L1 is located closer to the beam pipe at R=29 mm to improve the vertex resolution. With the total of 1856 modules, the Phase-1 pixel detector is almost a twofold increase in the sensor area (1.9 m^2) and the channel number (124 million), compared to the original tracker.

Being placed in an immediate proximity of the beam pipe, the BPIX L1 layer is subject to the most radiation damage. Thus, as scheduled, the BPIX L1 has been replaced during the LS2. As of Run 3, the replaced layer is already reaching the end of its operational life.

Phase-2 pixel tracker

For the Phase-2 upgrade [50], the pixel tracker is being completely redesigned. The Phase-2 pixel tracker will use sensors with a finer pixel pitch of 25×100 μm^2, providing 6 times the granularity of the original and Phase-1 pixel detectors and 2 billion readout channels in total. An improvement in radiation tolerance is achieved by using n^+-in-p sensors and radiation-hard readout chips, produced in a 65 nm CMOS process. While the barrel part of the Phase-2 pixel detector keeps the same layout, the endcaps will be extended to 12 layers (see Figure 5.12), and the pixel detector will cover the entire length of the tracking volume ($|z| < 265$ cm).[1] This will also increase the pseudorapidity coverage from $|\eta| < 2.5$ to to $|\eta| < 4$.

Strip tracker

The CMS strip tracker uses more than 24 thousand strip sensors, built into ∼15 thousand modules with 198 m^2 area coverage and more than 9 million channels. The strip tracker is organized as four parts (see Figure 5.8): the inner barrel (TIB), the outer barrel (TOB), the inner disks (TID), and the endcaps (TEC). For each track within the acceptance $|\eta| < 2.5$, the strip tracker typically provides four coordinate measurements in the inner section (TIB/TID) and six measurements in the outer section (TOB/TEC).

There are 15 different types of sensors, rectangular-shaped in the barrels and wedge-shaped in the endcaps. The strip pitch varies from 80 µm in the inner regions to 184 µm in the outer regions.

[1] Unlike the original and Phase-1 pixel detectors, that extended to $|z| < 50$ cm.

Strips are oriented to provide a precise momentum measurement: along z in the barrel and radially in the endcaps. Certain modules, however, have two sensor layers, placed at a so-called *stereo angle* of 100 mrad $\simeq 6^\circ$ w.r.t. each other and thus are also able to provide a coarse coordinate measurement in the direction along the strips, perpendicular to the momentum-resolving one.

The strip sensors are designed in a p^+-in-n layout. While this design is subject to type inversion, the sensors remain operable after sustaining radiation damage via depletion from backside. The strip tracker remained unchanged in the LHC Runs 1–3. In the framework of the Phase-2 upgrade, the strip tracker is being completely replaced with the new *Outer Tracker (OT)*. The design of the Phase-2 OT is presented in greater detail in Section 5.3.

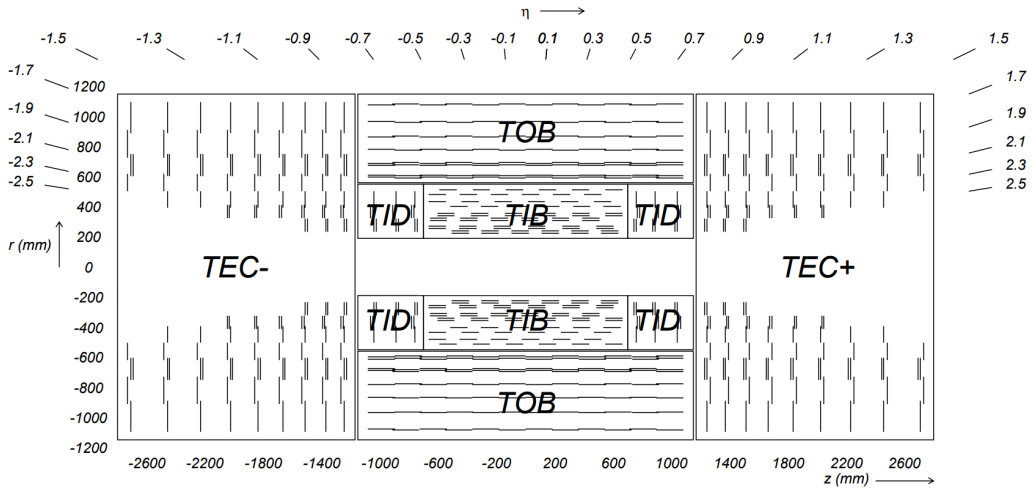


Figure 5.8: Layout of the CMS Strip Tracker. Stereo modules are shown with double lines. Adapted from [13].

5.2.6 Triggering and data handling

The volume of raw data acquired by the CMS detector is $\sim$1 MB per event. At the bunch crossing rate of 40 MHz it is impossible to transfer and store all the data collected by the detector. Therefore, filtering (or *triggering*) systems are used to only store data for selected events which are potentially interesting from the physics perspective. The selection criteria include: presence of certain particles (e.g. photons or muons), jets, or different combinations thereof; presence of particles with certain transversal momenta; global event observables (e.g. the total transverse momentum). The trigger system has two consecutive stages: *Level-1 Trigger* and *High-Level Trigger*.

The full data collected by detectors is stored in *buffers*, located in the front-end electronics of every detector. The depth of the buffers typically allows for storage of 128 consecutive events. Thus, a decision to store or to discard the next data fragment should be made within 3.2 µs. This low-latency filtering is accomplished by the **Level-1 Trigger (L1T)** [147, 148]. The L1T reduces the data flow by approximately a factor of 1000 with a typical output rate of 30 kHz and maximally allowed output of 100 kHz. The restrictions on latency define the implementation of the L1T: no extensive calculations are allowed; computing should be performed on dedicated low-latency hardware, such as FPGAs or ASICs; and the computing should be done as close to the detectors as possible to eliminate signal round-trip latency. Currently, only the data from muon system and

calorimeters is used in the L1T. After the Phase-2 upgrade of the trigger system [149], the maximal L1T output will be increased to 750 kHz, and the extended front-end buffers on detectors will allow for a more relaxed latency of 12.5 µs. Additionally, the Outer Tracker will be included in L1T by providing coarse but fast measurements of transversal momenta of charged particles (as discussed further in Section 5.3).

The data accepted by the L1T is passed further to the **High-Level Trigger** (HLT) [150, 151]. Given the lower data flow and higher latencies allowed (hundreds ms), much more convoluted algorithms can be employed. The higher latency tolerance also allows to implement the HLT as software running on commercial CPUs. The HLT performs a lightweight *event reconstruction* for every bunch crossing, i.e. it uses all the raw data to reconstruct particles' properties and to disentangle tracks, originating from different proton-proton collisions within one bunch crossing. The HLT can therefore apply more complex selection criteria, e.g. cuts on kinematic parameters of the collision products. The HLT provides the data reduction by a factor of 100 with a typical output rate of 1 kHz.

After passing the HLT, data is stored and packaged locally at CMS to be sent to the CERN Tier-0 computing center, where reconstruction is performed [152]. The reconstructed datasets are then distributed to Worldwide LHC Computing Grid centers for further storage, processing, and analysis [153].

5.3 CMS Phase-2 Outer Tracker

5.3.1 The concept of $p_{\mathrm{T}}-$discriminating modules

The design of the Phase-2 Outer Tracker (OT) [50] is strongly influenced by the necessity to contribute to the L1T by fast measurements of the transverse momentum p_{T} of the charged particles. Given the magnetic field of the CMS magnet of 3.8 T, a transverse curvature radius r_{T} for a particle of a unit charge (as follows from Eq. 1.4) reads

$$r_{\mathrm{T}}[\mathrm{m}] \simeq \frac{p_{\mathrm{T}}[\mathrm{GeV}]}{1.14}. \tag{5.1}$$

Assuming a cylindrical detector layer of a radius R, a charged particle with a transverse momentum p_{T} would cross the detector at an angle α_{T}, measured between the projection of the trajectory to the transverse plane and the normal to the detector plane:

$$\sin(\alpha_{\mathrm{T}}) \simeq 0.57 \cdot \frac{R[\mathrm{m}]}{p_{\mathrm{T}}[\mathrm{GeV}]}. \tag{5.2}$$

The detector can then provide a low-latency p_{T} estimate by measuring the angle α_{T} in every detector module. This is achieved with the help of double-layer modules. In such a module, two sensors are arranged one on top of another at a certain distance such that their strips or pixels are aligned to each other. Both sensors are connected to a common set of front-end chips. The module is able to estimate the α_{T} based on the coordinate offset between hits in its top and bottom sensors. The scheme is shown in Figure 5.9.

Since events with high p_{T} are of interest, the module is designed to select tracks with a small angle. Hits on two planes correlating within a certain programmable *window* signify a passage of a particle with a p_{T} above a given threshold. For every particle passing this selection criterion the module outputs a special data instance called a *stub*. Stubs are passed to the L1T at the bunch

crossing rate of 40 MHz and with low latency, and the L1T is able to use the stub data to make trigger decisions.[1] Meanwhile, full raw data is stored in buffers and is transmitted only if a L1T accept signal is received by the module.

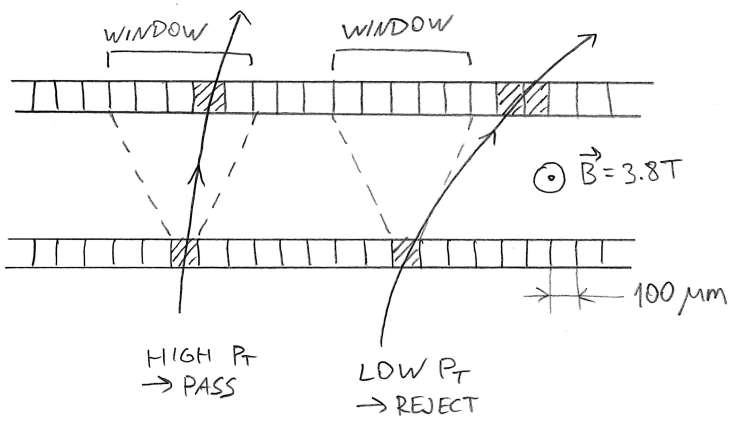


Figure 5.9: The concept of p_T-discrimination. The passage of particles with different p_T through a stack of sensor is shown. Adapted from [50].

A p_T selection threshold of 2 GeV is expected to be used since it provides the optimal balance between the trigger sensitivity and the required bandwidth for stub readout. For each module, depending on its location in the tracker, the sensor spacing and the stub window are chosen to ensure a uniform p_T threshold.

5.3.2 Module design

Unlike the original strip tracker, the Phase-2 OT uses only two major types of modules, as the design of p_T-modules is significantly more complex than of previously used ones. As the original strip tracker, the OT needs to provide a coarse coordinate measurement in the non-momentum-resolving direction as well. However, the concept of stereo modules is not viable, since sensors in each module must be parallel to each other for the p_T discrimination. Instead, several sensors in the inner layers are also segmented along the strips, and these segmented readout channels are referred to as *macro-pixels*. All sensors in the modules use a radiation-hard n^+-in-p design and have a nominal active thickness of 290 µm.

There are two types of OT modules: strip-strip (2S) and pixel-strip (PS). In total, the Phase-2 Outer Tracker will contain 7680 2S modules and 5616 PS modules.

PS modules

PS modules offer resolution in both coordinates. The top sensor of a PS module (called *PS-s*) is a strip sensor 98.1×49.2 mm^2 in size and segmented to 960×2 strips of 100 µm × 23.5 mm. The top sensor is read out with *Short-strip ASIC (SSA)* located at the edge of the sensor. The bottom sensor *PS-p* is 98.7×49.2 mm^2 and has 960×32 macro-pixels with a pitch of 100×1467 µm^2. The 16 readout chips in a 8×2 arrangement called *Macro-pixel ASIC (MPA)* are bump-bonded on the PS-p

[1] Nearly 80% of the available bandwidth will be used to transmit stubs to the L1T.

surface. Similarly to modules currently used in the pixel tracker, PS-p have double-pitch pixels of at the borders of the readout chips. A PS-p sensor with the attached MPAs is referred to as a *macro-pixel subassembly (MaPSa)*.

Both sensors in PS modules have a binary readout. Thresholds for PS-p and PS-s sensors are configured in the MPA and SSA chips respectively. Per-channel trimming of the threshold is available to ensure uniform threshold level across each of the sensors. Front-end chips of both sensors can perform *clustering* – recognition of groups of several adjacent strips/pixels being hit in a single event. The SSA chips transmit data on PS-s hits and clusters to the MPA chips. Similarly, the MPA chips perform clustering of hits from the PS-p sensor. Then, the MPA chips execute the angle discrimination logic using information on clusters on both sensors and produce stubs.

The mechanical design and the components of a PS module are shown in Figure 5.10. The engineering goal is to have a compact and lightweight design, while ensuring electrical insulation of the sensors, as they are kept at a bias voltage up to 800 V. Additionally, good thermal contact needs to be maintained to provide proper heat transfer from the sensors to the cooled support structure. A module is built on a carbon-fiber-reinforced polymer *baseplate* with a thickness of $\sim$200 µm that acts as a structural component. The baseplate has three holes with metal inserts which are used to position and mount a module in the OT. The baseplate is partially covered with a *Kapton* film for additional electric insulation. A MaPSa sensor is attached on top of the baseplate. A PS-s sensor is installed on top of the MaPSa with four aluminum nitride ceramic *spacers* in between. The spacers create the necessary separation between the sensors for the p_{T} discrimination to work, as well as provide electric insulation and heat dissipation for the PS-s sensor. Spacers are available in three different thicknesses to build PS modules with a sensor spacing of 1.6, 2.6, or 4.0 mm. The strip front-end chips and all the auxiliary electronic components are located on *hybrids* – subassemblies of flexible circuit boards and stiffeners. A PS module has four hybrids placed around the sensor stack and connected to each other:

- The *power hybrid (POH)* hosts a DC-DC voltage converter, that supplies power to all the chips on the module;
- Left and right *front-end hybrids (FEH-L and FEH-R)* hold eight SSA chips each that serve as a front-end for the PS-s sensor. The front-end hybrids also provide communication between SSAs and MPAs. Additionally, each FEH has one *Concentrator Integrated Circuit (CIC)* chip that transmits data to the readout hybrid;
- The *readout hybrid (ROH)* provides inbound (for configuration and control) and outbound (for data readout) communications for the module via an optical link. The communication is performed with a *Versatile Transceiver plus (VTRx+)* chip, paired with a serializer/deserializer chip called *Low-power Gigabit Transceiver (LpGBT)*. There are two available ROH versions with maximal data output rates of 5.12 Gbit/s and 10.24 Gbit/s, with the latter being used in modules where a higher particle flux is expected, i.e. located closer to the IP.

The assembly procedure for PS modules is discussed in detail in Chapter 6.

2S modules

The 2S modules (see Figure 5.11) are built of two identical strip sensors that measure $94.2\times102.7\,\mathrm{mm}^2$, approximately twice the size of sensors in the PS modules. Each sensor is divided into 1016×2 strips with a pitch of 90 µm and a length of 50.2 mm. In design and operation

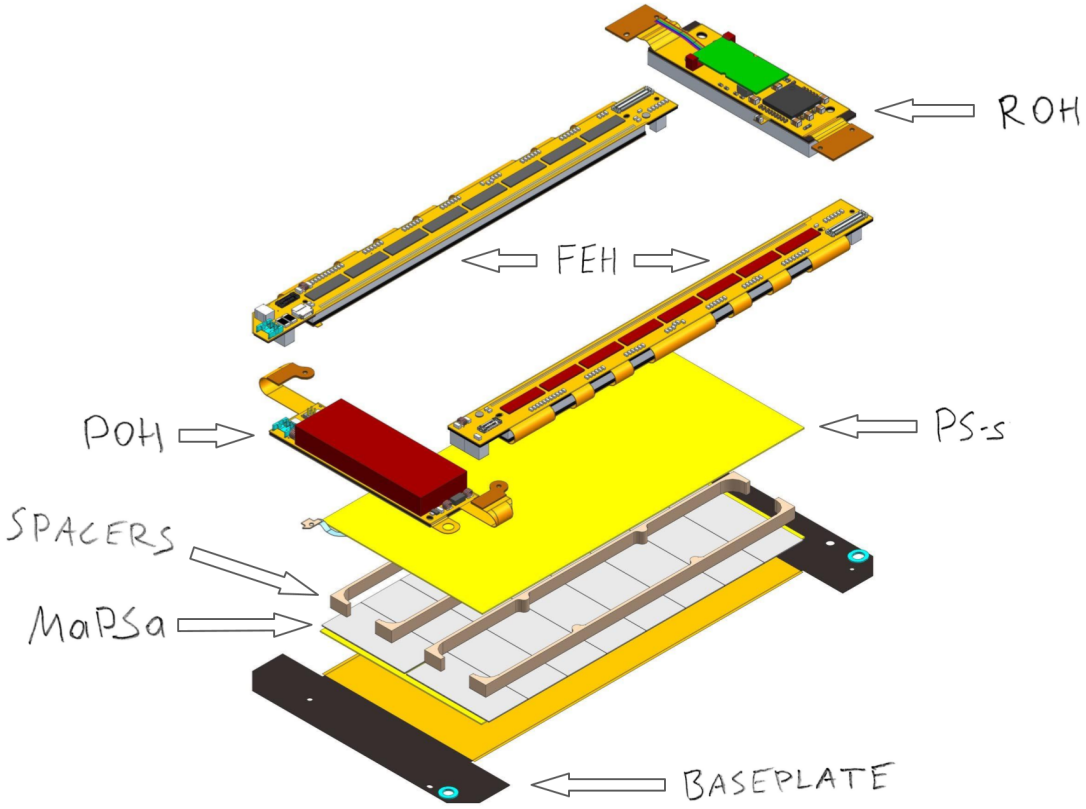


Figure 5.10: Design of a PS module (exploded view) [154].

aspects 2S modules are generally similar to PS modules. The design aspects, specific to 2S modules, are:

- The readout of both sensors and stub generation are realized through a single set of ASICs, called *CMS Binary Chips (CBCs)*, located on hybrids;
- Different hybrids are used: two 2S-specific front-end hybrids with CBC chips, and a *service hybrid (SEH)*, that merges power distribution and communication functionalities;
- 2S modules do not have a baseplate; three *bridges* made of aluminum carbon fiber composite, sandwiched between the sensors, are used as a support structure. The bridges also provide mounting points and cooling contacts;
- Available sensor spacings for 2S modules are 1.8 and 4.0 mm.

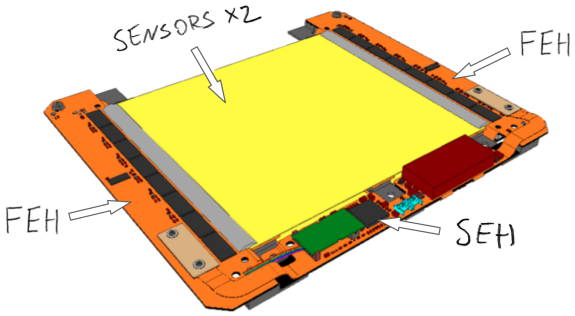


Figure 5.11: Design of a 2S module [50]. Front-end hybrids (FEHs) and a service hybrid (SEH) are placed at the edges of a sensor stack.

5.3.3 Tracker layout and mechanics

The general view of the entire Phase-2 Tracker is shown in Figure 5.12. The layout of the Outer Tracker is shown in Figure 5.13. The OT is divided into three parts: the inner barrel (TBPS), the outer barrel (TB2S), and the tracker endcap double disks (TEDDs). The OT provides pseudorapidity coverage up to $|\eta| \lesssim 2.5$. Every charged particle in this region would be registered by at least 6 modules (typically three PS and three 2S).

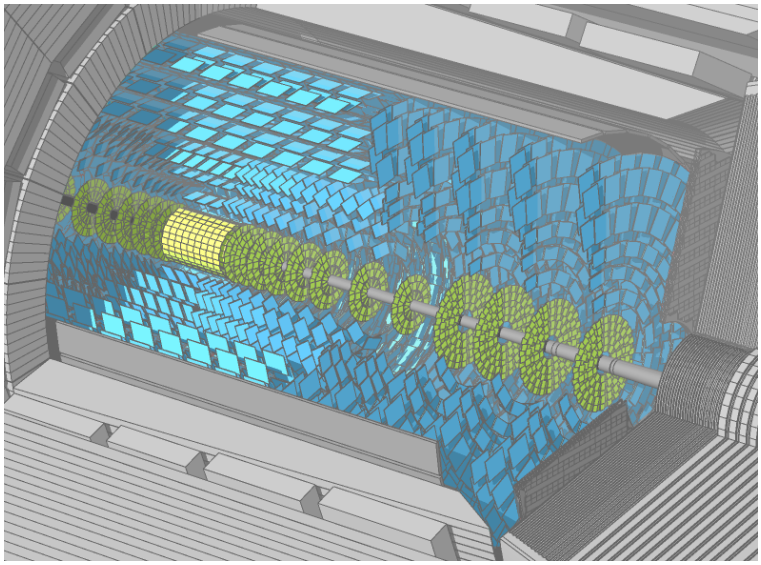


Figure 5.12: View of the CMS Phase-2 Tracker. The Pixel Tracker is shown in green, and the Outer Tracker in blue. Adapted from [155].

The inner barrel consists exclusively of PS modules and covers the region $200\,\text{mm} \leqslant R \leqslant 600\,\text{mm}$, $|z| < 1200\,\text{mm}$. There are three layers of modules (thus 6 sensor layers). In the center the arrangement is cylindrical, while the rest of the TBPS is arranged into tilted conical rings. This is done to reduce the required amount of modules, as well as to ensure that all modules are perpendicular to the tracks in order to optimize stub finding. The outer barrel ($600\,\text{mm} \leqslant R \leqslant 1200\,\text{mm}$, $|z| \leqslant 1100\,\text{mm}$) consists of three cylindrical layers of 2S modules.

Each endcap consists of five double-disks. Each double-disk consists of two disks with modules being placed on both sides, which ensures full coverage of the area with four overlapping module layers. Similarly to the barrel, the endcaps use PS modules in the inner region ($R \leqslant 700\,\text{mm}$) and 2S modules in the outer region ($700\,\text{mm} \leqslant R \leqslant 1200\,\text{mm}$).

The structures holding the modules are mainly made of carbon fiber composites to maximize their stiffness (to reduce vibrations and flex under load), while providing good thermal conductivity and keeping the weight as low as possible to reduce particle scattering. A pipe network for two-phase CO_2 cooling is embedded in the support structure to keep all sensors at the operational temperature of $-35\,^{\circ}\text{C}$. Additionally, the design allows to route services (e.g. optical fibers and low- and high-voltage supplies) to every module.

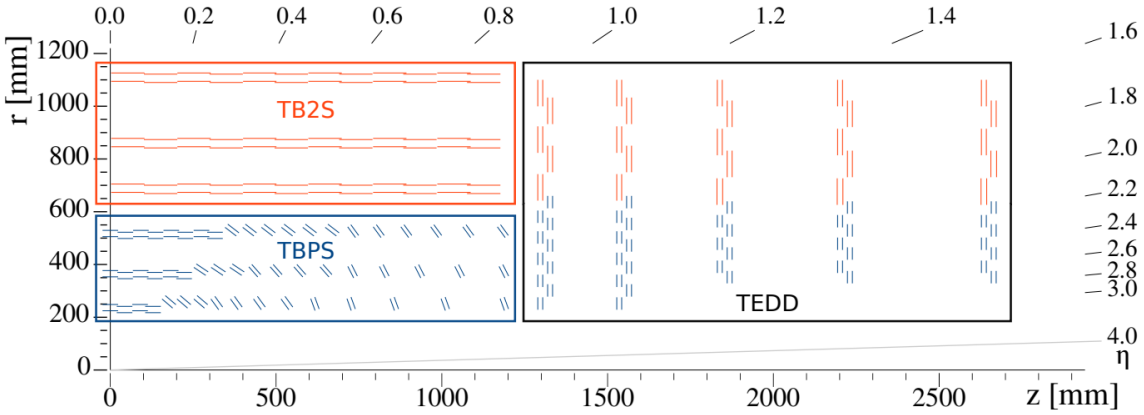


Figure 5.13: Cross-section of the CMS Phase-2 Outer Tracker. PS modules are shown in blue, and 2S modules in red. From [50].

Chapter 6

Construction of PS modules

In this chapter, the construction of PS modules is presented with a focus on the mechanical assembly procedures. Since the upgrade of the CMS experiment is an extremely complex endeavor, the workload is distributed among multiple research institutions contributing to the CMS experiment. For the Phase-2 upgrade of the Outer Tracker, DESY Hamburg acts as one of *assembly centers* for PS modules. This doctoral project contributes to designing, setting up and commissioning of the module production line, particularly to the step involving the robot-assisted mechanic assembly of the modules. Besides the module production activities, DESY Hamburg is responsible for the design and construction of one of the OT endcaps.

For the module assembly, the following coordinate system is used: the module is placed such that its long side with two inserts faces the operator; the x-axis is pointing to the right along the long side of the module; the y-axis points away from the operator; the z-axis points upwards; angle α denotes rotation around the z-axis.

6.1 Challenges of PS module production

The PS modules have to meet stringent specifications, both on operational characteristics and on mechanical properties. The following subsections cover the specific challenges of the module production with respective solutions.

Module functionality

To ensure proper functionality of every module being built, numerous quality assurance procedures are implemented. Various tests are performed both, on individual components before assembling them into modules, and on full modules, including:

- Electrical tests on sensors ensuring that their depletion characteristics and leakage current are within specification. These tests are used to identify and reject sensors that are prone to early breakdown or have an increased leakage current [156];
- Functionality tests on electronic subassemblies (i.e. MaPSas and hybrids) ensuring that they are fully operational [156];
- A *"skeleton"* test: communication is tested on a set of hybrids connected to each other but without sensors [157];

- After the assembly, every module undergoes a so-called *burn-in* procedure. The module is being thermally cycled between the room temperature and the operation temperature of $-35\,^{\circ}\mathrm{C}$ several times [158], with a series of functionality checks being performed repeatedly. This process also acts as a stress test to identify potential early failure of the modules and mechanical issues, related to the thermal stress.

Mechanical precision

The PS modules have to be as lightweight and compact as possible in order to fulfill the design requirements of the Outer Tracker on the module placement and the total material budget. This poses a significant engineering challenge and necessitates tight manufacturing tolerances of the components, and also puts stringent requirements on the assembly precision. Visual inspection and mechanical tests are performed on individual components to ensure compliance to the specifications.

Gluing is the main way of joining the module components together, as it ensures good thermal contact between the parts and does not require additional structural elements. Since module components do not feature alignment pins or other mechanical alignment features, the precise positioning of the parts for gluing them together is performed with help of dedicated tools. For some of the assembly steps such positioning is achieved with *jigs* – fixtures designed specifically for each step. Jigs are equipped with precisely machined pins, pushers, stops or vacuum holders that secure the components in place throughout the assembly process. Components usually remain fixed in jigs until the glue is fully cured.

Attaching a PS-s sensor to a MaPSa is an assembly step that requires a particularly high level of precision (exact tolerances are given in Table 6.1). This step defines a relative positioning of the PS-p and PS-s sensors, which directly affects the stub recognition. A transitional offset along the sensor plane can in principle be compensated for by the adjustment of a programmed stub window offset. Nevertheless, sensors need to be precisely positioned in x, as changing the stub window offset is not foreseen to be used for this purpose. Additionally, a misalignment in x may lead to wirebonding issues. A rotational offset, however, would result in effectively different stub windows at different positions on the module, which cannot be compensated by the electronics (see Figure 6.1). Thus, this rotation should be minimized by the assembly procedure. In order to meet these requirements, the assembly steps involving sensor placement are performed with a specialized automated manipulator (presented in detail in Section 6.3.1). Afterwards, the alignment of sensors is verified via optical metrology procedure (as discussed in Section 6.3.3).

Table 6.1: Alignment tolerances between PS-p and PS-s sensors.

Parameter	Description	Tolerance
$\lvert\Delta x\rvert$	Offset across the strips	50 µm
$\lvert\Delta y\rvert$	Offset along the strips	100 µm
$\lvert\Delta\alpha\rvert$	Rotation in xy-plane	800 µrad (max) 200 µrad (target)

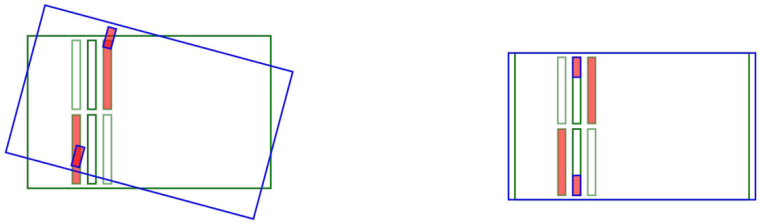


Figure 6.1: Impact of rotational misalignment between PS-p (shown in blue) and PS-s sensors (shown in green). If sensors are misaligned, as shown on the left, the two particles with perpendicular incidence would trigger channels highlighted in red. The readout electronics would however interpret the tracks at the top and the bottom halves of the module as if they had offsets of +1 and -1, since the sensors are assumed to be aligned to each other (as shown on the right). Not to scale. From [159].

Production volume

The assembly center at DESY Hamburg is expected to produce more than 1000 PS modules. To match the production goals and sustain the high quality, the module assembly procedure should be highly optimized. This includes simplifying individual procedures to make them more reliable and utilizing specialized tools and automation wherever possible. Several assembly steps (e.g. curing of glues) may take up to 2 days, while the required throughput of the assembly line is 4 modules per day. The amount of available instruments, quality control setups, and the planned use thereof must be enough to provide this production capacity.

Laboratory conditions

Exposing silicon sensors to dust particles and moisture is extremely undesirable, as it may result in damage. To minimize these risks, the module assembly workshop is situated in an ISO 6 cleanroom with controlled air temperature and humidity. Silicon sensors are also mechanically fragile, thus special tooling and handling procedures are to be used to avoid scratching or breaking the components.

6.2 Assembly sequence of the Phase-2 OT modules

While certain subassemblies, particularly, MaPSas and hybrids, are assembled by external vendors, the remaining components, i.e. baseplates, baseplate inserts, Kapton sheets, spacers, PS-s sensors, and HV-tails, are received as individual parts and assembled into complete modules in the assembly centers. The assembly sequence is as follows:

1. **HV-tail attachment.** A special flexible cable, called *HV-tail* (see Figure 6.2), is attached to the backside of the PS-s sensor to supply the bias voltage to it.
   1.1. The HV-tail is positioned on the backside of the PS-s sensor with a jig and secured with glue.
   1.2. Contact pads on the HV-tail are wire-bonded to the backside metalization of the sensor to provide electrical connection.
   1.3. Wire-bonds are encapsulated in a silicone compound [160] for mechanical protection.
2. **Baseplate preparation.**
   2.1. Three metal inserts are glued into the mounting holes at the corners of the baseplate. This is done using a jig with positioning stops and pins.
   2.2. A sheet of Kapton film is glued on top of the baseplate, in a position where the MaPSa will be attached later. Since the glue needs to be uniformly distributed over a large fraction of the baseplate area, a dosing robot (introduced in Section 6.3.2) is used. The Kapton sheet is positioned on the baseplate with a jig and is pressed with a weight piece to ensure the proper glue distribution.

   The assembled baseplate after these steps is shown in Figure 6.3.

Figure 6.2: HV-tail at the backside of the PS-s sensor.

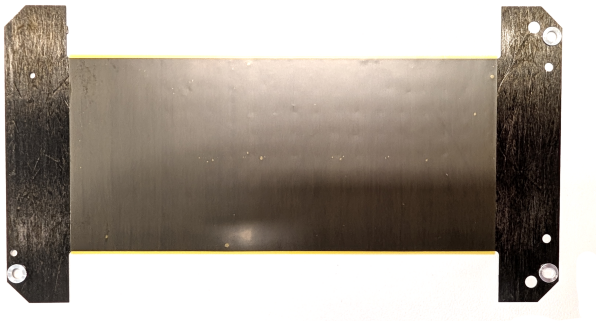


Figure 6.3: Baseplate with inserts and a Kapton layer.

3. **Bare module assembly.** A prepared baseplate, a MaPSa, spacers, and a PS-s sensor are built into a unit called a *bare module* (see Figure 6.4). This is performed in a single sequence, covered in detail in Section 6.3.
4. **Hybrid installation.**
   4.1. All four hybrids are glued to the bare module at once, using a jig (see Figure 6.5).

4.2. Sensors are electrically connected to hybrids with wire-bonds:
- PS-p backside to a high voltage pad on the power hybrid;
- each strip of the PS-s sensor to the front-end hybrids (to connect them to SSA chips);
- MPA chips to the front-end hybrids (to connect SSAs to MPAs).

4.3. Wire-bonds are encapsulated in a silicone compound [160] for electrical insulation and mechanical protection.

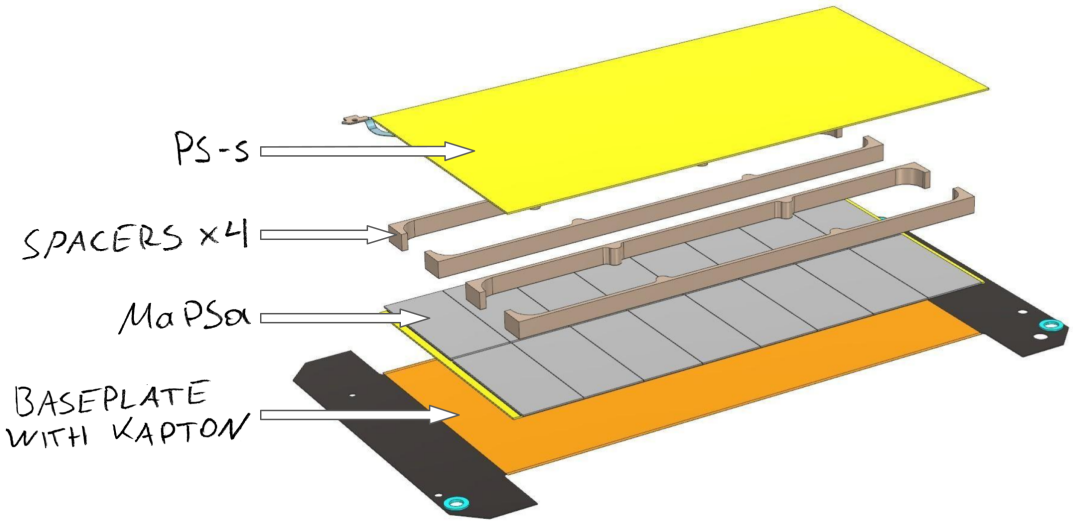


Figure 6.4: Exploded view of a bare module [161].

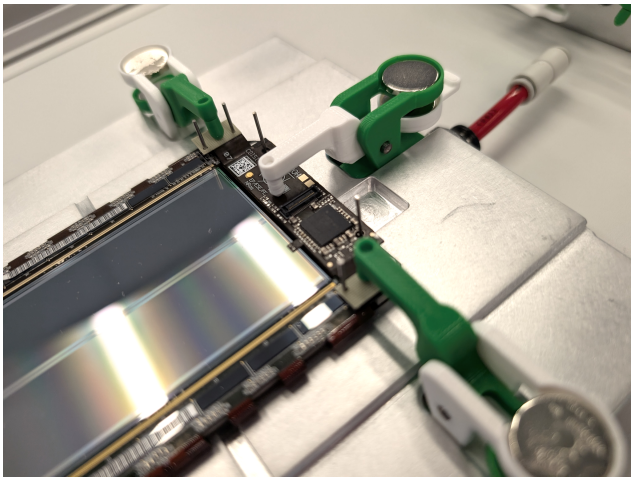


Figure 6.5: Hybrid installation process. The bare module is held on the jig with vacuum; the hybrids are secured with magnet-actuated pushers.

6.3 Bare module automated assembly

The bare module assembly relies on specific instrumentation and techniques, discussed in the following. In the course of the doctoral project, these were developed, improved and optimized in order to cope with the challenges of the module production.

6.3.1 The assembly machine

The assembly of bare modules is performed with a specially designed 4-axis CNC machine with vacuum holders and a microscopic camera with pattern recognition algorithms. Its purpose is to precisely position components atop each other. A drawing of the assembly machine is shown in Figure 6.6.

During the assembly procedure, the module or its components reside on the *assembly platform* (see Figure 6.7). Flat components (a baseplate, a MaPSa, or a PS-s) can be held on the platform by a set of vacuum holes. The baseplate can be secured at a predefined xy-position with removable pins that are put through the baseplate inserts. Spacers can be placed in dedicated machined slots in the platform, aligned against a set of stops and held via another set of vacuum holes, supplied by a second vacuum line. In order to align components in α, the platform is mounted on a motorized turntable [162], that provides an angular precision of 175 µrad. In a PS module, geometric centers of both sensors and of the baseplate coincide in xy, and the assembly platform is designed and mounted such that its rotation axis also passes through this point when the module is on the platform.

The machine can pick up, move, and place sensors with a vacuum *pickup tool*, which is installed on the assembly *arm* with a ball joint. The entire arm can be moved in x, y, and z by motorized linear stages [163, 164] with a precision of 4 µm. The pickup tool has a flat plastic surface with a set of vacuum holes, which is leveled to be parallel to the assembly platform.

The arm is also equipped with a *microscopic camera*, the main purpose of which is to determine the positions of sensors in order to precisely place them during the assembly process. Each sensor has four optical *markers* at its corners. The sensor's position can be measured by taking a frame of a marker and applying the template matching algorithm [165], as shown in Figure 6.8.

The entire setup is controlled by a software application [166], written in C++ with Qt libraries. The OpenCV library [167] is used to process camera images (including the template matching procedure mentioned above). The software allows to manually control parts of the setup (e.g. move the positioners or toggle the vacuum lines) and to execute basic routines, such as camera auto-focus or marker position measurements. The software also contains a list of pre-programmed actions, which directly correspond to the steps of the bare module assembly sequence (as given in Section 6.3.3). For the assembly, the operator would trigger them one by one.

Calibrations

Several mechanical parameters and offsets of the assembly machine need to be periodically calibrated to ensure high quality. These parameters are quantified, as described below, and then accounted for in the control software.

1. **Pickup tool calibration**. In order to optimize precision when picking up and placing the components and to ensure a uniform glue distribution over the entire surface when joining

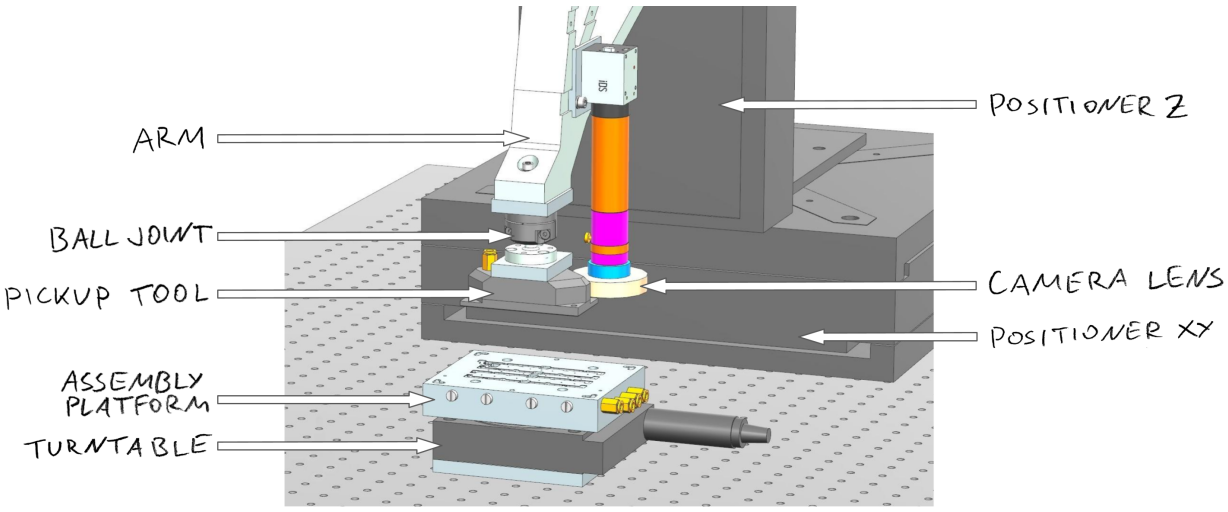


Figure 6.6: Scheme of the assembly robot [161].

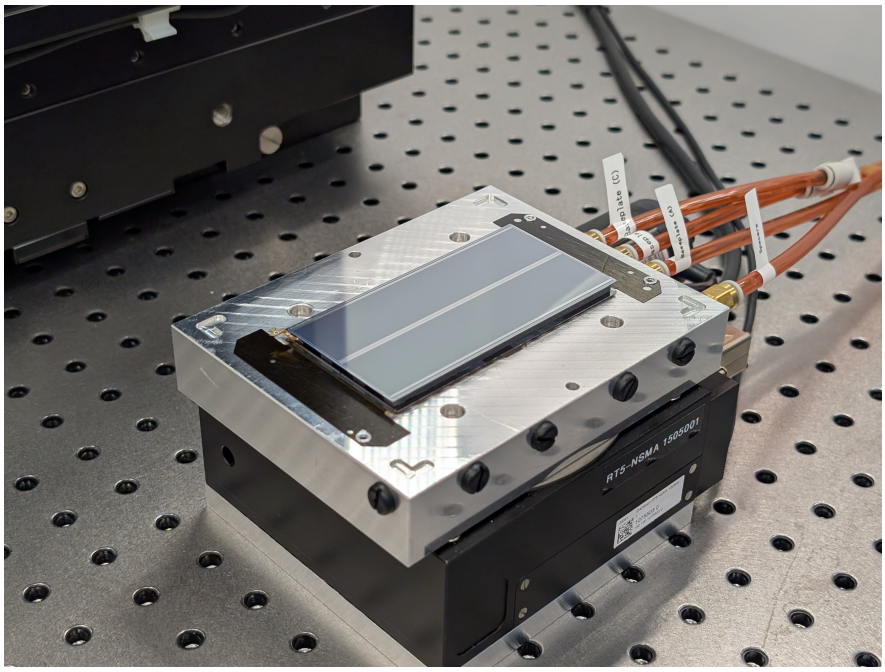


Figure 6.7: Assembly platform with a complete bare module on it.

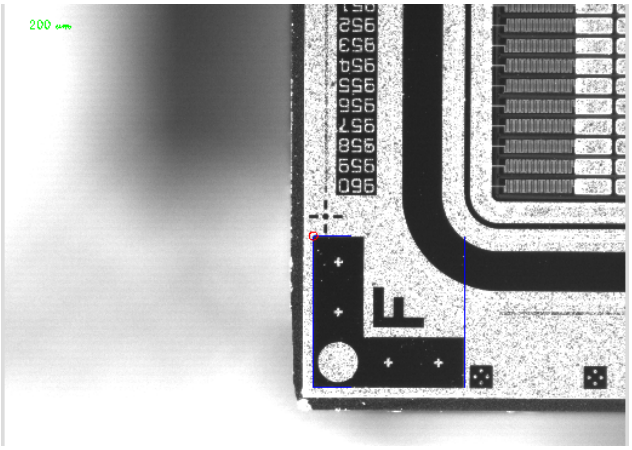


Figure 6.8: Result of the pattern recognition. A frame, taken with a microscopic camera at a corner of a PS-s sensor is shown. The best-match position of the template is shown with a blue square.

components together, the pickup tool needs to be properly leveled. First, the pickup tool is adjusted such that its surface is parallel to the assembly platform. This is done by loosening the ball joint of the pickup tool, lowering the tool onto the assembly platform and tightening the joint back. This allows to ensure that the tilt of the tool is less than 10 µm (over an 8 cm width). Second, the pickup height[1] is determined. Ideally, for the pickup procedure the tool should be lowered onto a component such that it is touching to minimize the shift of the component in xy during the procedure. At the same time, one also should avoid exerting stress on components by pressing the tool against them. The pickup height can be verified with a *multi-pickup test*: a sensor is repeatedly picked up and placed at the same location, and its position is measured on every iteration by the marker recognition procedure. A result of this test is shown in Figure 6.9 for various pickup heights. At the optimal value, the shift in xy is less than 2 µm per iteration. Picking up a sensor higher than the optimum results in an increase of the shift.

2. **Camera angle in the frame of reference.** The camera should be installed in such a way that the frames taken are aligned to x- and y-axes. A rotational offset of the camera w.r.t. the coordinate system can be calculated and compensated for. The scheme for this calibration step is shown in Figure 6.10. A sensor is placed on the assembly platform and two frames of fiducial markers at its corners are taken with the camera. The angle β between the sensor and the x-axis is derived from coordinates of the markers, obtained via pattern recognition: $\beta = \arctan\left((y_2 - y_1)/(x_2 - x_1)\right)$. By pattern recognition, one also obtains the angle α of a marker within a camera frame. The sought camera angle w.r.t. the x-axis is then the difference between the two: $\gamma = \beta - \alpha$.

3. **Coordinates of the assembly platform.** The coordinates of the assembly platform in the machine frame are calibrated by directing and focusing the camera at alignment markers, that are located in the corners of the platform. Additionally, the platform is rotationally

[1] Defined as vertical distance between arm positions "camera is focused on the component's surface" and "pickup tool is lowered onto the component".

aligned with the x-axis. This rotational alignment is done by scanning two corners, opposite in x, and rotating the platform so that the y-coordinates of the two coincide. This is needed to ensure proper placement on the components onto the baseplate.

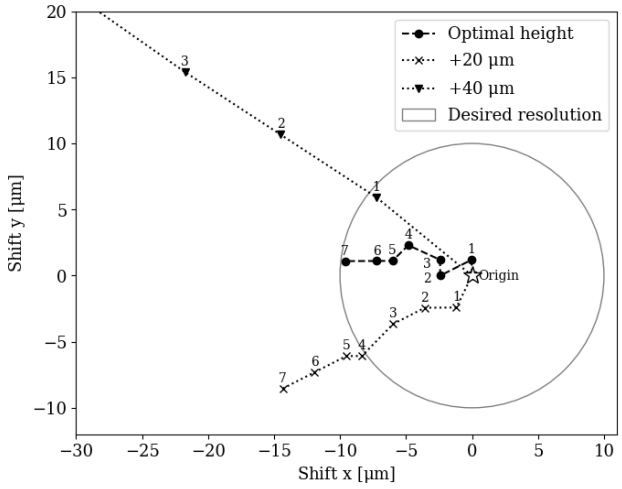


Figure 6.9: Results of several multi-pickup tests. Measured sensor positions in x and y after consecutive pickup-placement iterations are shown. With the optimal pickup setting, 7 iterations are needed for the total shift to exceed 10 µm. Increasing the pickup height significantly increases the systematic shift distance per iteration.

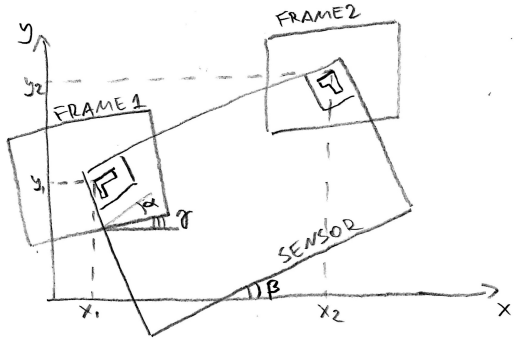


Figure 6.10: Scheme for calibrating the camera angle w.r.t. the reference frame.

6.3.2 Glues and the glue dispensing robot

A two-component epoxy glue [168] is used for the module assembly. This specific type was chosen because it provides the required thermal conductivity, electrical insulation, and it retains its properties after irradiation. For all glue joints in the bare module assembly, a coverage of >95% surface should be achieved. This poses a necessity to precisely dispense glue on the components

and to keep its quantity within a margin of ±10% to avoid a spillage around the glue joint, which might affect the assembly procedure.

Applying the epoxy glue on the components is performed with a volumetric dispensing unit [169], which is capable to reproducibly deposit amounts as small as few µl at a controlled flow rate. The dispensing unit is mounted on a 3-axis CNC robot [170], depicted in Figure 6.11. This facilitates process automation: a component is placed on a special jig on the table of the dispensing machine, and the required amount of glue is dispensed in a pre-programmed pattern (e.g. line segments). The dispensing process and related challenges are discussed further in Section 6.3.4.

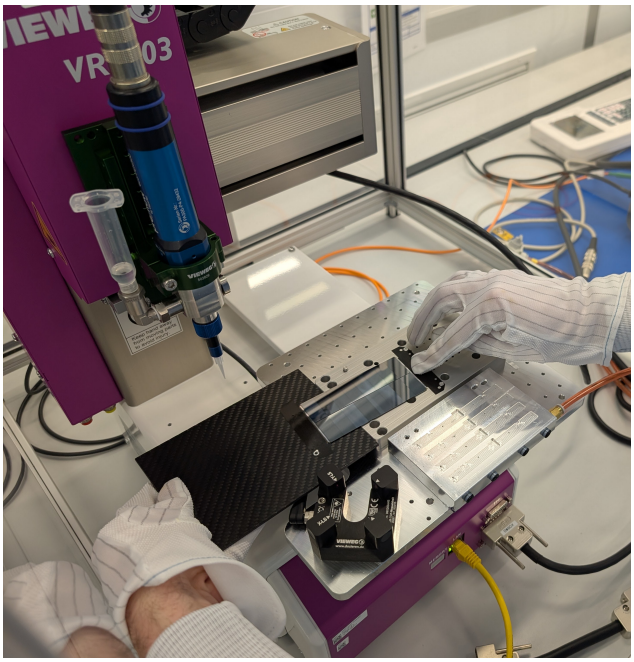


Figure 6.11: Glue application robot. The operator is placing baseplate+MaPSa subassembly onto the jig.

The epoxy requires 24–48 h to fully cure and is therefore referred to as *slow* glue. In order to be able to perform three gluing steps consecutively (as presented in the following section) and to achieve the required throughput of the assembly machine, in every joint the slow glue is complemented with a minor amount of a fast-curing (20 min) epoxy [171], referred to as *fast* glue. The fast glue allows one to handle subassemblies before the slow glue is fully cured. The fast glue is applied onto the components manually as several small drops, which would form connection points between the components. The application pattern is such that the mixing of the fast and the slow glues is minimal when the components are brought in contact. The fast glue is chosen such that it does not react with the slow glue on contact, and the fast glue drops are able to reliably cure even being submerged in the slow glue. Adding fast glue does not significantly alter the properties of the entire glue joint, as fast glue dots cover a minor fraction of its surface (<1% between a MaPSa and a baseplate and <5% on spacers).

6.3.3 Sequence

The bare module assembly is performed in three consecutive main steps: attaching a MaPSa on a baseplate, attaching a PS-s sensor to spacers, and joining these subassemblies together. The full assembly sequence is as follows:

MaPSa to baseplate

1. A MaPSa is placed on the assembly platform and secured with vacuum.
2. The *alignment* procedure is performed: the machine iteratively scans markers at two corners of the PS-p sensor and performs adjustments in α to ensure that the MaPSa is parallel to the x-axis. Its xy-position is also obtained in this process.
3. The *pickup* is performed: the pickup tool moves over the MaPSa and is lowered to it. The vacuum grip is engaged on the pickup tool and disengaged on the assembly platform. The arm with the MaPSa is lifted up. Thanks to the preceding alignment procedure, the MaPSa is positioned always in the same way w.r.t. the pickup tool regardless of its initial placement on the assembly platform. This ensures the reproducibility of the following placement.
4. Slow glue is dispensed onto the baseplate using the dispensing robot.
5. The baseplate is placed on the assembly platform, positioned with pins and secured with vacuum. For the baseplate-to-MaPSa alignment the mechanical positioning with pins yields sufficient precision. A minor amount of fast glue (four drops) is manually deposited on the baseplate.
6. The MaPSa is lowered onto the baseplate. The machine is programmed to leave a gap of 20 µm for a glue layer. The components are held in this position until the fast glue cures; vacuum stays on both on the platform and on the pickup tool.
7. Vacuum is disabled both on the platform and the pickup tool. The baseplate+MaPSa subassembly is removed from the machine.

PS-s to spacers

8. A PS-s sensor is placed on the assembly platform. The alignment and pickup routines are performed similarly to those at the previous stage.
9. Slow glue is dispensed on spacers.
10. The spacers are placed in the slots on the assembly platform and secured with vacuum. Two drops of fast glue are manually applied onto each spacer.
11. The PS-s sensor is lowered onto the spacers, maintaining a gap of 20 µm for glue, and left until the fast glue is cured, similarly to the previous stage.
12. Vacuum is disabled on the platform. The arm is lifted, and the PS-s+spacers subassembly stays suspended on the arm for the upcoming stage.

Baseplate+MaPSa to PS-s+spacers

13. Slow glue is dispensed on the surface of the baseplate+MaPSa subassembly; the pattern matches the shape of the spacers (as demonstrated in Section 6.3.4).
14. The subassembly is placed on the assembly platform, positioned with pins and secured with vacuum.
15. The alignment procedure is run to determine the MaPSa position, as aligning the full subassembly with pins does not ensure the required PS-p to PS-s alignment.
16. Fast glue is manually applied onto the MaPSa surface (two droplets at the location of each spacer).
17. The PS-s+spacers subassembly is lowered onto the baseplate+MaPSa subassembly. High precision of the placement is ensured by the preceding MaPSa alignment and the PS-s alignment, performed in step 8.
18. The complete bare module is left in this position until the fast glue is cured.
19. The bare module is removed from the machine and put in a ventilated cabinet for 48 h for the slow glue joints to cure.

After the bare module assembly, the mechanical specifications of the module (i.e. offsets between the two sensors and their positioning w.r.t. the baseplate) are verified with a dedicated automated optical metrology station [172]. This procedure is a standard part of the quality control and is done on every module. A simplified metrology procedure can also be performed by the assembly machine itself using the markers recognition feature.

6.3.4 Development of the assembly procedures

The module assembly procedure and related instrumentation were developed throughout the several years of production preparation. During this doctoral project, several routines, machinery and software were further improved to reach the required reproducibility. The time, required for the assembly of a single bare module, was reduced to less than 2 hours so the maximum throughput goal can be reached. Also, different variants of mechanical components (e.g. various spacer designs and glue types) were tested.

Most procedures were developed and tested by building mechanical prototypes, also called *glass dummies* (see Figure 6.12). In such bare modules, glass mockups are used instead of the sensors. These glass mockups are approximately of the same dimensions as the sensors, and they have fiducial markers engraved on them, which allows to use the same positioning procedures. The transparent glass panels allow for visual evaluation of the glue coverage for most of the joints.

Automated glue dispensing

In the course of the thesis, a particularly challenging step in the module assembly was developed, addressing application of the slow glue onto the bare module components.

The previously used glue application procedures lead to the following issues :

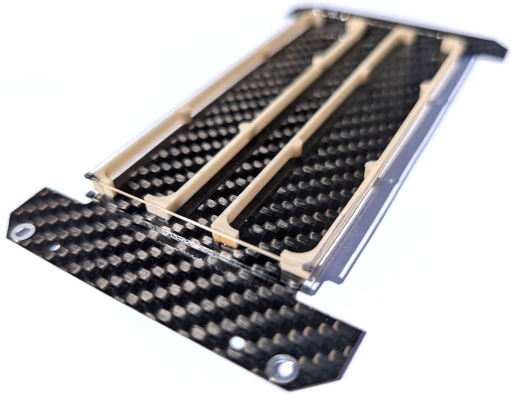


Figure 6.12: Glass dummy module. The baseplate does not have a Kapton insulator on it.

- For the MaPSa to baseplate step, slow glue was applied with a dispensing robot [170], but using a pneumatic dispenser with a metal needle. This resulted in inconsistencies in the total amount of glue dispensed (up to $\pm 20\%$), since the pneumatic dispenser does not allow to precisely control the flow rate.
- For the PS-s to spacers step, slow and fast glue were applied using a manual procedure: two types of glue were smeared on dedicated metal blocks, and spacers were manually brought to contact to this glue surface before being placed on the assembly platform. This procedure required a significant amount of time and manual effort, and was prone to errors.
- For the baseplate+MaPSa to PS-s+spacers step, the same set of blocks was used. The blocks were prepared in a similar way and installed in the assembly machine before joining the two subassemblies together. After that the machine lowered the PS-s+spacers subassembly onto the glue surface and lifted it up immediately. Besides the high amount of manual effort mentioned before, this routine had a disadvantage of low reliability, since it included operations with a semi-cured subassembly, suspended on a vacuum tool.

These issues were overcome in the course of the doctoral project by adopting a volumetric dispensing unit [169]. The controlled dosing, combined with the automation capability of the dispensing robot (introduced in Section 6.3.2), allows to achieve fine glue patterns with sub-mm precision. This requires a precise ($\sim$10 µm) calibration of the dispensing needle height, as due to high surface tension and adhesion of the slow glue the reproducibility of the dispensing is affected of the needle is either too low or too high. This calibration is performed with a dedicated laser tool of the dispensing robot [173].

The resulting glue patterns are shown on Figures 6.13 (for attaching a spacer to a PS-s sensor) and 6.14 (for joining PS-s+spacers to baseplate+MaPSa).

Figure 6.13: One spacer with slow glue dispensed on it. The automated procedure allows to reliably dispense a uniform line on a 1 mm wide spacer. Two gaps in the middle are left to apply the fast glue.

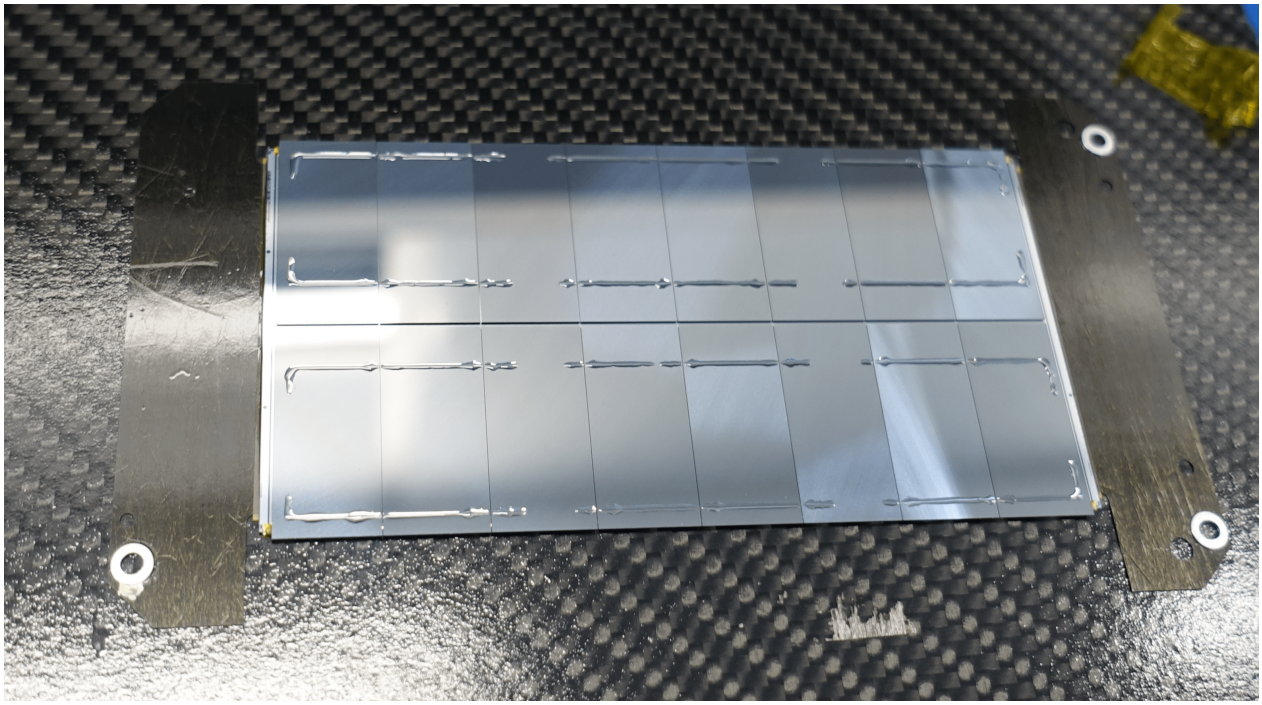


Figure 6.14: Baseplate+MaPSa subassembly with slow glue dispensed on it. The pattern matches positions of spacers. Minor non-uniformity of the lines is caused by variance of the MPA chip heights and by tendency of the slow glue lines to break down due to high surface tension. Gaps in the lines are left to apply the fast glue there.

6.3.5 Results of the bare module assembly

Accuracy of alignment between two sensors for the bare modules, assembled at DESY, is presented in Figures 6.15 and 6.16. Compliance to the sensor specifications is shown: for the translational offset in Figure 6.15 and for the rotational offset in Figure 6.16. Early prototypes and mechanical dummies are not included. The naming convention for PS modules is `PS_DD_R_CCC-NNNNN`, where:

- «`DD`» is sensor spacing: 16, 26 or 40 [mm/10];
- «`R`» is readout speed: 5 or 10 [Gbit/s];
- «`CCC`» is an abbreviation of the assembly center name;
- «`NNNNN`» is a serial number within that assembly center: `0000N` are prototype modules, `0010N` are kick-off modules, and `1NNNN` are production modules.

The procedure demonstrates consistent results and allows to achieve an alignment well within the limits. Exact reasons for the specifications not being met in two cases are well understood and fixed:

- The module `PS_26_5_DSY-00102` has a significant rotational misalignment of the sensors. This module was built with additional temperature probes installed inside the module for evaluation of the thermal characteristics of a kick-off module. The assembly procedure was altered to enable this, lowering the precision.
- The module `PS_26_5_DSY-00104` has an offset in y beyond the tolerance. This was caused by a software bug, affecting the PS-s sensor placement procedure. This bug was introduced before assembly of this module and fixed immediately after it.

Moreover, the optimizations of the procedure, performed during this doctoral project, significantly reduced time required for the bare module assembly to less than 2 h per bare module. This ensures timely production of PS modules at DESY assembly center.

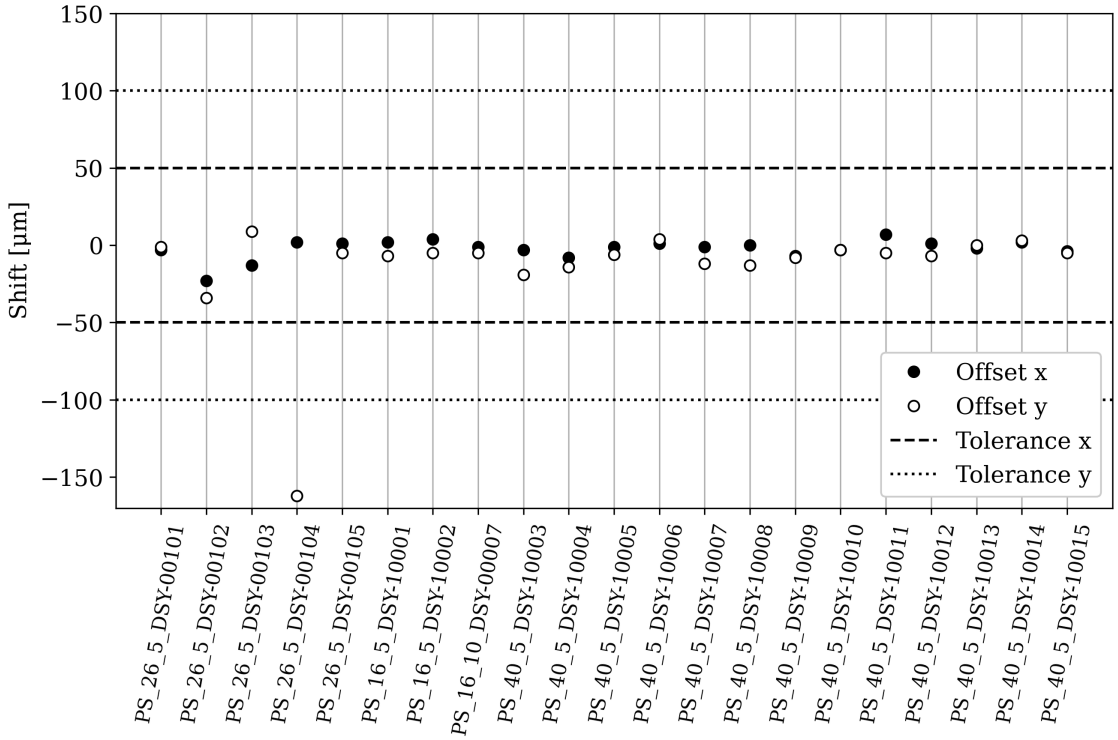


Figure 6.15: Offset between sensors in x and y in early production PS modules.

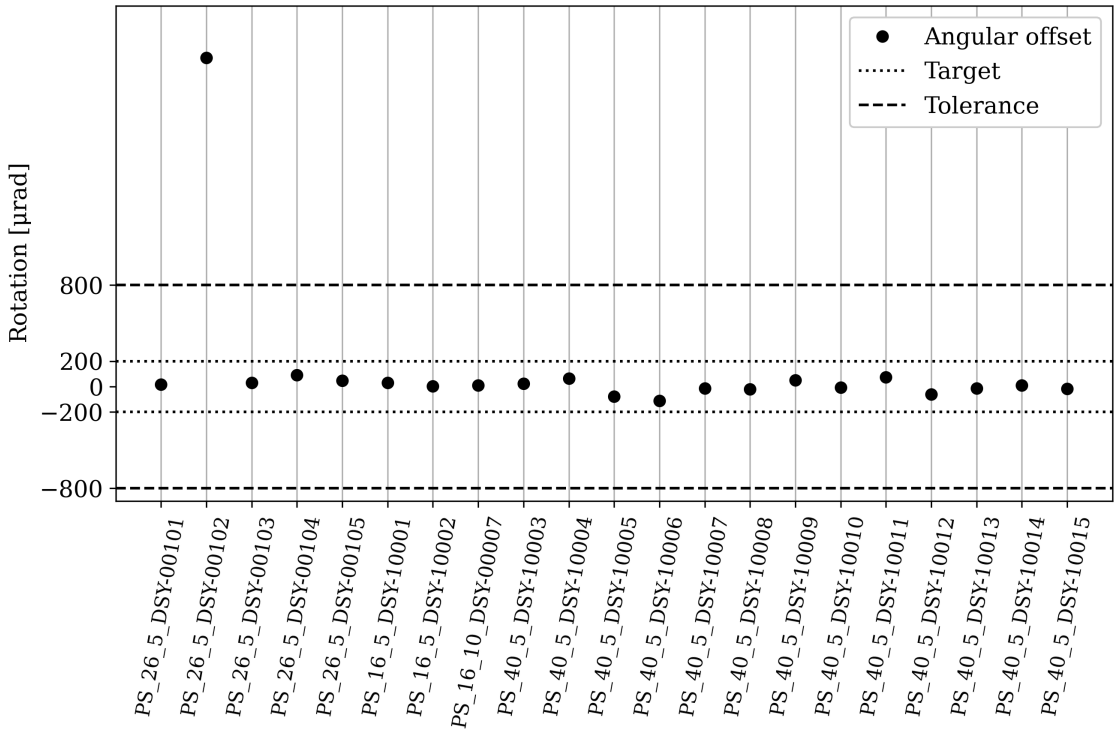


Figure 6.16: Rotational offset between sensors in early production PS modules.

Chapter 7

Test beam qualification of PS modules

This chapter presents the test beam qualification measurements of Outer Tracker modules, carried out in February 2024 at the DESY II test beam facility [174]. The main goal of the campaign was to characterize full modules and evaluate their production readiness, whereas previous studies [175, 156, 157] focused on testing of individual components and early prototypes. Throughout two weeks of data taking six modules were tested:

- `PS_26_5_DSY-00101`[1] (later referred to as `DSY-101`), a DESY-built (as presented in Chapter 6) PS module. An extensive set of tests was performed on this module to fully evaluate its characteristics;
- `PS_16_10_BRN-00009` (later referred to as `BRN-009`), a prototype PS module with a PS-s sensor irradiated to 1.1×10^{15} n_{eq}/cm^2 to measure characteristics of a radiation-damaged sensor, built into an otherwise non-irradiated module;
- `PS_16_10_IPG-00103`, a PS module equipped with a 10 Gbit/s ROH. Tests were focused on evaluation of a firmware update, resovling a previously encountered data transfer issue (not presented in this thesis);
- A set of three 2S modules (outside the scope of this work, presented in Ref. [176].

In Section 7.1 the general concept of test beams and their use for detector testing is explained. Section 7.2 is devoted to the DESY II test beam facility and the instrumentation used. Section 7.3 covers the data reconstruction techniques and the analysis chain, implemented with the Corryvreckan framework [85]. Results of the test beam measurements are presented in Section 7.4.

7.1 Test beams for detector R&D

When developing a new type of particle detectors, one must verify that it matches the expected performance before deploying it in a large-scale experiment. Ideally, such testing should be performed under conditions resembling those in the real experiment as close as possible.

The general concept of evaluating the detector performance is as follows: the device under test (DUT) is exposed to particles, properties of which ideally should be known *a priori*, e.g. from an

[1] Here, the name signifies a module with 2.6 mm sensor spacing and 5 GHz readout speed. The exact naming convention for modules is given in Section 6.3.5.

independent measurement by another trusted detector. By comparing the measurement results of the DUT to the reference values, one may infer the required characteristics of the DUT such as resolution on certain observables.

In the context of ionization tracking detectors the key characteristics are *spatial resolution* and *detection efficiency*. To study those, one requires a source of minimum-ionizing particles and a reliable way to determine particles' trajectories. Such an environment can be implemented in specialized *test beam* facilities [177]. Particle accelerators typically serve as m.i.p. sources, with the dedicated accelerator outputs being called *beamlines*. In order to provide reference measurements, beamlines are often equipped with tracking systems, referred to as *beam telescopes* or *hodoscopes*. These telescopes consist of multiple tracking detectors similar to tracking detectors in particle physics experiments in their form and function. For every particle passing through the DUT the intersection coordinates can be precisely interpolated using a particle trajectory reconstructed by the telescope.

7.2 DESY II test beam facility

7.2.1 DESY II accelerator and beamlines

The test beam facility used in this work [174] is situated at the DESY II synchrotron [178]. The DESY II machine was commissioned in 1985 as an electron and positron pre-accelerator for PETRA [179] and DORIS II [180] e^+e^- storage rings. Nowadays, DESY II accelerates only electrons and is used as an injector for the PETRA III synchrotron light source [181] and also generates beams for a test beam facility.

DESY II operates in cycles, with the current in all its magnets oscillating sinusoidally at 12.5 Hz (period of 80 ms), corresponding to a cyclic change of the beam energy between 450 MeV and 6.3 GeV (see Figure 7.1). A single bunch of approximately 10^{10} electrons is injected from a pre-accelerator complex at the minimum energy of the cycle. After the first half-period the bunch energy reaches its maximum. At this moment, the beam can be extracted to be sent to PETRA III. Otherwise, it continues circulating in DESY II. During the second half-period the bunch is decelerated back to its injection energy. The bunch is stored in the ring for two cycles and is discarded thereafter.

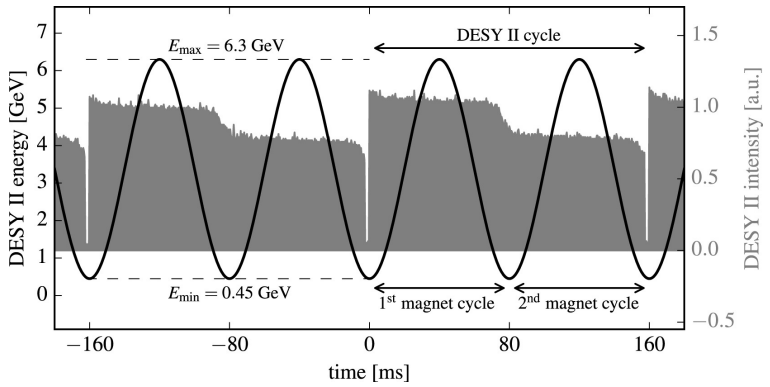


Figure 7.1: Operation cycles of the DESY II accelerator [174]. Two consecutive bunch injection cycles (four magnet cycles) are shown.

The DESY II test beam facility has three independent beamlines, referred to as TB21, TB22, and

TB24. *Secondary* electron or positron beams with an adjustable energy of 1–6 GeV and a particle rate up to few kHz are available to users. These beams are generated via double conversion (see Figure 7.2). On every revolution (period of 976 ns) in the DESY II ring (circumference of 292.8 m) the bunch interacts with a carbon-fiber *primary* target installed directly in the beam pipe, which produces bremsstrahlung with a continuous spectrum ranging up to 6.3 GeV, directed along the beam. Bremsstrahlung photons leave the beam pipe and after traveling a certain distance interact with a *secondary* target in a form of a metal plate. Via pair production the photon beam is converted to electrons and positrons with a continuous energy spectrum. Particles with desired charge and momentum can be selected with a user-controlled dipole magnet and a set of collimators.

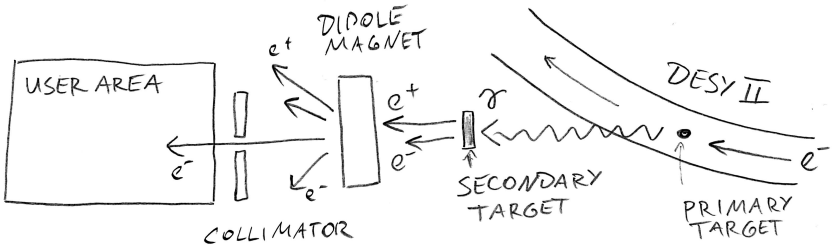


Figure 7.2: Scheme of the secondary beam generation at the DESY II test beam facility.

For the measurements presented in this thesis, an electron beam at the beamline TB21 was used. The beam energy was set to 5 GeV in order to ensure a sufficient particle rate (which is lower for higher beam energies), while reducing the negative effect of multiple scattering on the tracking precision of the telescope.[1] The average secondary beam rate for these settings was approximately 1 kHz. An example measurement of the rate over time is shown in Figure 7.3. A periodic structure corresponding to the cycling of the DESY II accelerator is visible, since secondary electrons with the required energy can only be generated when the bunch energy, which defines the upper limit on the secondary beam energy, exceeds the target value of 5 GeV. The difference of the rate in even and odd periods is related to the decrease of the beam intensity in the second DESY II magnet cycle after the bunch injection (as explained in Ref. [174] and shown in Figure 7.1).

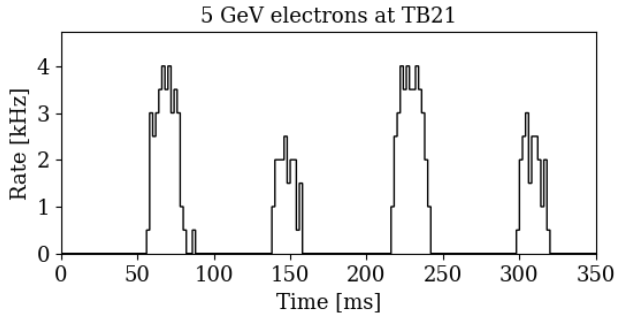


Figure 7.3: Particle rate, measured with the trigger system of beamline TB21.

[1] Similar to impact parameter resolution in tracking detectors, discussed in Section 1.5.3.

A system of collimators is employed to create a nearly parallel beam. Nevertheless, as the beam travels from the source, it is subject to spreading, or *divergence*. The divergence in x and y at the beamline TB21, measured with three upstream telescope planes via straightline tracking (as explained in Sections 7.2.2 and 7.3.1) is presented in Figure 7.4. In either axis, the divergence does not typically exceed 0.1°.

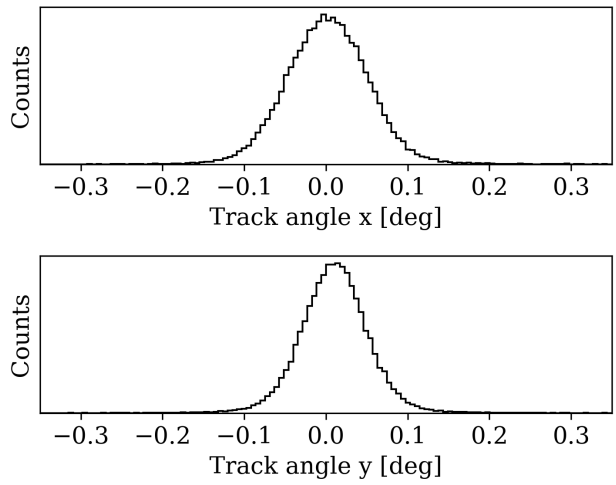


Figure 7.4: Beam divergence, measured with the telescope of beamline 21.

7.2.2 Instrumentation

Telescope

Beamline TB21 at the DESY II test beam facility is equipped with DATURA, a EUDET-type beam telescope (see Figure 7.5) [182]. The telescope consists of six parallel tracking planes (three in the *upstream arm* and three in the *downstream arm*) each being one MIMOSA26 detector [183, 184].

MIMOSA26 is monolithic active pixel sensor produced in AMS 350 nm lithography. Sensors are thinned down to ∼50 µm, so they have a low material budget, which is beneficial for reducing multiple scattering in the telescope and thus improving the tracking resolution. MIMOSA26 has a very fine pitch of 18.4 µm and provides an excellent spatial resolution of better than 4 µm. The entire telescope, in turn, can provide a tracking resolution at the DUT plane (between the telescope arms) of up to 1.5 µm [182].[1] Each sensor has 1152×576 pixels and a total area of approximately 21×11 mm^2. The data is read out continuously in a *rolling shutter* mode (row by row) with the full readout cycle taking 115 µs.

Each sensor is installed in a water-cooled metal enclosure, and the telescope mechanics allows for independent adjustment of the position of each plane along the beam, as well as repositioning the entire arm.

In the data analysis the telescope planes are referred to as MIMOSA26_0 – MIMOSA26_5, with MIMOSA26_0 being the furthest plane upstream (rightmost in Figure 7.5). Cartesian coordinates

[1] The telescope resolution significantly depends on the material budget of the DUT and positions of the telescope planes. Resolution estimates for configurations used in this work are provided in Section 7.3.3.

$\{x, y, z\}$ with the origin at the center of MIMOSA26_0 are used; the z-axis is along the beam; the y-axis points up; the x-axis is chosen such that the coordinate system is right-handed. The PS module is mounted with the strips parallel to the y-axis. The *track incidence angle* (for the rotated module) is measured in the xz-plane between the normal to the module surface and the beam direction.

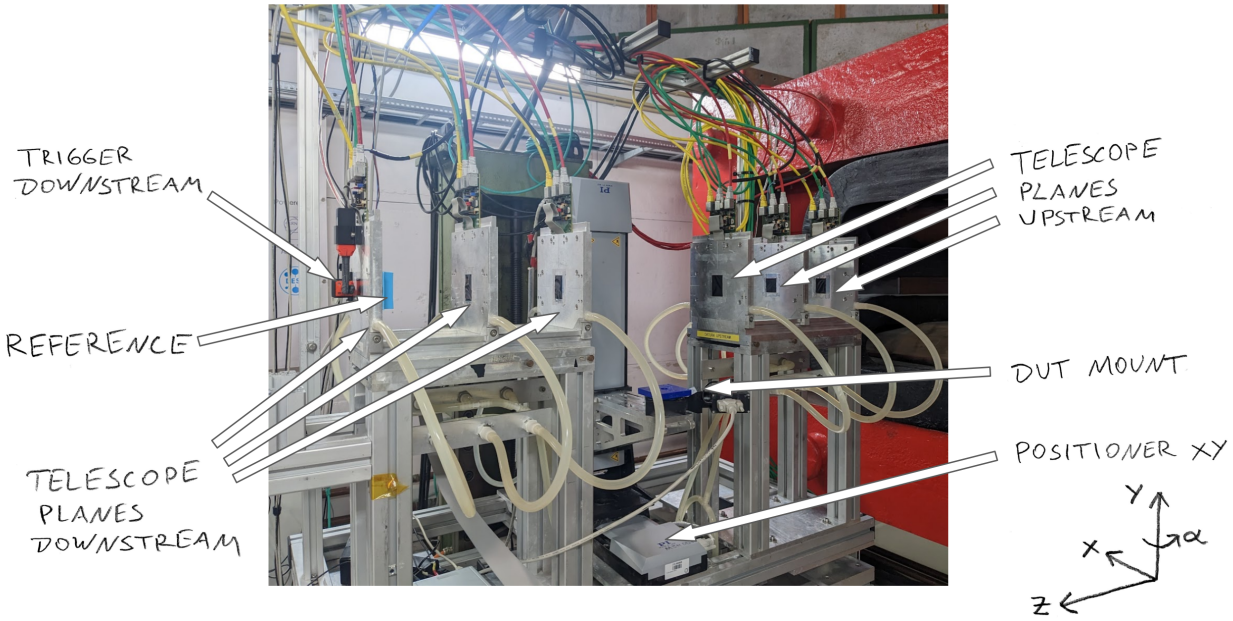


Figure 7.5: View of the DATURA beam telescope in the area 21 of the DESY II test beam facility. A DUT is not installed. The beam is coming from the right.

Triggering and synchronization

The telescope and the DUT rely on an external *trigger* signal to initiate data readout upon a particle passage. These signals are generated with a help of four scintillating tiles coupled to photomultiplier tubes, which provide a fast ($\mathcal{O}$(ns)), self-sufficient, and robust way to generate triggers. Two tiles are installed on MIMOSA26_0, and two behind MIMOSA26_5 (see Figure 7.5, left). The configuration of the tiles yields a geometrical acceptance of approximately $2\times1\,\mathrm{cm}^2$, matching the size and the position of the MIMOSA26 sensors in the telescope.

Generating triggers and distributing them to all detectors in the setup is performed by dedicated hardware, the AIDA Trigger Logic Unit (TLU) [185]. The trigger signal is produced if all four scintillating detectors simultaneously register a particle passage. Besides generating triggers from scintillator outputs, the TLU may also provide other means of synchronization between individual detectors in the test beam setup: distributing clock signals if necessary, declaring a start of the run etc. The TLU also stores information on triggers it produced (e.g. IDs and timestamps), which is used to match the data from different detectors in post-processing.

Reference detector

When a MIMOSA26 sensor receives a trigger signal, two readout frames from it are stored to avoid data loss[1]: the one currently scanned and the subsequent one. Merging two frames leads to an effective frame duration of 230 µs. During that time, the telescope sends a *busy* signal to the TLU, and the TLU would suppress any incoming triggers. With typical particle rates it is not uncommon to have two or more particles arriving within one telescope frame, and all of them would be recorded by the telescope. Therefore, not every telescope track would necessarily correspond to to a trigger, send to the DUT. An additional *reference* detector is used to select the telescope track that actually caused the trigger, if multiple tracks are present in the telescope frame. The reference detector is required to a sufficient time resolution to be able to reliably register beam particles one by one. For each event, the track intersection position measured by the reference detector is used to select the true telescope track during the analysis (see Figure 7.6).

For the measurements in this thesis, a CMS Phase-1 BPIX module (introduced in Section 5.2.5) with a pixel pitch of $150 \times 100\,\mu m^2$ and 25 ns readout frames was used as a reference detector. The module is mounted to the enclosure of MIMOSA26_5 and read out with a Pixel DTB board [186]. In the analysis, the reference plane is designated as REF.

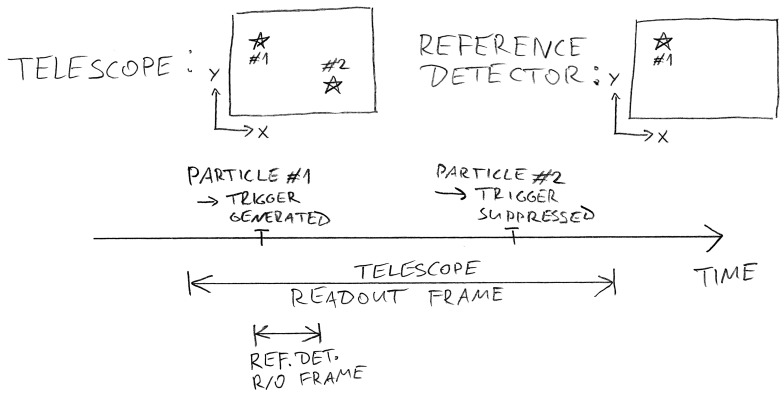


Figure 7.6: Principle of track selection with a reference detector. Here, passages of two tracks within one telescope frame are shown with stars. Track #1 will be kept, and track #2 will be discarded.

Integration of the PS module into the setup

The PS module secured on a metal carrier plate is mounted in a metal enclosure, covering the module from ambient light (see Figure 7.7). The beam passes through cutouts in the box, which are covered with a thin opaque film. The box is mounted on motorized positioners (see Figure 7.5) that move the module in the directions transverse to the beam. Since the acceptance of the telescope is only $2 \times 1\,cm^2$, not the full module area can be studied at once. Typically, the module is positioned such that the beam passes through its center (see Figure 7.8). To mimic curved particle tracks necessary for testing the stub functionality, the module box can be rotated around the y-axis.

[1]Depending on the position of the particle hit w.r.t. the rolling shutter the hit data can be read out in the current or in the next shutter cycle.

The box is flushed with room-temperature dry nitrogen to cool the module down and to prevent condensation. Non-irradiated modules can be operated at room temperature with the nitrogen supply as the only means of cooling. For the irradiated module, which must be operated at low temperatures ($\simeq -30\,^{\circ}\mathrm{C}$), a similar box is used, additionally equipped with thermal insulation and a copper cooling loop. The coolant (silicone oil) at $-40\,^{\circ}\mathrm{C}$ is supplied by an external chiller.

The module is controlled with the Micro Data, Trigger and Control (μDTC) unit [187], which provides communication between the module (via optical link), the TLU, and the data acquisition PC. High voltage for biasing the PS-p and PS-s sensors and low voltage for powering the on-module electronics is provided with external power sources. The connection scheme is shown in Figure 7.9.

Data acquisition and control

The slow control of the setup (e.g. configuring instruments and starting/finishing runs) and data acquisition is performed with the EUDAQ2 framework [188, 189]. EUDAQ2 uses a modular architecture, which provides high flexibility (software modules can be chosen to match the setup) and allows for distributed operation (running modules on different computers in the local network). The following software modules were used:

- `RunControl`: the main process of EUDAQ2, which also provides a graphical user interface;
- multiple `Producer` modules that send commands to and acquire data from all detectors: `NiProducer`, `AidaTluProducer`, `CMSPixelProducer`, and `CMSPhase2Producer` for communication with the telescope, the TLU, the reference detector, and the PS module respectively;
- `StdEventMonitor` and `LogCollector` for live monitoring of the data acquisition;
- `DataCollector` for data collection and storage.

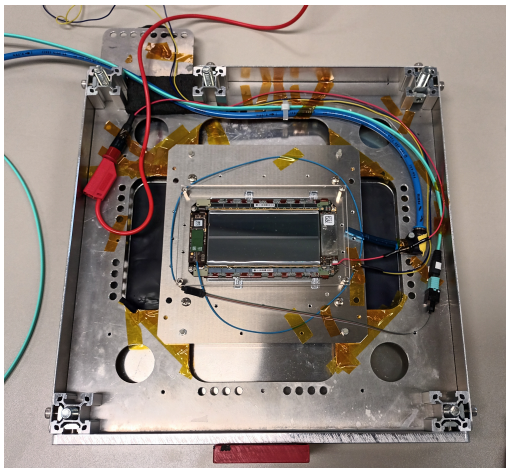


Figure 7.7: PS module in a module box with the upper lid of the box removed.

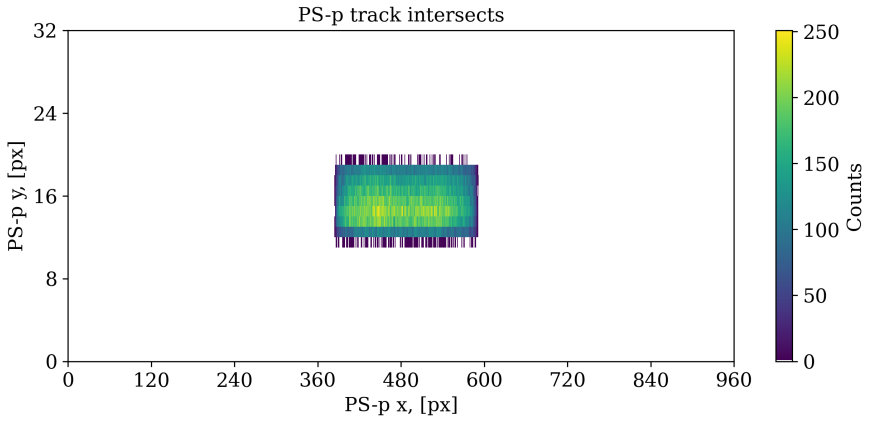


Figure 7.8: Spatial distribution of *telescope* tracks intersecting the PS module. Range of the axes shows the full area of the PS-p sensor.

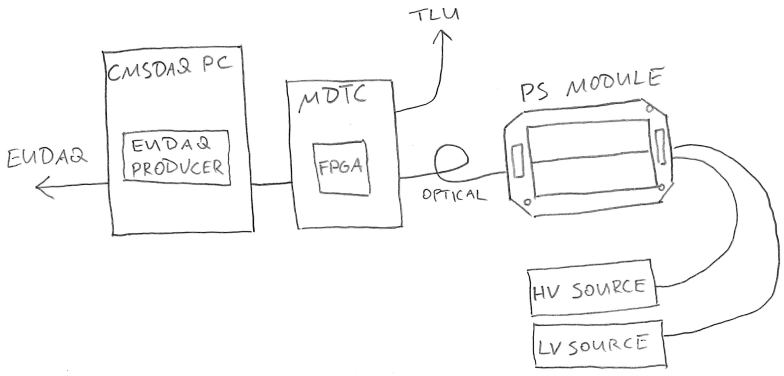


Figure 7.9: Scheme of the PS module connection.

7.3 Data reconstruction with Corryvreckan framework

Data processing and analysis were performed with the Corryvreckan framework [85, 190]. This software is written in C++ and implements various algorithms and procedures commonly used for data reconstruction in test beam studies (as discussed in Section 7.3.1). Features that are unique for studies of the CMS Outer Tracker modules (i.e. analysis of stub performance) were specifically implemented for this [191] and previous qualification campaigns [157, 156].

Besides the data files running the framework also requires a *configuration* file and a *geometry* file, as the philosophy and software architecture of Corryvreckan are significantly inspired by the Allpix Squared framework (introduced in Section 3.1.2). The configuration file contains a list of routines to be performed and their parameters. The geometry file describes all detectors in the setup and their coordinates. The reconstruction algorithm requires knowledge of detectors' characteristics, i.e. pitch and thickness, as well as their precise positions in the laboratory coordinate system. These positions are obtained via the *alignment* procedure, described in Section 7.3.2.

When run, Corryvreckan framework processes data event by event, with each event being defined by one TLU trigger. For every routine performed, outputs and statistics are recorded in a ROOT file [83].

7.3.1 General concepts and procedures

Parameters listed for the routines are selected specifically for this analysis.
Data processing in Corryvreckan typically consists of the several routines, in the following order:

Event loading

Raw data is read from files and `Corryvreckan::Event` instances are created, one per each TLU trigger. Corresponding data from the detectors is assigned to each event. In these measurements, data is already split into events in EUDAQ outputs, which performs online data alignment using the TLU trigger IDs.

Telescope masking

Certain pixels of the MIMOSA26 detectors are *noisy*, i.e. prone to fake hits. Such pixels are identified and *masked*, i.e. hits from them are excluded from further analysis. Here, a frequency cut is used as a selection criterion: a pixel is considered noisy if it has a hit occupancy 25 times higher than average. This cut excludes $\mathcal{O}(100)$ pixels (out of more than 600 thousand) in every MIMOSA26 plane. Masking is performed once and the maps of the masked pixels are stored and used further on.

Clustering

The data acquired from detectors is a list of pixel and strip hits for each event. In every detector, groups of adjacent firing pixels or strips, called *clusters*, are identified. Single hits are considered clusters as well. Most clusters are caused by passage of charged particles, and the coordinate of the passage can then be estimated as the cluster center. For single-pixel clusters the center is the geometrical center of that pixel. For multi-pixel clusters a barycenter (for detectors with binary readout, i.e. the MIMOSA26 and the PS module) or a weighted mean (for the charge-resolving

BPIX reference detector) is taken. `Corryvreckan::Cluster` instances are created and assigned to events. The subsequent procedures primarily use clusters and not the individual hits.

Correlations

In order to check correctness of the data and verify proper synchronization between detectors, *correlation* between cluster positions on different planes within one event is investigated. Since most particle tracks are perpendicular to the sensors, the strong correlation is expected. Examples of correlation between both sensors of the PS module and telescope plane 0, presented in Figure 7.10, signify proper functioning of the setup and synchronization of data streams.

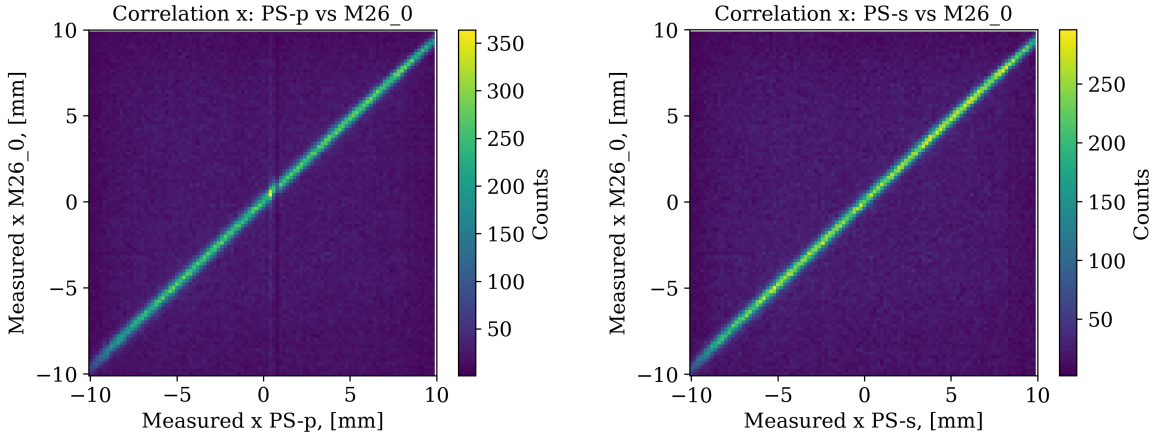


Figure 7.10: Correlation between the particle x-coordinate measurements in PS-p (left) and PS-s (right) sensors and telescope plane 0.

Tracking

Clusters on telescope planes are used to reconstruct particle tracks. Data from the reference detector and from the DUT are not used for tracking. For highly energetic particles, the trajectories can be approximated with straight lines. However, for GeV-range electrons multiple scattering on the telescope planes and the DUT is non-negligible. More precise tracking can be achieved with the General Broken Lines (GBL) method [192], which allows for track kinks at intersections with materials, thus providing improved resolution w.r.t. straight-line approximation. To properly estimate possible scattering angles, this algorithm needs information on positions, materials, and thicknesses of the setup's components. The track parameters are obtained by the algorithm with a χ^2-fit [193], such that the reconstructed track is the best match to the experimental data. The optimization also ensures that the track kink on every plane is in line with the Highland formula estimate (Eq. 1.9). As a result, the `Corryvreckan::Track` objects are created and assigned to events. The tracking procedure is highly dependent on the telescope alignment, discussed further.

DUT and reference association

After tracking, the clusters on the PS module and on the reference detector are analyzed. Clusters on the DUT and REF within a 2-pixel range from the track intersection are assigned to this track as `Corryvreckan::AssociatedCluster` objects. The exact selection criteria are as follows:

$$\begin{aligned} &\text{PS-p:} && |x_{\text{PS-p}} - x_{\text{track}}| \leqslant 200\,\mu\text{m}, \\ &&& |y_{\text{PS-p}} - y_{\text{track}}| \leqslant 3\,\text{mm}; \\ &\text{PS-s:} && |x_{\text{PS-s}} - x_{\text{track}}| \leqslant 200\,\mu\text{m}; \\ &\text{REF:} && |x_{\text{REF}} - x_{\text{track}}| \leqslant 300\,\mu\text{m}, \\ &&& |y_{\text{REF}} - y_{\text{track}}| \leqslant 200\,\mu\text{m}. \end{aligned} \tag{7.1}$$

To properly perform this procedure, the coordinates of the DUT and REF in the laboratory frame should be known to a sub-pixel level to match clusters to tracks. This is achieved by the DUT alignment routine, discussed below.

Associated clusters on the PS-p and PS-s planes are used to evaluate the performance of the module by comparing the cluster parameters (i.e. measurements by the module) to the track parameters (i.e. the reference information). Association of the REF clusters serves as a way to filter out telescope tracks outside of the DUT acquisition frame (see Figure 7.6). Tracks without an associated cluster on the REF are discarded.

7.3.2 Setup alignment

Coordinates and orientations of the setup components should be known with a sub-µm precision, which is crucial for performing tracking and associating DUT hits to tracks. As a first approximation, mechanical measurements are taken to provide a starting point for offline *data-based alignment*, performed in two steps: *prealignment* and *track-based alignment*, discussed in the following. First, the telescope itself is aligned so that tracks can be fitted. Second, the DUT is aligned using the reconstructed telescope tracks, so that its position w.r.t. the telescope is precisely known.

Telescope prealignment

The prealignment routine performs a coarse alignment of the telescope planes without reconstructing tracks. Instead, it relies on an assumption that all particle trajectories are almost parallel to the z-axis (see Figure 7.4). This algorithm analyzes correlations between cluster coordinates of telescope plane 0 and all the other telescope planes. Offsets in x and y are then applied to telescope planes 1–5, such that distributions of the values $(x_{\text{M26\_i}} - x_{\text{M26\_0}})$ and $(y_{\text{M26\_i}} - y_{\text{M26\_0}})$ are centered at 0. As a result, a new geometry file with updated coordinates is generated. The prealignment procedure is able to position telescope planes in x and y with a µm-level precision, which is sufficient to perform tracking and track-based alignment.

Telescope track-based alignment

A set of $\mathcal{O}(10^4)$ tracks is then reconstructed to perform the second alignment step. A built-in optimizer [194] is used to adjust the coordinates x and y and extrinsic rotational angles $(\alpha_x, \alpha_y, \alpha_z)$

of telescope planes 1–5. The total χ^2 of the entire set of tracks is used as the optimization metric. The z-positions are not adjusted since they are *weak modes* for this configuration, i.e. they do not put constraints on tracking and therefore cannot be optimized via track-based alignment.

The impact of the alignment on the track fitting performance is presented in Figure 7.11. The χ^2/DoF (χ^2 per number of degrees of freedom) values of approximately 1 signify a proper telescope alignment.

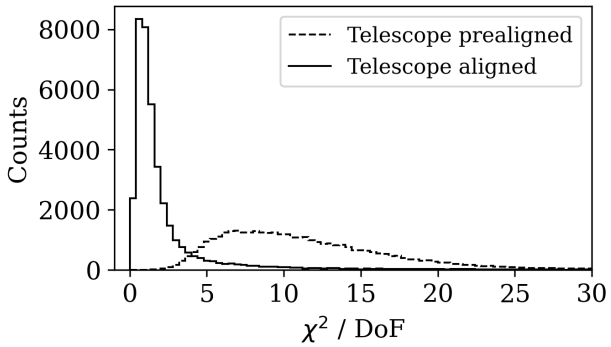


Figure 7.11: Distribution of the χ^2/DoF values of the same set of tracks, reconstructed before and after the track-based alignment of the telescope.

DUT and REF alignment

Similar prealignment and track-based alignment routines are then performed for the PS module and the reference detector. The prealignment procedure is fully identical to the one used for the telescope. The track-based alignment uses the same approach, but a different optimization metric. Here, the optimizer minimizes the variance of track residual distributions, i.e. the difference between track intersection coordinates (interpolated from the telescope data) and measurements by the DUT itself. Additional weak modes are present for the PS-s sensor: the y-coordinate (along the strips) and angle α_x (between the strips and the z-axis). These values are copied from the PS-p alignment, as the PS-s sensor is rigidly attached to the PS-p.

7.3.3 Telescope resolution

Assuming that systematic offsets are excluded by the alignment, one may estimate the GBL track resolution at the DUT plane, expressing how precisely the telescope can measure the position of a passage traversing the DUT. This calculation takes into account multiple factors, such as the intrinsic resolution of the telescope planes, the beam energy, the material budgets of all the detector planes, and the z-coordinates of the telescope planes. Regarding the latter, moving the telescope planes closer to the DUT is typically beneficial, as it reduces the negative impact of multiple scattering in the DUT on tracking precision.

Given the relatively large material budget of the PS module ($\sim$1 mm of silicon), moving the planes as close as possible to the DUT is preferred, which is referred to as the *narrow* geometry. For tests of the irradiated module the material budget is even higher, as the module is mounted on a solid 3 mm aluminum plate to improve cooling. For the measurements including rotation of

the module, however, the arms should be moved further away from the DUT to provide enough clearance (*wide* geometry).

Provided all the setup parameters, the telescope resolution can be calculated with the GBL resolution simulator tool [195]. Resolution estimates for the used configurations are presented in Table 7.1.

Table 7.1: Telescope tracking resolution at the DUT plane.

Setup geometry	Δz telescope–DUT	Al carrier	Resolution @ DUT
Narrow	50 mm	No	4.1 µm
Wide	160 mm	No	8.6 µm
Narrow	50 mm	Yes	4.4 µm
Wide	160 mm	Yes	9.7 µm

7.3.4 Observables definitions

Efficiency

The detection efficiency of the DUT can be measured in the test beam environment as a probability to produce an associated cluster in a response to a passing particle:

$$\text{Efficiency} = \frac{\text{Number of selected tracks with an associated DUT cluster}}{\text{Total number of selected tracks}}. \tag{7.2}$$

Uncertainty in the efficiency measurement can be estimated via Clopper-Pearson intervals [196] at 68.3% confidence (1σ). Efficiency can be also defined for stubs (discussed in Section 5.3.1), since stubs contain coordinate information and can be associated to tracks similarly to clusters. Tracks should be filtered in a certain way to exclude effects, related to tracking in triggering in the test beam setup. To be included in the efficiency calculation, a track should match all the following criteria:

- presence of an associated REF cluster, which ensures that the track corresponds to the particle that triggered readout of the DUT;
- goodness of the track fit ($\chi^2/\text{DoF} \leqslant 3$);
- arrival of the trigger at the proper phase of the PS module clock, as detailed in the following.

Since PS modules are designed to operate in the LHC environment with particles arriving every 25 ns, sampling of the signal in the sensors is performed synchronously with the LHC clock, once per cycle. At the test beam, however, particle hits are asynchronous. This effectively leads to a uniformly distributed random delay between the passage of the particle and sampling of the pulse in the front-end chips of the PS module. This results in a deterioration of the *measured* detection efficiency, e.g. if a pulse is sampled at its falling edge. To overcome this issue, the μDTC provides a dedicated time-to-digital converter which measures the trigger phase w.r.t. the module clock cycle (dividing it into 8 bins) and stores this information alongside with other data. Therefore, events can be filtered by the delay between the particle passage and the sampling.

Sensor efficiencies for tracks arriving at different phases of the clock are shown in Figure 7.12. In the presented example, for both sensors the measured efficiency is the highest if the delay is in

the bin 5, indicating that the pulse is being sampled near its maximum. For a realistic estimate of efficiency of a module operating synchronously to the LHC clock, only events in the most efficient bin are selected. For measurements, requiring higher statistics but not a precise efficiency estimate, i.e. spatial resolution, events with phase in other bins can be used as well (e.g. in the second and third most efficient), as the pulse is still sampled at a sufficiently high amplitude.

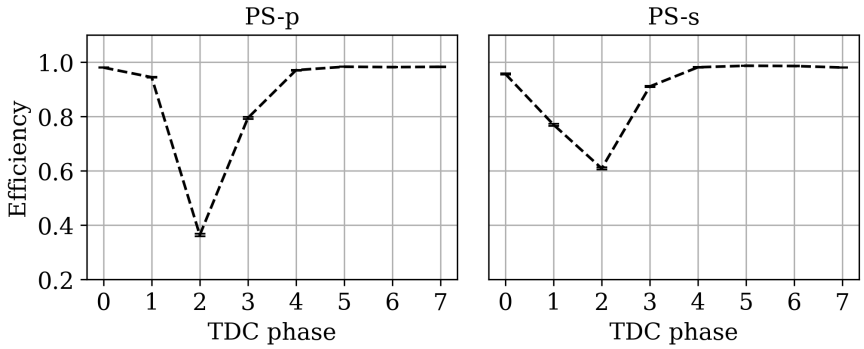


Figure 7.12: Efficiencies of the PS-p (left) and PS-s (right) sensors for triggers at different phases of the module clock cycle.

Spatial resolution

The spatial resolution of the detector is evaluated by analyzing the distribution of track *residuals*, i.e. differences between track coordinates measured by the DUT and interpolated from the telescope track:

$$\text{residual}_x \equiv x_{\text{track}} - x_{\text{DUT}} \quad \text{(similarly for } y\text{)}. \tag{7.3}$$

The width of the residual distribution is affected by coordinate measurement uncertainties in both telescope and DUT. Assuming that these uncertainties are uncorrelated, one obtains the intrinsic DUT spatial resolution:

$$\sigma_{\text{DUT}} = \sqrt{\sigma^2_{\text{residual}} - \sigma^2_{\text{telescope}}}, \tag{7.4}$$

where σ_{residual} is the width of the obtained residual distribution and $\sigma_{\text{telescope}}$ is the telescope resolution (see Table 7.1).

7.3.5 Trigger offset

An issue with synchronization between the telescope and the DUT/REF was encountered during the test beam campaign. During data taking the trigger offset between the telescope and the rest of the setup would randomly change, resulting in data de-synchronization between the telescope and the other detectors. A possible explanation for this effect is the electromagnetic interference caused by switching modes of the telescope cooling system, which temporarily disrupts triggering of the telescope [197]. Such a disruption would result in wrong distribution of the data into events in EUDAQ output files, which is later used to define Corryvreckan events. In total, approximately 20% of runs are affected.

An example of an affected run is presented in Figure 7.13. If the data is synchronized in EUDAQ events, tracks are being properly associated to DUT clusters ($t < 65\,\text{s}$). At $t \simeq 65\,\text{s}$ and $t \simeq 75\,\text{s}$ trigger disruptions happen, therefore shifts of the DUT and REF data need to be applied to the subsequent events: by 1 event for $65\,\text{s} < t < 75\,\text{s}$ and by 2 events for $t > 75\,\text{s}$. Such correction procedure is performed manually by identifying the disruptions, splitting the affected run into blocks and adjusting the trigger shift in each. This was done for runs with a high fraction of desynchronized events. It is often sufficient to only keep the longest continuous unaffected block and discard the rest. Given a typical run size, approximately 10000 tracks are selected for the efficiency calculation after applying all cuts. Assuming sensor efficiency of 0.99, this yields Clopper-Pearson errors of $\simeq$0.10%. If 50% of events are discarded, this error is increased to $\simeq$0.14%.

Presence of a minor fraction of off-sync events in a run does not significantly affect the observables, mentioned in Section 7.3.4, because off-sync events are automatically excluded by the DUT and REF association. The efficiency calculation is not affected either, as the desynchronized parts of the run do not have associated REF clusters.

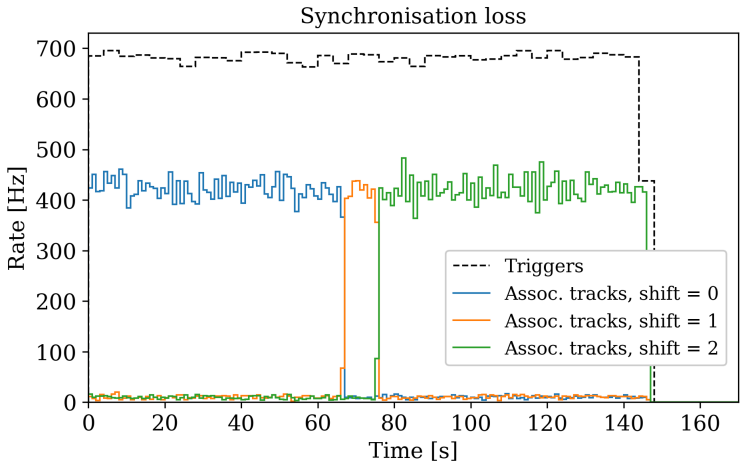


Figure 7.13: Example of a run affected by the synchronization disruption. Trigger rate over time is shown with a dashed line. Rates of PS-p track associations with various trigger offsets applied are shown with colored lines.

7.4 Results

7.4.1 Detection efficiency studies

Detection efficiency is a crucial aspect of PS modules performance, which should be kept as high as possible. Therefore, the efficiency of both sensors needs to be studied as a function of main parameters affecting pulse formation and signal detection in sensors, namely, bias voltage and discrimination thresholds. Such series of efficiency measurements was performed for the `DSY-101` module. The results were used to find the *working point* for the module, i.e. operational conditions that yield optimal detection efficiency. For this measurement series the module is mounted perpendicularly to the beam. For each data point, 10^5 events are accumulated, which results in $\mathcal{O}(10^4)$ tracks being selected for efficiency calculation per data point.

Threshold scan

Discrimination thresholds for PS-p and PS-s sensors are adjusted by 8-bit digital-to-analog converters (DACs) in MPA and SSA chips respectively. The full DAC range of 256 units allows to set thresholds equivalent to amplitudes of pulses up to 17.5 ke^- in the MPA and up to 60 ke^- in SSA. The conversion is as follows:

$$\begin{aligned} \text{Threshold}_{\text{MPA}}[e^-] &= (\text{Threshold}_{\text{MPA}}[\text{DAC units}] - 70) \times 94, \\ \text{Threshold}_{\text{SSA}}[e^-] &= \text{Threshold}_{\text{SSA}}[\text{DAC units}] \times 250. \end{aligned} \tag{7.5}$$

To optimize the detection efficiency, thresholds are to be set as low as possible. However, lowering the threshold is limited by the increasing the rate of noise events. The *noise occupancy* (i.e. the probability to register a noise hit per channel) should not exceed 10^{-6} for pixels and 10^{-5} for strips. The noise occupancy as a function of threshold can be estimated by setting the TLU to send *emulated* triggers to the module with no beam present.

Threshold scans should be performed at a bias voltage that exceeds the full depletion voltage. Otherwise, less charge would be collected since the active thickness of sensors is reduced due to under-depletion. Here, the bias voltage is set to 350 V which exceeds the full depletion voltage in both sensors[1] of 280 V.

The obtained dependence of sensors' efficiencies on the threshold, as well as their noise occupancies, are presented in Figure 7.14 for PS-p and in Figure 7.15 for PS-s.

For the PS-p sensor, the efficiency reaches a plateau of above 98% for the threshold in the range of 90–150 units. The PS-p efficiency, however, does not exceed ∼99% due to presence of biasing structures in every pixel, which renders a fraction of the pixel area intrinsically inefficient for perpendicular tracks (as discussed in Section 7.4.3). At the thresholds below 90 units the efficiency drops due to higher noise occupancy, which causes saturation of data buffers. Increasing the threshold above 150 units results in a decrease in efficiency, as not all pulses can cross it anymore.

For the PS-s sensor the behavior is similar with a plateau of >99.5% for the thresholds below 50 units and a gradual decrease thereafter. The available threshold adjustment range in SSA chips for PS-s is significantly higher than in MPA chips for PS-p, so the PS-s threshold can be set so high so that no pulses can cross it and the sensor efficiency drops to zero.

Threshold values of 110 units in MPA and 30 units in SSA are chosen for the working point.

[1]The full depletion voltage is obtained from a capacitance-voltage characteristic of sensors; the technique is explained in Section 3.2.1.

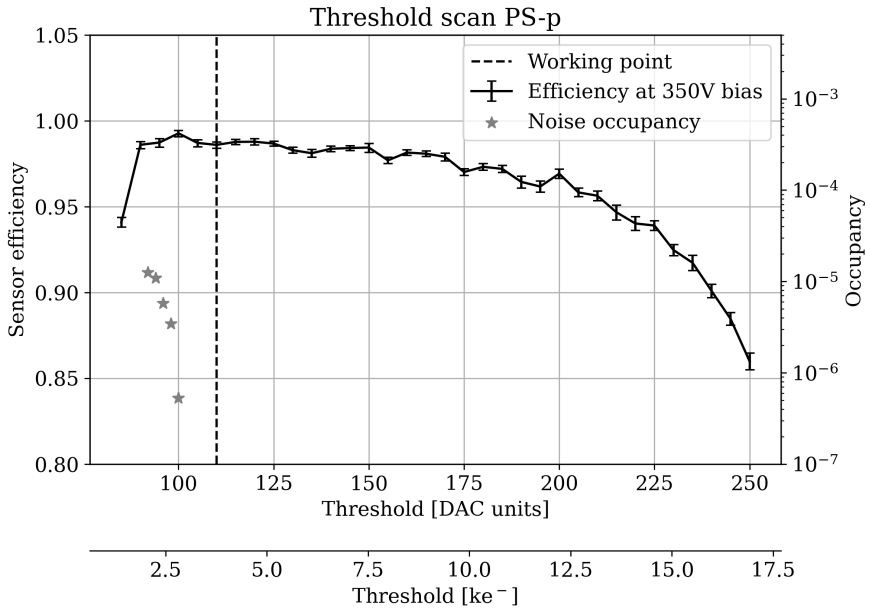


Figure 7.14: Efficiency (solid line) and noise occupancy (star markers) of the PS-p sensor as a function of threshold. The dashed line indicates the working point.

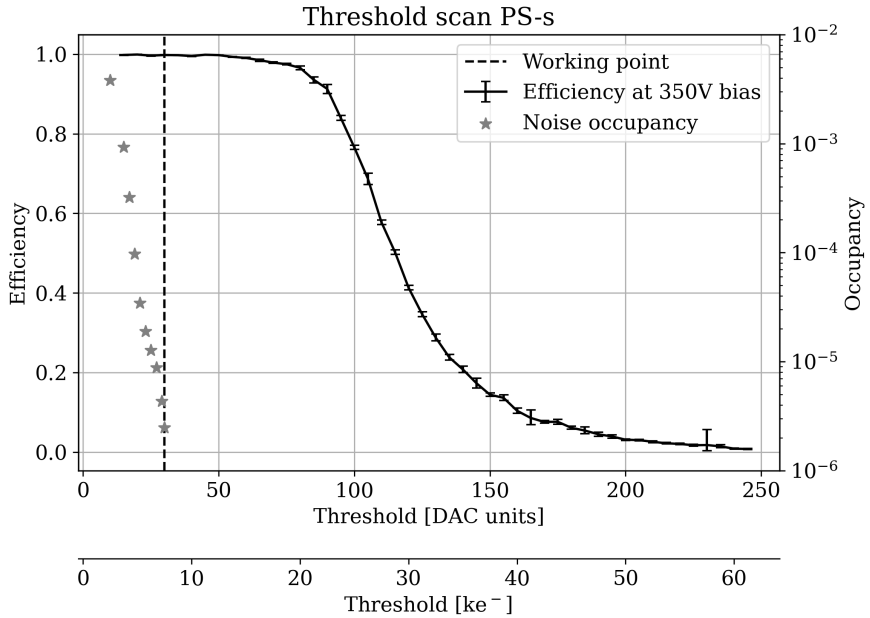


Figure 7.15: Efficiency (solid line) and noise occupancy (star markers) of the PS-s sensor as a function of threshold. The dashed line indicates the working point.

Bias voltage scan

A bias voltage scan is performed up to 350 V, since a further increase of the bias above the full depletion voltage would not affect the detection efficiency. For this measurement, thresholds on PS-p and PS-s sensors are set to previously obtained optimal values of 110 and 30 units respectively. The voltage supply scheme is designed such that the same voltage is applied to both sensors. The results of the scan are shown in Figure 7.16. The depleted zone thickness is the main factor that determines the dependence of efficiency on bias for non-irradiated sensors. At voltages lower than the full depletion voltage the active thickness is reduced, leading to lower amount of collected charge and thus lower efficiency. For the selected thresholds, however, even under-depleted sensors stay fully efficient within a certain margin.

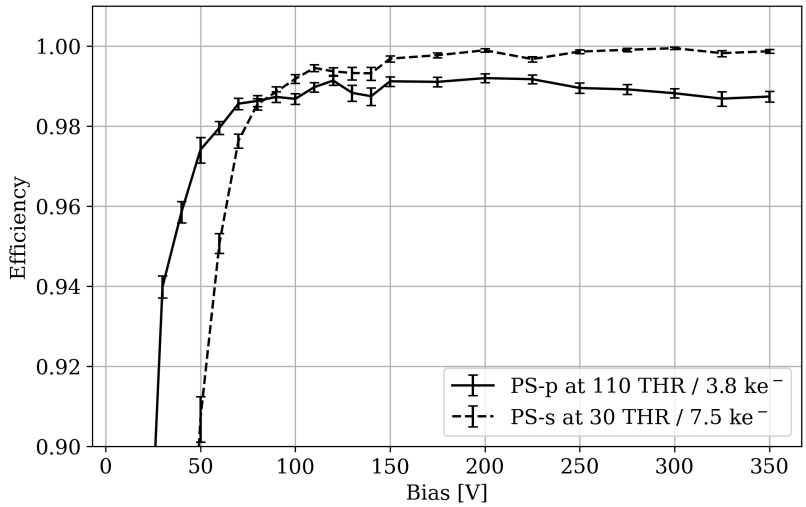


Figure 7.16: Efficiencies of PS-p (solid line) and PS-s (dashed line) sensors as a function of the bias voltage.

Working point and efficiency

The working point efficiencies of both sensors in the `DSY-101` module are given in Table 7.2. The PS-s sensor reaches nearly 100% efficiency under optimal working conditions.

Table 7.2: Working point and sensors efficiencies of the non-irradiated PS module.

Sensor	Bias [V]	Threshold [DAC]	Threshold [ke^-]	Efficiency
PS-p	350	110	3.8	98.7±0.1%
PS-s	350	30	7.5	99.8±0.1%

Incidence angle and full depletion

This section showcases the impact of bias voltage (and, in turn, of the active thickness) and threshold on the detection efficiency of particles arriving at an angle. The PS-s sensor is used as an example. Efficiency measurements at various bias voltages, thresholds, and particle incidence angles (achieved by rotation of the module) are presented in Figure 7.17.

The performance of segmented silicon sensors for tracks crossing at an angle is determined by the geometrical configuration of the sensor, namely, by the ratio between its pitch and its thickness.

- In *thick* detectors (pitch $\ll$ active thickness), e.g. in a fully depleted PS-s sensor, the efficiency typically decreases for non-perpendicular tracks. Even though the total induced charge is higher, it is split between multiple channels (as discussed in Section 7.4.4), thus the amount of charge per channel is reduced. If the detection threshold is low enough, this has a minor effect, but for higher thresholds it may result in a significant efficiency decrease even for a fully depleted sensor.
- In *thin* detectors (pitch $\gg$ active thickness) increasing the incidence angle similarly increases the total charge, but it is still collected in one channel, therefore the efficiency is higher. This behavior can be partially demonstrated by running a PS-s sensor at very low bias voltages, such that the depletion layer is thin (see curves with triangular markers in Figure 7.17 at $V < 50\,\mathrm{V}$).

The bias scan with a threshold at 60 units and 45° incidence angle demonstrates the full depletion of the sensor: increasing the bias above the full depletion voltage of 280 V does not increase the active thickness and thus detection efficiency anymore.

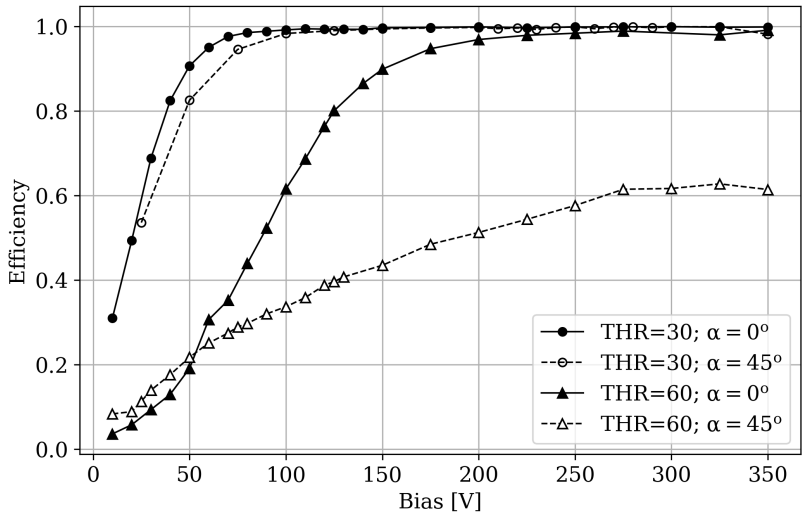


Figure 7.17: Efficiency of the PS-s sensor as a function of bias voltage at various thresholds and particle incidence angles. Threshold value is shown with marker shape: round for 30 units and triangular for 60 units. Particle incidence angle is presented with lines: solid ones for perpendicular and dashed for 45°. Error bars are not shown.

7.4.2 Efficiency of an irradiated sensor

The efficiency study was also performed for the irradiated PS-s sensor of the BRN-009 module. The sensor was irradiated to the equivalent non-ionizing dose of 1.1×10^{15} n_{eq}/cm^2. This reflects the maximum dose for PS modules during the HL-LHC operation, which is estimated to be in the range of $5 \times 10^{14} - 1 \times 10^{15}$ n_{eq}/cm^2 depending on the position within the Outer Tracker [50]. On this module the front-end was not irradiated, which allows one to study the effects of the radiation damage of the sensor separately from the front-end, since front-end irradiation can have an impact on the module operation as well.

The efficiency measurement procedure is similar to the one described in Section 7.4.1. The voltage scan is performed up to 600 V, which is the nominal operation voltage foreseen for irradiated sensors in the CMS detector. Threshold and bias scans for the irradiated PS-s sensor are presented in Figures 7.18 and 7.19 respectively. The following effects of radiation damage, explained in Section 1.5.4, can be observed:

- Increase of the effective bulk doping concentration[1], which results in higher full depletion voltage. This is reflected in a necessity for higher bias voltage to reach the full efficiency;
- Charge trapping, which decreases charge collection efficiency and therefore reduces amplitudes of the induced pulses. This is reflected in a drop in efficiency when increasing the threshold, as well as in generally lower efficiency, compared to non-irradiated sensors. Charge trapping also influences the efficiency-vs-bias dependency. Higher bias voltage results in faster charge collection, therefore, less charge carriers are trapped/recombined.

The tested irradiated PS-s sensor retains the efficiency of 99.5% under optimal conditions (see Table 7.3), although margins around the working point are reduced, compared to the non-irradiated sensor. This indicates that PS-s sensors would sustain the radiation environment of the HL-LHC and remain fully operational.

Table 7.3: Working point and efficiency of the irradiated PS-s sensor.

Sensor	Bias [V]	Threshold [DAC]	Threshold [ke^-]	Efficiency
PS-s irradiated	600	23	5.8	99.5±0.1%

[1] PS-s sensors have a p-type bulk, and radiation-induced defects are also p-like.

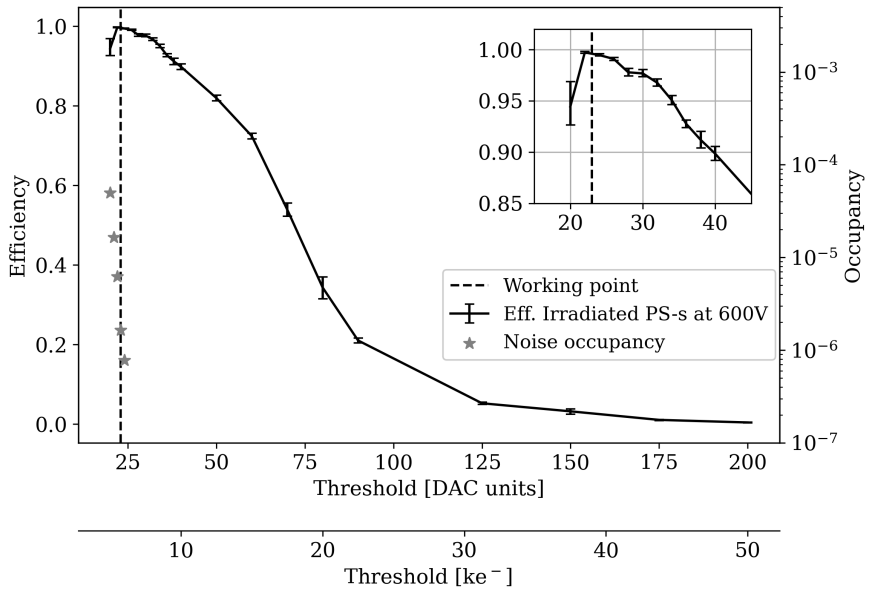


Figure 7.18: Efficiency (solid line) and noise occupancy (star markers) of the irradiated PS-s sensor biased at 600 V as a function of the threshold. The dashed line indicates the working point. The inset axes show zoom to the region around the working point.

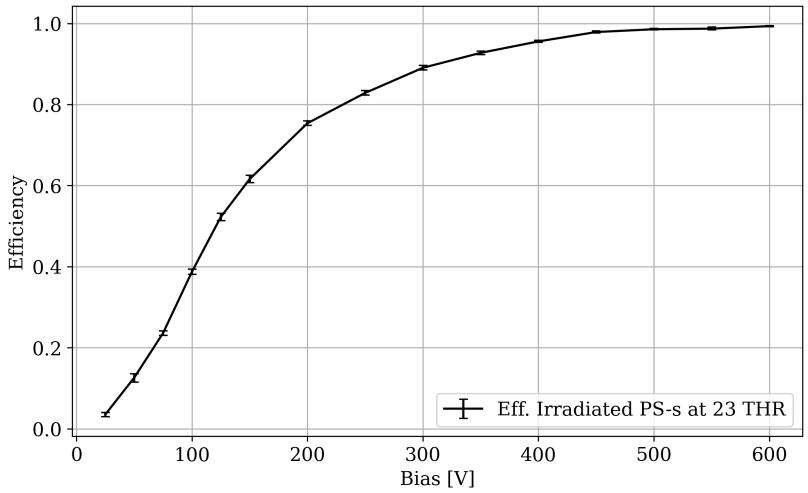


Figure 7.19: Efficiency of the irradiated PS-s sensor at the threshold of 23 units as a function of the bias voltage.

7.4.3 In-pixel studies

Micron-level resolution of the telescope and the precise alignment of the DUT enable studies of various observables, such as efficiency, as a function of in-pixel position. For this purpose, the pixel area of the DUT can be divided into bins, and observables can be calculated independently for each bin by selecting only tracks that pass through a particular bin. Data from all pixels is then analyzed together to form an in-pixel map. These measurements would have a position uncertainty (smearing) of the order of telescope resolution $\sigma_{\text{telescope}}$.

The obtained in-pixel (in-strip) efficiency maps of the PS-p and PS-s sensors are presented in Figure 7.20. The PS-p sensor was operated at the working point. The sensor shows uniformly high efficiency across the entire area except for regions where the biasing structures are located [198, 199]. Bias rails at the short edges of macro-pixels and bias dots at ($x = 0\,\mu\text{m}, y = \pm 700\,\mu\text{m}$) create intrinsically inefficient spots in a pixel. The placement of the bias dots is alternating in even and odd macro-pixel columns (see the layout of the macro-pixel matrix in Figure 7.21). The biasing structures are kept at the same electric potential as the readout implants, therefore they do not significantly distort the electric field and still may collect the charge carriers. However, they are not electrically connected to the readout implants, so if charge carriers drift directly towards the biasing structures, no signals would be registered by the front-end. This inefficiency appears almost exclusively for tracks at an exactly perpendicular incidence. Otherwise, for particles impinging at an angle,[1] the deposited charge would be spread in xy across a certain area of the pixel, which excludes the scenario when a major fraction of charge is collected by a biasing structure and not by a readout implant. Therefore, the operation of PS modules in the CMS environment is mostly unaffected since tracks typically cross sensors at an angle in the yz-plane.

On PS-s sensors, a minor decrease in efficiency is observed in regions closer to the long strip edges due to charge sharing. If a particle crosses near a border of a strip, it induces a signal of smaller amplitude, since a fraction of the charge is collected in the adjacent strip. To demonstrate this behavior, the threshold was increased to 60 units. A projection along y of the in-strip PS-s efficiency map, i.e. efficiency as a function of. in-strip x, is shown in Figure 7.22. In the central part (in x) of the strip the efficiency stays high at above 99%. On the efficiency map, shown in the Figure 7.20, no data is present for central part ($|y| < 8\,\text{mm}$) due to the beam geometry. Given the placement of the PS module in the setup and the layout of the PS-s sensor, the beam spot only covers the upper part of the lower strip row and the lower part of the upper strip row.

[1] The effect is less pronounced for rotation in the xz-plane (as explained in the following sections), since such a rotation would not eliminate the inefficiency at the bias rail.

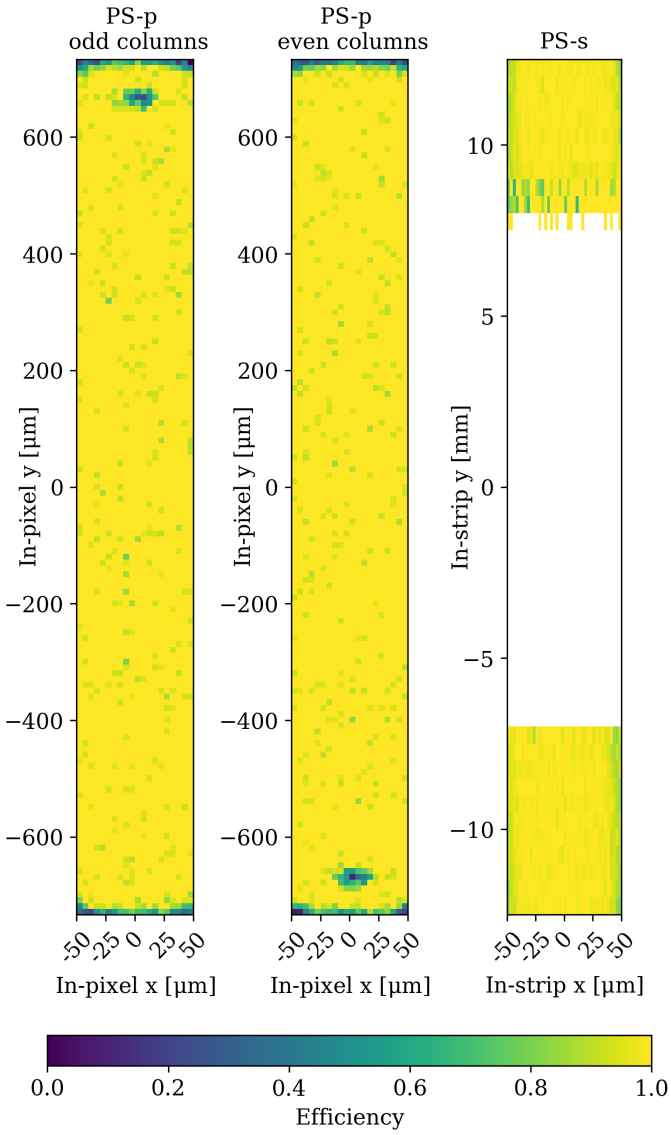


Figure 7.20: In-pixel (in-strip) efficiency maps of PS module sensors. Left: PS-p odd columns; middle: PS-p even columns; right: PS-s.

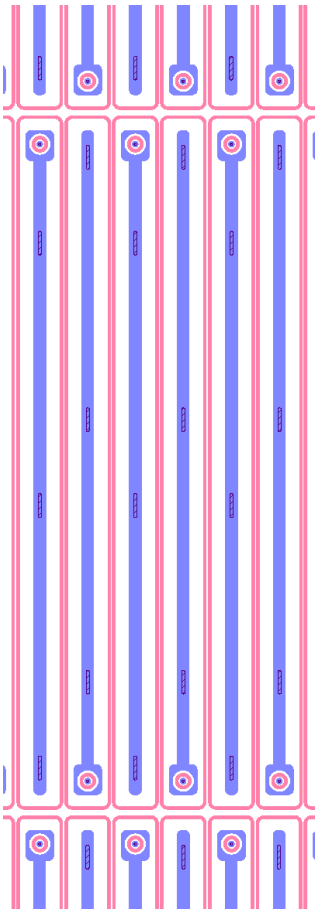


Figure 7.21: Layout of PS-p macro-pixels [199].

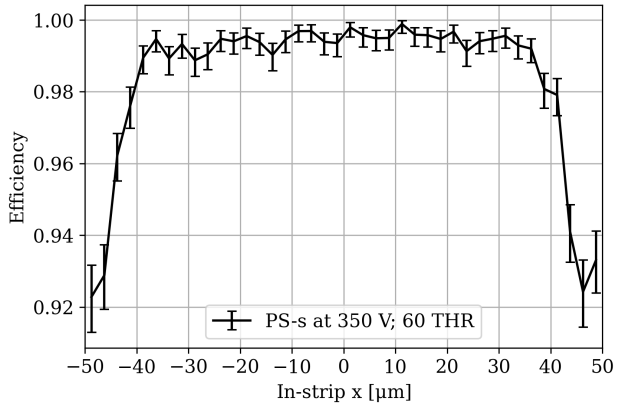


Figure 7.22: Efficiency of the PS-s sensor at THR=60 as a function of distance (in x) between the hit position and the strip center.

7.4.4 Cluster size and angle calibration

In this and following sections, the performance of the PS module is studied as a function of the particle incidence angle.

One of the key characteristics of segmented silicon detectors is the cluster size, i.e. the amount of pixels/strips hit per particle passage, as it directly affects coordinate measurements and detection efficiency. In both sensors of the PS module, given the layout of strips and macro-pixels, the majority of particle hits produce clusters with a size of 1–2 in x and 1 in y. Increasing the track incidence angle increases the mean cluster size (as discussed in Section 7.4.5) due to the increased probability of the particle to traverse several sensor cells. A dependence of the mean cluster size for both sensors on the incidence angle is shown in Figure 7.23.

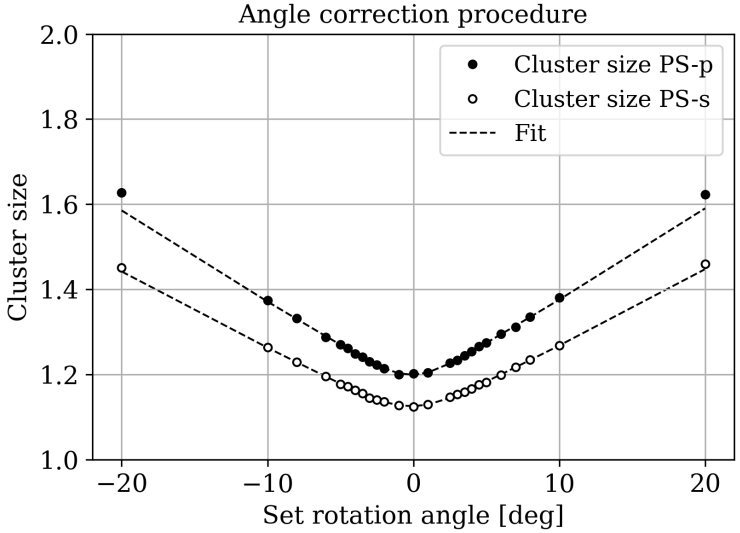


Figure 7.23: Mean cluster size in PS-p (solid markers) and PS-s sensors (open markers) at various track incidence angles. Data is shown together with a fit (dashed line).

The angle shown on the x-axis in Figure 7.23 is the reading of the encoder of the motorized rotation stage. Albeit precise, this angle might have a systematic offset due to mounting of the module box on the turntable and of the stage itself. This offset can be quantified via the mean cluster size measurement shown above and compensated for in the subsequent measurements. This procedure is highly robust, since the investigated observable, the cluster size, does not require track reconstruction.

The angle offset is obtained by fitting the cluster size data with a function [157]:

$$\text{size}(\alpha) = s_0 + k \cdot \int_{-\infty}^{\infty} d\xi \cdot \tan|\xi - \alpha_0| \cdot \exp\left(-\frac{(\xi - \alpha)^2}{2\sigma^2}\right), \tag{7.6}$$

where s_0 and α_0 are the offsets along x- and y-axes, k is the vertical scale, defined by the sensor geometry, the tangent factor describes the geometrical probability to deposit charge in two adjacent channels, and the Gaussian smearing of an arbitrary width σ accounts for applying thresholds.

The offset of -0.14° is obtained as an average of fitted α_0 values for PS-p and PS-s sensors. In the following, this offset is applied as correction to the incidence angle. An additional uncertainty in the track incidence angle may originate from the beam divergence, however, in this work this effect is negligible, as shown previously.

7.4.5 Spatial residual and coordinate resolution

The spatial residual distribution in a PS-p sensor for perpendicular track incidence is shown in Figure 7.24. The principle is the same for the PS-s sensor, although the structure is less pronounced. There are two main components in the residual distribution:

- A box-shaped distribution of a width equal to the pixel pitch (100 µm). It originates from events with one-pixel clusters, since tracks are distributed uniformly across a pixel, and for each of these tracks the center of the pixel is considered the cluster position. The edges of the box are smeared, since events at $(x_{\text{track}} - x_{\text{cluster}}) \simeq \pm 50$ µm are likely to produce a two-pixel cluster instead;

- A peak, originating from two-pixel clusters. At a perpendicular incidence such clusters are only produced by tracks in a close proximity of a pixel border. Therefore, a fraction of these events is relatively minor, but yields fine spatial resolution, because such clusters only appear close to the actual track and thus provide a good estimate of particle passage coordinate. For higher incidence angles (see Figure 7.25) this fraction increases, and the one-pixel distribution is reduced correspondingly. Additionally, a box distribution with a width of 200 µm is observed. It originates from double-pitch pixels on MPA borders, which are treated in the analysis as pairs of pixels with a normal pitch, always firing together. This effect would appear in normal operation of the PS-p sensor, although it would appear in residual for one-pixel clusters.

Also, a negligible amount of events with cluster size of three pixels is present, which is omitted in the analysis. Two sources of these events are hits including double-pitch pixels and hits near pixel borders in y.

The spatial resolution of the sensors is then obtained from a fit. Distributions for each cluster size are fitted separately. The one-pixel distribution is fitted with a box of a height h and borders x_{left} and x_{right}, smeared with a Gaussian of a width σ_{smear}:

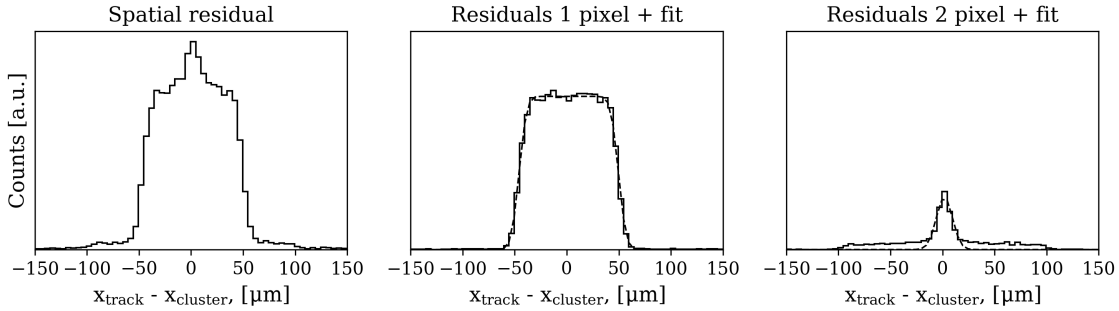


Figure 7.24: Spatial residuals in x for the PS-p sensor at THR=110 and 0° track incidence (left). Separate contribution from one-pixel clusters with a box function fit is shown in the middle, and from two-pixel clusters with a Gaussian fit on the right.

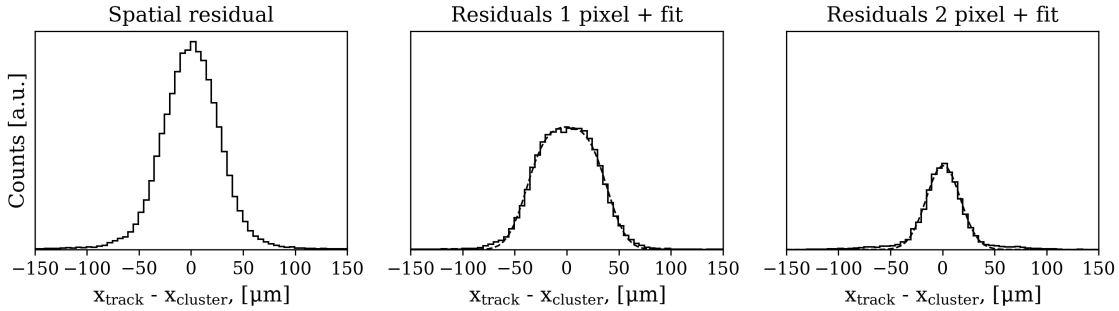


Figure 7.25: Spatial residuals in x for the PS-p sensor at at THR=110 and 10° track incidence (left). Separate contribution from one-pixel clusters with a box function fit is shown in the middle, and from two-pixel clusters with a Gaussian fit on the right.

$$\mathrm{box}(x) = \frac{h}{2} \cdot \left[\mathrm{erf}\left(\frac{x - x_{\mathrm{left}}}{\sqrt{2} \cdot \sigma_{\mathrm{smear}}} \right) - \mathrm{erf}\left(\frac{x - x_{\mathrm{right}}}{\sqrt{2} \cdot \sigma_{\mathrm{smear}}} \right) \right]. \tag{7.7}$$

The two-pixel residual distribution is fitted with a Gaussian curve with a variance $\sigma_{2\mathrm{px}}$. The total standard deviation is then given with a weighted sum:

$$\sigma_{\mathrm{residual}} = \frac{N_{1\mathrm{px}} \cdot \sqrt{\frac{(x_{\mathrm{right}} - x_{\mathrm{left}})^2}{12} + \sigma_{\mathrm{smear}}^2} + N_{2\mathrm{px}} \cdot \sigma_{2\mathrm{px}}}{N_{1\mathrm{px}} + N_{2\mathrm{px}}}, \tag{7.8}$$

where $N_{<..>\mathrm{px}}$ is the number of entries in corresponding distributions.

The spatial resolution of both sensors in the PS module is presented in Figure 7.26. At $\alpha = 0°$, the sensors' resolution is close to the binary resolution limit of pitch$/\sqrt{12} = 28.8\,\mu\mathrm{m}$ due to a relatively low amount of two-pixel clusters under these conditions. The increase of the fraction of two-pixel events by rotation of the module improves the resolution. The effect is slightly stronger in the PS-p sensor due to intrinsically higher charge sharing. As the angle reaches 20°, the effect decreases: the fraction of events with cluster size 2 grows, effectively reverting the detector to the scenario with binary readout. If all clusters are the same size and no charge weighting is employed, a resolution close to the binary limit is expected for any cluster size.

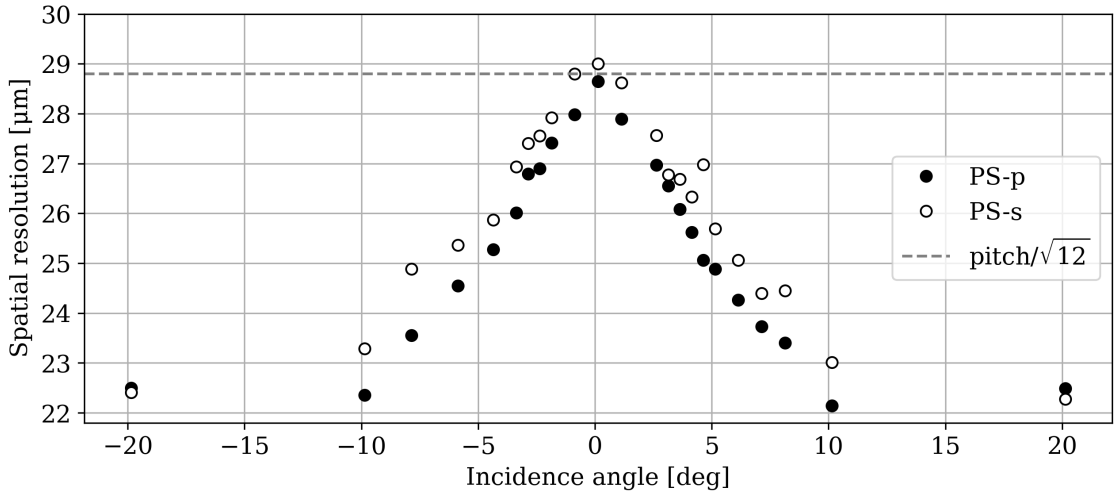


Figure 7.26: Spatial resolution in x of PS-p and PS-s sensors at the working point as a function of the incidence angle.

7.4.6 Stub performance

The angle discrimination mechanism is tested via measuring the *stub efficiency* for various track incidence angles. The stub window size and its offset can be configured in the module front-end with a step of a half-strip, as the MPA and SSA chips perform clustering and may assign half-integer coordinates to clusters. The module is then expected to produce stubs for all tracks at incidence angles α such that:

$$\arctan \frac{p \cdot (-\frac{w}{2} + s)}{\Delta z} \leqslant \alpha \leqslant \arctan \frac{p \cdot (\frac{w}{2} + s)}{\Delta z}, \tag{7.9}$$

where w and s are the configured stub window and its offset, and p and Δz are the pitch and sensor spacing of the module.

A measurement series was performed with a window of 3 strips at an offset of 0. Thus, the `DSY-101` module with a pitch of 100 µm and sensor spacing of 2.6 mm is expected to produce stubs for track incidence $|\alpha| \leqslant 3.3°$. The measured dependence of stub efficiency on the track angle is shown in Figure 7.27. The observed effects are explained in the following.

- Within the acceptance region the module yields high stub efficiency of $98.0 \pm 0.2\%$. Since both sensors need to register the particle to produce a stub, the stub efficiency is limited by the product of sensors' efficiencies:

$$\text{efficiency}_{\text{stub}} \leqslant \text{efficiency}_{\text{PS-p}} \times \text{efficiency}_{\text{PS-s}}. \tag{7.10}$$

- Outside of the acceptance region the stub efficiency drops to 0% within $\sim 1°$ of the border. The transition behavior is explained by the fact that tracks at the same angle may produce 1- or 2-pixel clusters on either plane depending on the in-pixel hit coordinate, which would result in different response of the stub discrimination logic.

- No shift or asymmetry in the stub acceptance region w.r.t. zero is observed. This is ensured by a precise mechanical alignment between the PS-p and PS-s sensor during the module assembly, as for the `DSY-101` module $|\Delta x_{\text{PSp-PSs}}| = 3\,\mu\text{m}$ (see Figure 6.15).

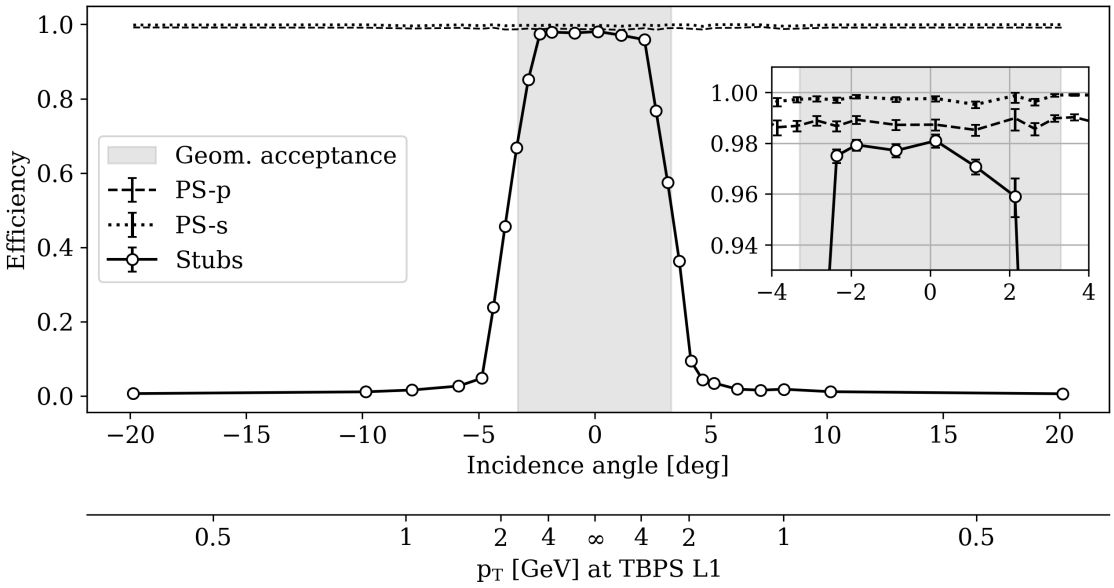


Figure 7.27: Stub efficiency as a function of track incidence angle. Efficiencies of PS-p and PS-s sensors are shown with dashed and dotted lines correspondingly; both sensors are at the working point. The inset axes show a zoom into the central region. The set window size $w = 3$ corresponds to the geometrical acceptance $|\alpha| \leqslant 3.3°$, shown with shading.

Therefore, the module is able to reliably select tracks in the desired range of angles, which validates the p_T-discrimination concept. Transverse momenta, corresponding to given incidence angles (as follows from Eq. 5.2), if the module would have been placed in layer 1 of the TBPS at $R = 248\,\text{mm}$ are shown on the auxiliary x-axis on Figure 7.27. To ensure a p_T cut on 2 GeV, window size of 5 strips is foreseen for TBPS layer 1.

The functionality of the stub window adjustment is presented in Figure 7.28. For various window and offset parameters the measured stub efficiency matches the expectations (as given by Eq. 7.9). A minor decrease in stub efficiency is observed within the acceptance regions at high angles ($|\alpha| \gtrsim 10°$), which is caused by the PS-s sensor being operated at a higher threshold during this measurement series. Thus, the efficiency for angled tracks is reduced (shown in Figure 7.17), therefore the stub efficiency decreases as well (as determined by Eq. 7.10).

7.4.7 Direct measurement of the incidence angle

Each stub incorporates explicit information on the offset (in half-strips) between the cluster centers on PS-p and PS-s planes, called a *bend code*. The bend code can be translated to a track incidence angle via geometrical relation, taking the sensors pitch and spacing in z into account. In a 2.6 mm module it yields an angle measurement at a granularity of $\sim 1.1°$. This on-module track angle estimate will be used in the Level-1 Trigger in the HL-LHC operation of the CMS experiment.

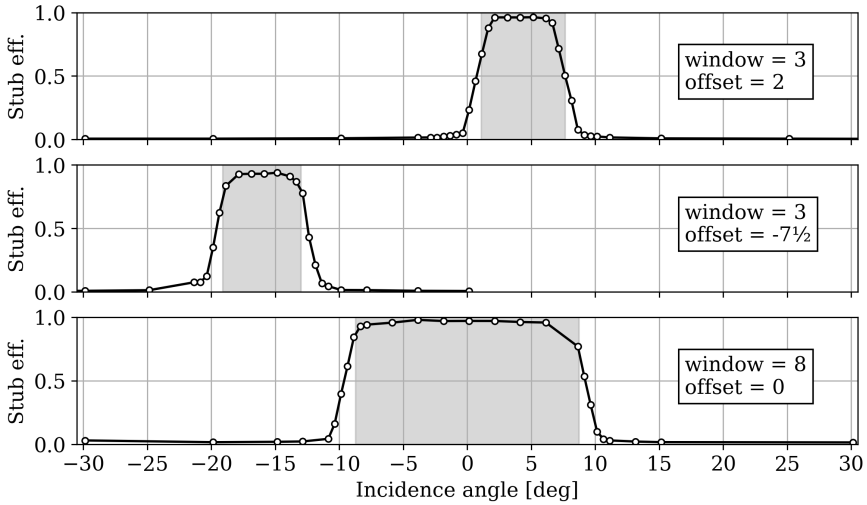


Figure 7.28: Stub efficiency as a function of track incidence angle for various windows and offsets. The PS-s sensor is operated at the threshold of 60 units. The stub acceptance region (Eq. 7.9) is shown with a gray-shaded area.

For a set of parallel tracks at the incidence angle can be estimated more precisely as the mean value of the bend code distribution. Such measurement for a set of tracks with a module with a configured stub window size of 3 strips is presented in Figure 7.29. Within the acceptance, the module can provide a reasonable measurement of the incidence angle of a set of tracks. At the edges and out of acceptance the stub efficiency drops to zero. Therefore, the angle outside the stub acceptance is calculated from a minor amount of stubs randomly generated from noise events. Given the uniform random distribution of noise, the mean angle starts approaching 0°. This technique was proven functional for other window and offset values.

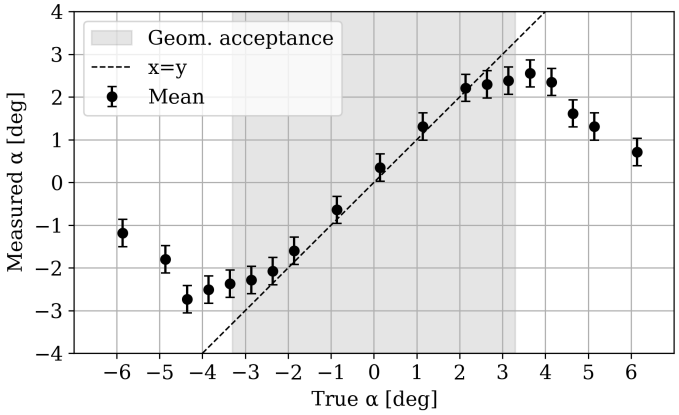


Figure 7.29: Mean track incidence angle measured by the PS module (window of 3 strips, offset=0). The error bars only represent systematic error of $1.1° / \sqrt{12}$, i.e. they do not reflect the mean value being affected by noise outside of the acceptance.

Summary of Part II

The Phase-2 upgrade for the CMS Outer Tracker is an extremely complex endeavor, which requires well-conceived design and massive R&D efforts. Additionally, rigorous testing and quality control are mandatory to ensure excellent performance of the detector. This part of the thesis particularly focused on novel PS detector modules. During the course of this doctoral project, module development and production at DESY was advanced by several stages.

The *prototyping* phase of was completed in 2023. By then, design of the modules and production procedures were mostly established. In particular, the robot-assisted bare module assembly procedure was significantly improved and optimized throughout this work. Thanks to these improvements, the procedure reached the goals on the mechanical precision and throughput. It is ready to meet the challenges of the full-scale module production.

The prototyping was followed by a so-called *kick-off* stage, the goal of which was to build modules with final versions of components in order to verify the production readiness of the design. A batch of five kick-off modules was built at DESY with a major contribution of the author of this work. These kick-off modules were thoroughly tested. One of the modules was extensively characterized at the DESY II test beam facility as a part of this doctoral project. The key results on the module performance are listed in the following.

1. The tested PS module demonstrated all the required functionality; no critical issues were encountered. This serves as a validation for module design and for the production procedures.

2. The performance of both sensors in a PS module was studied in great detail. Under the optimal operation conditions the PS-p and PS-s sensors reach efficiencies of >98% and >99% respectively. In-pixel (in-strip) efficiency variations are studied; both sensors demonstrate uniform efficiency except for specific regions in the PS-p sensor, as expected due to presence of biasing structures.

3. An irradiated PS-s sensor was evaluated. The sensor remains operational and fully efficient having obtained a non-ionizing radiation dose of 1.1×10^{15} n_{eq}/cm^2. This indicates that the PS-s sensors have sufficient radiation hardness for sustained operation at the HL-LHC.

4. Spatial resolution of $\lesssim$28 µm was measured for both sensors, improving with the increase of the incidence angle.

5. The novel p_T-discrimination was thoroughly tested. The module demonstrates $\sim$98% stub efficiency within the target acceptance regions, and 0% outside the acceptance.

6. The on-module track angle measurement was tested. For a single track, the module can estimate the particle passage angle with a granularity of 1.1°.

These findings show that the modules reach all the design goals and demonstrate full functionality. A particularly important feature tested is the p_{T}-discrimination mechanism, which was among the main design drivers for the Outer Tracker and its modules. The results achieved in this work were a significant contribution to the confirmation of the production readiness of the module design and to giving a green light to the module mass production phase.

The module production phase at DESY begun in early 2025. The full production capacity is expected to be reached within 2025.

Conclusion

In this thesis, methods for characterization of silicon sensors are explored. The author has contributed to the development, optimization, and applications of the techniques themselves for advancement of the silicon detector technology for future particle physics experiments.

PS modules for the CMS Outer Tracker upgrade

One of the two projects addressed in this thesis is the Phase-2 Upgrade of the CMS Outer Tracker for the High-Luminosity LHC. On the one hand, the Outer Tracker mostly relies on well-established technology, i.e. hybrid silicon detectors. On the other hand, it features a novel p_{T}-discrimination technique that allows it to contribute to the triggering in the CMS experiment. The development of the upgrade requires significant efforts in production, testing, and quality control to assure excellent performance of the detectors. This thesis is focused on PS modules – individual units of the Outer Tracker, enabled with a novel feature of on-module track angle measurement, which should allow for online event filtering, based on a low-latency cut on the transverse momentum p_{T} of the passing particles. This mechanism will help the CMS experiment to handle significantly higher data rates in the HL-LHC environment.

In the course of this thesis, an automated module construction procedure was developed, optimized, and proven capable of a reliable assembly with $\mathcal{O}(\mu\mathrm{m})$ accuracy, which is crucial for the functioning of the p_{T}-discrimination. The procedure also reaches the required throughput of four modules per day.

One of the modules, built by the author of this work, was then extensively characterized in a test beam, which is a renowned method of testing particle detectors under conditions, resembling those of a real experiment. The macro-pixel and strip sensors of the module have shown high detection efficiencies of 98.7$\pm$0.1% and 99.7$\pm$0.1% respectively, and coordinate resolutions better than 29 µm in both sensors, improving as the particle incidence angle increases. Moreover, proper functioning of the p_{T}-discrimination was verified: the module was shown able to select tracks with a p_{T} above a set threshold with an efficiency of 98.0$\pm$0.2%. This testing campaign served as a validation of the module design and core functionality before the beginning the large-scale production. Thanks to these efforts, the Phase-2 Outer Tracker has entered the production stage and is expected to start operation in 2030.

Sensor characterization with pulsed lasers

Another innovative research direction of the presented thesis is testing silicon sensors with charge injection by the means of pulsed lasers. This flexible and effective technique can be of great use when investigating properties of novel sensor designs. For the purpose of this thesis,

a dedicated setup, the Laserbox, capable of performing such laser-based testing procedures, was commissioned and upgraded by the author of this work.

This thesis highlights the performance of the Laserbox by presenting a testing campaign for the dSiPM prototype – a monolithic silicon pixel single-photon detector with precise timing capabilities. This project investigates a novel technology approach that combines CMOS SPADs with a digital readout. This detector may have potential applications in scenarios requiring detection of ionizing radiation with simultaneously high spatial and temporal resolution. In this thesis, properties of the device were characterized via laser injection with a focus on timing characteristics. The timing resolution of the device was found to be 53 ± 4 ps in the optimal case. Additionally, specific in-pixel variations of the device properties were revealed thanks to the ability of the laser injection method to resolve µm-scale features of the device under test. These variations are linked to the design of SPADs and their arrangement within the individual pixels of the chip. This work delivers a proof of concept for the dSiPM design.

Monte Carlo simulations have become an indispensable tool in detector R&D. In particular, simulations allow one to improve the interpretation of results of testing campaigns and to better understand the sensor under investigation. In the course of this work, a Monte Carlo simulation approach was developed to model laser injection studies, such as those performed in the Laserbox. A dedicated module for the major simulation tool for semiconducting sensors, the Allpix^2 framework, was developed, capable of simulating interaction of pulsed laser beams with semiconductor detectors. The simulation was validated through a detailed comparison with experimental data, obtained with the Laserbox for a silicon strip sensor, similar to those used in the Outer Tracker PS modules. The simulation is able to fairly accurately reproduce the shape of the signal, induced in the sensor, which indicates that major physical effects are properly modeled. The availability of such a simulation tool facilitates investigations of signal formation processes in novel silicon sensors in greater detail by performing cross-correlation between MC and experimental studies.

Detectors with excellent characteristics are always a result of extensive development efforts. Such R&D inevitably involves rigorous prototype characterizations, which is the main focus of this work. This research contributes to pushing the performance limits of silicon detectors, which are a key technology for current and next-generation experiments in particle physics.

Research, presented in this thesis, directly led to the following publications:

[49] I. Diehl et al. “Monolithic MHz-frame rate digital SiPM-IC with sub-100 ps precision and 70 μm pixel pitch”. In: *JINST* 19.01 (Jan. 2024), P01020. DOI: 10.1088/1748-0221/19/01/P01020.

[202] D. Rastorguev et al. “Timing performance of a digital SiPM prototype measured with a picosecond injection laser”. In: *JINST* 20 (June 2025), p. C06035. DOI: 10.1088/1748-0221/20/06/C06035.

Results on the CMS OT Upgrade are expected to be published in 2026 in collaboration-wide Detector Notes:

[200] CMS Tracker authors. “Semi-automated precision assembly of silicon detector modules”. CMS Detector Note. Draft.

[201] CMS Tracker authors. “Test beam performance of an irradiated PS module for the Phase-2 CMS Outer Tracker upgrade”. CMS Detector Note DN-2024/021. Draft.

Appendix A

Allpix^2 configuration for TCT simulation

In this appendix, content of an example main Allpix^2 configuration file for simulating TCT experiments (see Section 3.2.1) is listed. The list of the modules is given, as well as typical parameter values for them.

```
[Allpix]
number_of_events = 1
root_file = "output.root"
# Set detectors geometry
detectors_file = "geometry.conf"

[DepositionLaser]
# Set geometry of the beam
source_position = 0 -0 -2mm
beam_direction = 0 0 1
beam_geometry = "converging"
beam_waist = 100um
focal_distance = 2mm
beam_convergence_angle = 30deg
# Set properties of the pulse
number_of_photons = 100000
wavelength = 1064nm
pulse_duration = 50ps


[DopingProfileReader]
name = "pss"
# Upload from TCAD and properly map to the sensor
model = "mesh"
field_mapping = PIXEL_FULL_INVERSE
field_scale = 0.0005 1
file_name = "doping_4e12.apf"
```

```
[ElectricFieldReader]
name = "pss"
# Upload from TCAD and properly map to the sensor
model = "mesh"
field_mapping = PIXEL_FULL_INVERSE
field_scale = 0.0005 1
# Use a specific file for every bias voltage
file_name = 'efield_300v.apf'

[WeightingPotentialReader]
name = "pss"
# Upload from TCAD and properly map to the sensor
model = "mesh"
field_mapping = PIXEL_FULL
field_scale = 0.0005 1
file_name = "wp.init"
ignore_field_dimensions = true


[TransientPropagation]
# Ensure that each propagation simulation step is short enough
timestep_max = 0.01ns
# Set sufficient integration time to propagate all carriers
integration_time = 20ns

[PulseTransfer]
output_pulsegraphs = true
```

Bibliography

[1] H. Dembinski et al. “Data-driven model of the cosmic-ray flux and mass composition from 10 GeV to 10^{11} GeV”. In: *PoS* ICRC2017 (2017), p. 533. DOI: 10.22323/1.301.0533.

[2] D. J. Bird et al. “Detection of a Cosmic Ray with Measured Energy Well beyond the Expected Spectral Cutoff due to Cosmic Microwave Radiation”. In: *Astrophysics journal* 441 (Mar. 1995), p. 144. DOI: 10.1086/175344.

[3] C. D. Anderson. “The positive electron”. In: *Physical review* 43.6 (1933), p. 491.

[4] C. D. Anderson and S. H. Neddermeyer. “Cloud chamber observations of cosmic rays at 4300 meters elevation and near sea-level”. In: *Physical Review* 50.4 (1936), p. 263.

[5] C. M. G. Lattes et al. “Processes involving charged mesons”. In: *Nature* 159.4047 (1947), pp. 694–697.

[6] CMS Collaboration. “Commissioning of the CMS experiment and the cosmic run at four tesla”. In: *Journal of Instrumentation* 5.03 (Mar. 2010), T03001. DOI: 10.1088/1748-0221/5/03/T03001.

[7] CMS Collaboration. “Alignment of the CMS silicon tracker during commissioning with cosmic rays”. In: *Journal of Instrumentation* 5.03 (Mar. 2010), T03009. DOI: 10.1088/1748-0221/5/03/T03009.

[8] L. W. Alvarez et al. “Berkeley Proton Linear Accelerator”. In: *Review of Scientific Instruments* 26.2 (Feb. 1955), pp. 111–133. ISSN: 0034-6748. DOI: 10.1063/1.1771253.

[9] J. Larmor. “On the theory of the magnetic influence on spectra; and on the radiation from moving ions”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 44.271 (1897), pp. 503–512. DOI: 10.1080/14786449708621095.

[10] R. Fitzpatrick. *Maxwell's Equations and the Principles of Electromagnetism*. Reference, Information and Interdisciplinary Subjects Series. Infinity Science Press, 2008. ISBN: 9781934015209. URL: https://books.google.de/books?id=4QaVSxxnyWwC.

[11] B. H. Wiik. “HERA: accelerator performance and physics results”. In: *From Supersymmetry To The Origin Of Space-time. Proceedings Of The International School Of Subnuclear Physics* 31 (1995), p. 439.

[12] F. Takasaki. “Status of KEKB accelerator and detector, BELLE”. In: *International Journal of Modern Physics A* 15 (2000), pp. 12–21.

[13] CMS Collaboration. “The CMS experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08004–S08004. ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/s08004.

[14] L. Evans and P. Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[15] D. Barney. *CMS Slice*. 2015. URL: https://cds.cern.ch/record/2628641.

[16] The LHCb RICH Collaboration. "Performance of the LHCb RICH detector at the LHC". In: *The European Physical Journal C* 73.5 (May 2013), p. 2431. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-013-2431-9.

[17] A. Bingül. "The ATLAS TRT and its Performance at LHC". In: *Journal of Physics: Conference Series* 347.1 (Feb. 2012), p. 012025. DOI: 10.1088/1742-6596/347/1/012025.

[18] E. Rutherford. "The scattering of α and β particles by matter and the structure of the atom". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 21.125 (1911), pp. 669–688.

[19] F. Salvat. "Bethe stopping-power formula and its corrections". In: *Phys. Rev. A* 106 (3 Sept. 2022), p. 032809. DOI: 10.1103/PhysRevA.106.032809.

[20] H. Kolanoski and N. Wermes. *Particle Detectors: Fundamentals and Applications*. Oxford University Press, 2020. ISBN: 9780198858362.

[21] L. Landau. "On the energy loss of fast particles by ionization". In: *J. Phys. (USSR)* 8 (1944), pp. 201–205.

[22] M. J. Berger. *ESTAR, PSTAR, and ASTAR: Computer programs for calculating stopping-power and range tables for electrons, protons, and helium ions*. NIST Report. 1992.

[23] L. Baudis et al. "Review of Particle Physics". In: *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 2020), p. 083C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptaa104.

[24] *NIST Atomic Spectra Database Ionization Energies Form*. Accessed: 2024-09-05. URL: https://physics.nist.gov/PhysRefData/ASD/ionEnergy.html.

[25] R. Adolphi. "Construction and calibration of the laser alignment system for the CMS tracker". PhD thesis. Aachen, Tech. Hochsch., 2006.

[26] H. Geiger and W. Muller. "Technical considerations concerning the electron counting tube". In: *Phys. Zeit.* 29 (1928), pp. 839–841.

[27] J. G. Wilson. *The principles of cloud-chamber technique*. CUP Archive, 1951.

[28] D. A. Glaser. "The bubble chamber". In: *Scientific American* 192.2 (1955), pp. 46–51.

[29] M. Dracos. "The OPERA experiment". In: *Physics of Atomic Nuclei* 67 (2004), pp. 1092–1096.

[30] A. Ariga et al. "FASER: forward search experiment at the LHC". In: *arXiv:1901.04468* (2019).

[31] G. Charpak and F. Sauli. "Multiwire proportional chambers and drift chambers". In: *Nuclear Instruments and Methods* 162.1-3 (1979), pp. 405–428.

[32] C. Kittel. *Introduction to Solid State Physics*. Wiley series on the science and technology of materials. Wiley, 1959. URL: https://books.google.de/books?id=-Hy0AAAAIAAJ.

[33] C. Leroy and P. Rancoita. *Silicon solid state devices and radiation detection*. World scientific, 2012.

[34] B. D. Milbrath et al. "Radiation detector materials: An overview". In: *Journal of Materials Research* 23.10 (2008), pp. 2561–2581. DOI: 10.1557/JMR.2008.0319.

[35] A. Tyazhev et al. “GaAs radiation imaging detectors with an active layer thickness up to 1mm”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 509.1 (2003). Proceedings of the 4th Internatonal Workshop on Radiation Imaging Detectors, pp. 34–39. ISSN: 0168-9002. DOI: 10.1016/S0168-9002(03)01545-6.

[36] M. Krammer et al. “Status of diamond particle detectors”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 418.1 (1998), pp. 196–202. ISSN: 0168-9002. DOI: 10.1016/S0168-9002(98)00733-5.

[37] S. M. Sze, Y. Li, and K. K. Ng. *Physics of semiconductor devices*. John Wiley & Sons, 2021.

[38] M. A. Green and M. J. Keevers. “Optical properties of intrinsic silicon at 300 K”. In: *Progress in Photovoltaics: Research and Applications* 3.3 (1995), pp. 189–192. DOI: 10.1002/pip.4670030303.

[39] F. Hartmann. *Evolution of silicon sensor technology in particle physics*. Vol. 43. Springer, 2009.

[40] P. Drude. “Zur Elektronentheorie der Metalle”. In: *Annalen der Physik* 306.3 (1900), pp. 566–613. DOI: 10.1002/andp.19003060312.

[41] W. Shockley. “Currents to Conductors Induced by a Moving Point Charge”. In: *Journal of Applied Physics* 9.10 (1938), pp. 635–636. DOI: 10.1063/1.1710367.

[42] S. Ramo. “Currents Induced by Electron Motion”. In: *Proceedings of the IRE* 27.9 (1939), pp. 584–585. DOI: 10.1109/JRPROC.1939.228757.

[43] M. Klein et al. “Single chip bumping and reliability for flip chip processes”. In: *Microelectronics Reliability* 39.9 (1999), pp. 1389–1397. ISSN: 0026-2714. DOI: 10.1016/S0026-2714(99)00067-0.

[44] F. Reidt. “Upgrade of the ALICE ITS detector”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1032 (2022), p. 166632. ISSN: 0168-9002. DOI: 10.1016/j.nima.2022.166632.

[45] R. L. Gluckstern. “Uncertainties in track momentum and direction, due to multiple scattering and measurement errors”. In: *Nuclear Instruments and Methods* 24 (1963), pp. 381–389.

[46] G. Lutz. *Semiconductor Radiation Detectors*. Springer, 2007. ISBN: 9783540716792. DOI: 10.1007/978-3-540-71679-2.

[47] C. Claeys and E. Simoen. *Radiation Effects in Advanced Semiconductor Materials and Devices*. Springer, 2002. ISBN: 9783662049747. DOI: 10.1007/978-3-662-04974-7.

[48] R. Wunstorf. “Systematische Untersuchungen zur Strahlenresistenz von Silizium-Detektoren für die Verwendung in Hochenergiephysik-Experimenten”. PhD thesis. University of Hamburg, 1992. URL: https://bib-pubdb1.desy.de/record/153817.

[49] I. Diehl et al. “Monolithic MHz-frame rate digital SiPM-IC with sub-100 ps precision and 70 μm pixel pitch”. In: *JINST* 19.01 (Jan. 2024), P01020. DOI: 10.1088/1748-0221/19/01/P01020.

[50] *The Phase-2 Upgrade of the CMS Tracker*. Tech. rep. Geneva: CERN, 2017. DOI: 10.17181/CERN.QZ28.FLHW.

[51] J. Braach. "Functional Tests of 2S Detector Modules for the CMS Phase 2 Outer Tracker Upgrade including the Development of an IR LED Array". Master's thesis. Karlsruhe Institute of Technology (KIT), 2020.

[52] O. Svelto, D. C. Hanna et al. *Principles of lasers*. Vol. 1. Springer, 2010.

[53] Y. Unno et al. "Development of n+-in-p large-area silicon microstrip sensors for very high radiation environments – ATLAS12 design and initial results". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 765 (Nov. 2014), pp. 80–90. ISSN: 0168-9002. DOI: 10.1016/j.nima.2014.06.086.

[54] G. Kramberger et al. "Investigation of Irradiated Silicon Detectors by Edge-TCT". In: *IEEE Transactions on Nuclear Science* 57.4 (2010), pp. 2294–2302. DOI: 10.1109/TNS.2010.2051957.

[55] R. Geertsema et al. "Charge and temporal characterisation of silicon sensors using a two-photon absorption laser". In: *Journal of Instrumentation* 17.02 (Feb. 2022), P02023. ISSN: 1748-0221. DOI: 10.1088/1748-0221/17/02/P02023.

[56] M. G. Fernández Garcıa et al. "High-resolution three-dimensional imaging of a depleted CMOS sensor using an edge Transient Current Technique based on the Two Photon Absorption process (TPA-eTCT)". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 845 (2017), pp. 69–71.

[57] M. G. Fernández Garcıa and S. Otero Ugobono. *The Transient Current Technique: laser characterization of silicon detectors.* Seminar. CERN, 2017. URL: https://indico.cern.ch/event/684193/attachments/1567682/2471443.

[58] G. Kramberger et al. "Determination of detrapping times in semiconductor detectors". In: *Journal of Instrumentation* 7.04 (Apr. 2012), P04006. ISSN: 1748-0221. DOI: 10.1088/1748-0221/7/04/p04006.

[59] G. Kramberger. "Advanced TCT setups". In: *The 23rd International Workshop on Vertex Detectors*. Vol. 227. SISSA Medialab. 2015, p. 032.

[60] T. Vanat. "Caribou – A versatile data acquisition system". In: *PoS* TWEPP2019 (2020), p. 100. DOI: 10.22323/1.370.0100.

[61] *Particulars Pulsed Lasers*. Particulars, advanced measurement systems. URL: https://www.particulars.si/products.php?prod=lasers.html.

[62] *ALS PiLas Picosecond Laser Diode System*. Advanced Laser Diode Systems A.L.S. GmbH. URL: http://www.optophase.com/als_laser.html.

[63] *BE15M-B Beam expander*. Thorlabs Inc. URL: https://www.thorlabs.de/thorproduct.cfm?partnumber=BE15M-B.

[64] *DET08CFC/M 5GHz InGaAs FC/PC-Coupled Photodetector*. Thorlabs Inc. URL: https://www.thorlabs.de/thorproduct.cfm?partnumber=DET08CFC/M.

[65] *DET02AFC 1GHz Si FC/PC-Coupled Photodetector*. Thorlabs Inc. URL: https://www.thorlabs.de/thorproduct.cfm?partnumber=DET02AFC.

[66] *8MT30-50 Motorized Translation Stages*. STANDA LTD. URL: https://www.standa.lt/products/catalog/motorised_positioners?item=348.

[67] *BT-01 Bias-T*. Particulars, advanced measurement systems. URL: https://particulars.si/products.php?prod=biast.html.

[68] *AM-01A RF Amplifier*. Particulars, advanced measurement systems. URL: https://particulars.si/products.php?prod=amplifiers.html.

[69] *LeCroy WR640Zi oscilloscope*. TELEDYNE LECROY. URL: https://cdn.teledynelecroy.com/files/pdf/waverunner-6zi-datasheet.pdf.

[70] *Peary – DAQ framework for the Caribou system*. URL: https://caribou-project.docs.cern.ch/docs/05_peary_software/.

[71] D. Rastorguev. *TCT controller software repositiory*. URL: https://github.com/dannrastor/tct-controller.

[72] S. Ruiz Daza et al. *Test Beam characterization of the H2M chip designed in a 65 nm CMOS imaging process*. Poster. 12th Beam Telescopes and Test Beams Workshop, Edinburgh, 2024. URL: https://indico.cern.ch/event/1323113/contributions/5823792/.

[73] G. Vignola. *Private communication*. DESY. 2024.

[74] A. Simancas et al. "Simulations and performance studies of a MAPS in 65 nm CMOS imaging technology". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1064 (2024), p. 169414. ISSN: 0168-9002. DOI: 10.1016/j.nima.2024.169414.

[75] N. Metropolis et al. "Equation of State Calculations by Fast Computing Machines". In: *The Journal of Chemical Physics* 21.6 (June 1953), pp. 1087–1092. ISSN: 0021-9606. DOI: 10.1063/1.1699114.

[76] D. Dannheim et al. "Combining TCAD and Monte Carlo methods to simulate CMOS pixel sensors with a small collection electrode using the Allpix Squared framework". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 964 (June 2020), p. 163784. ISSN: 0168-9002. DOI: 10.1016/j.nima.2020.163784.

[77] P. Li. "Simulation and measurement of charge transport in the guard region of planar silicon sensors". Bachelor's thesis. 2024. DOI: 10.18452/28144.

[78] I. Bloch et al. *Simulation of charge transport at the surface of planar silicon sensors*. Preprint. 2025. URL: https://arxiv.org/abs/2505.14203.

[79] S. Agostinelli et al. "GEANT4—a simulation toolkit". In: *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303.

[80] L. W. Nagel and D. O. Pederson. *SPICE (Simulation Program with Integrated Circuit Emphasis)*. Tech. rep. UCB/ERL M382. UC Berkeley, Apr. 1973. URL: http://www2.eecs.berkeley.edu/Pubs/TechRpts/1973/22871.html.

[81] S. Spannagel et al. "Allpix^2: A modular simulation framework for silicon detectors". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 901 (Sept. 2018), pp. 164–172. ISSN: 0168-9002. DOI: 10.1016/j.nima.2018.06.020.

[82] S. Spannagel, H. Wennlöf, and P. Schütze. *Allpix Squared - Generic Pixel Detector Simulation Framework*. Version 3.0.0. May 2023. DOI: 10.5281/zenodo.8171803.

[83] I. Antcheva et al. “ROOT—A C++ framework for petabyte data storage, statistical analysis and visualization”. In: *Computer Physics Communications* 182.6 (2011), pp. 1384–1385.

[84] C. Hagmann, D. Lange, and D. Wright. “Cosmic-ray shower generator (CRY) for Monte Carlo transport codes”. In: *2007 IEEE Nuclear Science Symposium Conference Record*. Vol. 2. 2007, pp. 1143–1146. DOI: 10.1109/NSSMIC.2007.4437209.

[85] D. Dannheim et al. “Corryvreckan: a modular 4D track reconstruction and analysis software for test beam data”. In: *Journal of Instrumentation* 16.03 (2021), P03008.

[86] T. Bisanz et al. “EUTelescope: A modular reconstruction framework for beam telescope data”. In: *Journal of Instrumentation* 15.09 (2020), P09020.

[87] D. Rastorguev. *DepositionLaser User Manual*. URL: https://allpix-squared.docs.cern.ch/docs/08_modules/depositionlaser/.

[88] D. Rastorguev. *DepositionLaser module of Allpix Squared*. Software Repository. URL: https://gitlab.cern.ch/allpix-squared/allpix-squared/-/tree/master/src/modules/DepositionLaser.

[89] A. Schaelicke. *Optical Physics. Geant4 Tutorial*. CERN, 2010. URL: https://indico.cern.ch/event/58317/contributions/2047457/.

[90] Y. Liang and B. A. Barsky. “A New Concept and Method for Line Clipping”. In: *ACM Transactions on Graphics* 3.1 (Jan. 1984), pp. 1–22. ISSN: 1557-7368. DOI: 10.1145/357332.357333.

[91] T. V. Eichhorn and J. Rübenach. *DESY FH E-Lab Probe Station*. Software repository and user manual. 2019. URL: https://github.com/DESY-FH-ELab/probestation.

[92] A. Velyka. *Private communication*. DESY, 2025.

[93] *Sentaurus TCAD*. Synopsis Inc. URL: https://www.synopsys.com/manufacturing/tcad.html.

[94] H. Wennlöf et al. “Simulating monolithic active pixel sensors: A technology-independent approach using generic doping profiles”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1073 (2025), p. 170227. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2025.170227.

[95] H. Spieler. *Semiconductor detector systems*. Vol. 12. Oxford university press, 2005.

[96] H. Sadrozinski, A. Seiden, and N. Cartiglia. “4D tracking with ultra-fast silicon detectors”. In: *Rep. Prog. Phys.* 81.2 (Dec. 2017), p. 026101. DOI: 10.1088/1361-6633/aa94d3.

[97] P. Fischer, R. K. Zimmermann and B. Maisano. “CMOS SPAD sensor chip for the readout of scintillating fibers”. In: *NIMA* 1040 (2022), p. 167033. DOI: 10.1016/j.nima.2022.167033.

[98] J. Zhou. “A Review of LiDAR sensor Technologies for Perception in Automated Driving”. In: *Academic Journal of Science and Technology* 3.3 (Nov. 2022), pp. 255–261. DOI: 10.54097/ajst.v3i3.2993.

[99] M. N. Ullah et al. “Instrumentation for Time-of-Flight Positron Emission Tomography”. In: *Nucl. Med. Mol. Imaging* 50.2 (Feb. 2016), pp. 112–122. DOI: 10.1007/s13139-016-0401-5.

[100] M. Ferrero et al. *An Introduction to Ultra-Fast Silicon Detectors*. CRC Press, July 2021. DOI: 10.1201/9781003131946.

[101] W. Riegler. *Time resolution limits in silicon sensors from Landau fluctuations and electronics noise*. 2025. arXiv: 2504.02570.

[102] 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.

[103] G. Dalla Betta and J. Ye. "Silicon Radiation Detector Technologies: From Planar to 3D". In: *Chips* 2.2 (Apr. 2023), pp. 83–101. DOI: 10.3390/chips2020006.

[104] P. Windischhofer and W. Riegler. "The statistics of electron–hole avalanches". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1003 (2021), p. 165327. ISSN: 0168-9002. DOI: 10.1016/j.nima.2021.165327.

[105] J. Townsend. *The theory of ionization of gases by collision*. Constable, Limited, 1910.

[106] G. Giacomini. "LGAD-Based Silicon Sensors for 4D Detectors". In: *Sensors* 23.4 (Feb. 2023), p. 2132. DOI: 10.3390/s23042132.

[107] G. E. Stillman and C. M. Wolfe. "Chapter 5. Avalanche Photodiodes." In: *Semiconductors and Semimetals*. Ed. by R. K. Willardson and A. C. Beer. Vol. 12. Semiconductors and Semimetals. Elsevier, 1977, pp. 291–393. DOI: 10.1016/S0080-8784(08)60150-7.

[108] F. Carnesecchi et al. "Direct detection of charged particles with SiPMs". In: *Journal of Instrumentation* 17.06 (June 2022), P06007. DOI: 10.1088/1748-0221/17/06/p06007.

[109] R. Klanner and F. Sauli, eds. *Silicon Photomultipliers: Technology, Characterisation and Applications* 926 (May 2019). ISSN: 0168-9002. URL: https://www.sciencedirect.com/journal/nuclear-instruments-and-methods-in-physics-research-section-a-accelerators-spectrometers-detectors-and-associated-equipment/vol/926/suppl/C.

[110] M. Ngo et al. "Shallow trench isolation process". US US7648886B2. Filed: January 14, 2003. Priority date: January 14, 2003. Jan. 2010. URL: https://patents.google.com/patent/US7648886B2/en.

[111] I. Diehl et al. "Readout ASIC for fast digital imaging using SiPM sensors: Concept study". In: *2015 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. IEEE, Oct. 2015, pp. 1–3. DOI: 10.1109/nssmic.2015.7581816.

[112] I. Diehl et al. "Readout of digital SiPMs". In: *2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)*. IEEE, Nov. 2018, pp. 1–3. DOI: 10.1109/nssmic.2018.8824395.

[113] F. Acerbi and M. Perenzoni. "High Sensitivity Photodetector for Photon-Counting Applications". In: *Photon Counting - Fundamentals and Applications*. InTech, Mar. 2018. DOI: 10.5772/intechopen.71940.

[114] H. Xu et al. "Design and characterization of a p+/n-well SPAD array in 150nm CMOS process". In: *Optics Express* 25.11 (May 2017), p. 12765. DOI: 10.1364/oe.25.012765.

[115] I. Diehl et al. "The DESY digital silicon photomultiplier: Device characteristics and first test-beam results". In: *NIMA* 1064 (2024), p. 169321. DOI: 10.1016/j.nima.2024.169321.

[116] F. King et al. *Test Beam Characterization of a Digital Silicon Photomultiplier*. 2024. URL: https://arxiv.org/abs/2412.06687.

[117] S. Lachnit. "Time Resolution of a Fully-Integrated Digital Silicon Photo-Multiplier". MA thesis. Hamburg U., 2023. DOI: 10.3204/PUBDB-2024-00529.

[118] N. G. Worku and H. Gross. "Propagation of truncated Gaussian beams and their application in modeling sharp-edge diffraction". In: *J. Opt. Soc. Am. A* 36.5 (May 2019), pp. 859–868. DOI: 10.1364/JOSAA.36.000859.

[119] G. Vignola. "Studies of a Digital SiPM and MAPS Prototypes as Key Technologies for Future High-Energy Physics Experiments". PhD thesis. Rheinische Friedrich-Wilhelms-Universität Bonn, Apr. 2025. DOI: 10.48565/bonndoc-546.

[120] S. Baker and R. D. Cousins. "Clarification of the use of chi-square and likelihood functions in fits to histograms". In: *Nuclear Instruments and Methods in Physics Research* 221.2 (1984), pp. 437–442.

[121] *Allpix Squared v2.4.0 release notes*. Jan. 2023. URL: https://allpix-squared.docs.cern.ch/post/2023-01-12-allpix-squared-2.4-released/.

[122] CMS Collaboration. "Evidence for the direct decay of the 125 GeV Higgs boson to fermions". In: *Nature Physics* 10.8 (June 2014), pp. 557–560. ISSN: 1745-2481. DOI: 10.1038/nphys3005.

[123] ATLAS collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 1–29. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.020.

[124] J. Womersley. *The LHC physics program*. Tech. rep. Fermi National Accelerator Lab.(FNAL), Batavia, IL (United States), 1997.

[125] E. Lopienska. *The CERN accelerator complex, layout in 2022*. 2022. URL: https://cds.cern.ch/record/2800984.

[126] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[127] LHCb Collaboration. "The LHCb Detector at the LHC". In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[128] ALICE Collaboration. In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.

[129] O. Aberle et al. "High-luminosity large hadron collider (HL-LHC): Technical design report". In: (2020).

[130] *Plan of LHC/HL-LHC*. CERN. URL: https://hilumilhc.web.cern.ch/content/hl-lhc-project.

[131] D. Contardo et al. *Technical Proposal for the Phase-II Upgrade of the CMS Detector*. Tech. rep. Geneva, 2015. DOI: 10.17181/CERN.VU8I.D59J.

[132] J. Butler et al. *CMS Phase II Upgrade Scope Document*. Tech. rep. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/2055167.

[133] T. Sakuma. *Cutaway diagrams of CMS detector*. 2019. URL: https://cds.cern.ch/record/2665537.

[134] CMS Collaboration. *CMS: letter of intent by the CMS Collaboration for a general purpose detector at LHC*. Tech. rep. Geneva: CERN, 1992. URL: https://cds.cern.ch/record/290808.

[135] *The CMS magnet project: Technical Design Report*. Technical design report. CMS. Geneva: CERN, 1997. DOI: 10.17181/CERN.6ZU0.V4T9.

[136] J. G. Layter. *The CMS muon project: Technical Design Report*. Technical design report. CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/343814.

[137] *The Phase-2 Upgrade of the CMS Muon Detectors*. Tech. rep. This is the final version, approved by the LHCC. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2283189.

[138] *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical design report. CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/349375.

[139] *The Phase-2 Upgrade of the CMS Barrel Calorimeters*. Tech. rep. This is the final version, approved by the LHCC. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2283187.

[140] *The CMS hadron calorimeter project: Technical Design Report*. Technical design report. CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/357153.

[141] J. Mans et al. *CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter*. Tech. rep. 2012. URL: https://cds.cern.ch/record/1481837.

[142] *The Phase-2 Upgrade of the CMS Endcap Calorimeter*. Tech. rep. Geneva: CERN, 2017. DOI: 10.17181/CERN.IV8M.1JY2.

[143] V. Karimäki et al. *The CMS tracker system project: Technical Design Report*. Technical design report. CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/368412.

[144] *The CMS tracker: addendum to the Technical Design Report*. Technical design report. CMS. Geneva: CERN, 2000. URL: https://cds.cern.ch/record/490194.

[145] A. Dominguez et al. *CMS Technical Design Report for the Pixel Detector Upgrade*. Tech. rep. Additional contacts: Jeffrey Spalding, Fermilab, Jeffrey.Spalding@cern.ch Didier Contardo, Universite Claude Bernard-Lyon I, didier.claude.contardo@cern.ch. 2012. URL: https://cds.cern.ch/record/1481838.

[146] M. Miñano Moya. "CMS pixel upgrade for the phase I: Module production and qualification". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 831 (2016). Proceedings of the 10th International "Hiroshima" Symposium on the Development and Application of Semiconductor Tracking Detectors, pp. 137–139. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2016.03.030.

[147] CMS Collaboration. *CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems*. Technical design report. CMS. URL: https://cds.cern.ch/record/706847.

[148] A. Tapper and D. Acosta. *CMS Technical Design Report for the Level-1 Trigger Upgrade*. Tech. rep. 2013. URL: https://cds.cern.ch/record/1556311.

[149] *The Phase-2 Upgrade of the CMS L1 Trigger Interim Technical Design Report*. Tech. rep. This is the CMS Interim TDR devoted to the upgrade of the CMS L1 trigger in view of the HL-LHC running, as approved by the LHCC. Geneva: CERN, 2017. DOI: 10.17181/CERN.UUWZ.FMIS. URL: https://cds.cern.ch/record/2283192.

[150] S. Cittolin, A. Rácz, and P. Sphicas. *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project*. Technical design report. CMS. Geneva: CERN, 2002. URL: https://cds.cern.ch/record/578006.

[151] CMS Collaboration. *The Phase-2 Upgrade of the CMS Data Acquisition and High Level Trigger*. Tech. rep. This is the final version of the document, approved by the LHCC. Geneva: CERN, 2021. URL: https://cds.cern.ch/record/2759072.

[152] D. Hufnagel et al. "The architecture and operation of the CMS Tier-0". In: *Journal of Physics: Conference Series* 331.3 (Dec. 2011), p. 032017. DOI: 10.1088/1742-6596/331/3/032017.

[153] I. Bird. "Computing for the large hadron collider". In: *Annual Review of Nuclear and Particle Science* 61.1 (2011), pp. 99–118.

[154] A. Mussgiler. *Private communication*. DESY, 2024.

[155] D. Barney and T. McCauley. *Sketchup images of HL-CMS highlighting the sub-detectors*. CERN. URL: https://cms-docdb.cern.ch/cgi-bin/PublicDocDB//ShowDocument?docid=14785.

[156] C. Cheng. "Test Beam Studies of Prototype Module and Test Measurements of MaPSA for the CMS Outer Tracker Phase-2 Upgrade". PhD dissertation. Universität Hamburg, 2025, p. 174. DOI: 10.3204/PUBDB-2025-01019.

[157] Y. Otarid. "Pixel-Strip Modules for the CMS Tracker Phase-2 Upgrade: From DAQ test system development to module assembly and qualification". PhD thesis. 2023. DOI: 10.3204/PUBDB-2023-06263.

[158] A. Ventura Barroso. "From detectors to data: Thermal Characterization of CMS Outer Tracker modules and Measurement of the top quark mass using $t\bar{t}$ + jet events". PhD dissertation. Bergische Universität Wuppertal, 2025, p. 248. DOI: 10.3204/PUBDB-2025-00758.

[159] L. Sommer. "Qualification of Silicon Detectors for the Phase 2 Upgrade of the CMS Outer Tracker". Master's thesis. Universität Wuppertal, 2023, p. 103. DOI: 10.3204/PUBDB-2023-07550.

[160] *SYLGARD 186 Silicone Elastomer*. Dow Inc. URL: https://www.dow.com/en-us/pdp.sylgard-186-silicone-elastomer-kit.01040286z.html.

[161] Y. Gavrikov. *Private communication*. DESY, 2025.

[162] *Rotary Stage RT5-NSMA*. Walter Uhl technische Mikroskopie GmbH & Co. KG. URL: https://www.walteruhl.com/index.php/en/products/positioning-systems/rotary-stages.

[163] *XY Stage GT9-NSMA*. Walter Uhl technische Mikroskopie GmbH & Co. KG. URL: https://www.walteruhl.com/index.php/en/products/positioning-systems/xy-stages.

[164] *Linear Stage LT8-LBMA*. Walter Uhl technische Mikroskopie GmbH & Co. KG. URL: https://www.walteruhl.com/index.php/en/products/positioning-systems/linear-stages.

[165] Open Source Vision Foundation. *OpenCV: Template Matching reference*. URL: https://docs.opencv.org/3.4/de/da9/tutorial_template_matching.html.

[166] CMS Tracker Collaboration. *CMS Phase-2 OT Module Assembly software repository*. URL: https://github.com/DESY-FH-ELab/cmstkmodlab/tree/master/assembly.

[167] Open Source Vision Foundation. *OpenCV: open-source computer vision library*. URL: https://opencv.org/.

[168] *EP 601-LV adhesive*. Polytec PT GmbH. URL: https://www.polytecstore.fr/polytec_images/documents/FT_PT/ep/polytec-ep-601-lv_en.pdf.

[169] *preeflow eco-PEN300 dispensing unit*. ViscoTec Pumpen- u. Dosiertechnik GmbH. URL: https://dosieren.de/media/2d/a1/8a/1631521333/eco-pen300_dat_de.pdf.

[170] *VR3203 table dispensing robot*. VIEWEG GmbH. URL: https://dosieren.de/produkte/dosierroboter/tischroboter/505935/tischroboter-vr3200-purple.

[171] *LOCTITE EA 3430 adhesive*. Henkel AG & Co. URL: https://datasheets.tdx.henkel.com/LOCTITE-EA-3430-en_GL.pdf.

[172] *SmartScope Flash CNC 300*. Optical Gaging Products Inc. URL: https://www.ogpnet.com/products/metrology-systems/multisensor-metrology-systems/smartscope-video-multisensor-systems/smartscope-flash/smartscope-flash-302/.

[173] *Needle sensor for Dispensing Systems*. VIEWEG GmbH. URL: https://dosieren.de/produkte/dosierroboter/zubehoer/sonstiges/504161/nadelsensor.

[174] R. Diener et al. "The DESY II test beam facility". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 922 (2019), pp. 265–286. DOI: 10.1016/j.nima.2018.11.133.

[175] A. Harb. "Novel Trigger-Capable Modules for the Future CMS Tracking Detector and Inclusive Top Quark Pair Production Cross Section at $\sqrt{s} = 13$TeV". Dissertation. Hamburg: Universität Hamburg, 2017, p. 200. DOI: 10.3204/PUBDB-2017-07788.

[176] L. Stockmeier. "System and Integration Tests with 2S Module Prototypes for the Phase-2 Upgrade of the CMS Outer Tracker". PhD thesis. Karlsruhe Institute of Technology (KIT), 2025. URL: https://publish.etp.kit.edu/record/22323.

[177] *CERN Test Beam Facilities Database*. URL: https://test-beam-facilities.web.cern.ch/.

[178] G. Hemmie. "DESY II, a New Injector for the DESY Storage Rings PETRA and DORIS II". In: *IEEE Transactions on Nuclear Science* 30.4 (1983), pp. 2028–2030. DOI: 10.1109/TNS.1983.4332706.

[179] H. Schopper. *The start of PETRA and its experimental programme*. Tech. rep. International symposium perspective and future projects in high energy physics, Tokyo, Japan. 1979.

[180] P. Waloschek. "DORIS II — ein (fast) neuer Speicherring". In: *Physikalische Blätter* 38.1 (1982), pp. 7–7. DOI: 10.1002/phbl.19820380107.

[181] K. Balewski et al. "Commissioning of Petra III". In: *Proceedings of 1st International Particle Accelerator Conference: IPAC*. Vol. 10. 2010.

[182] H. Jansen et al. "Performance of the EUDET-type beam telescopes". In: *EPJ Techniques and Instrumentation* 3.1 (Oct. 2016). DOI: 10.1140/epjti/s40485-016-0033-2.

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

[184] C. Hu-Guo et al. "A ten thousand frames per second readout MAPS for the EUDET beam telescope". In: *Topical Workshop on Electronics for Particle Physics*. CERN, 2009. DOI: 10.5170/CERN-2009-006.47.

[185] P. Baesso, D. Cussans, and J. Goldstein. "The AIDA-2020 TLU: a flexible trigger logic unit for test beam facilities". In: *Journal of Instrumentation* 14.09 (2019), P09019.

[186] S. Spannagel. H. C. Perrey, and B. Meier. *The pxarCore Library-Technical Documentation, Reference Manual, and Sample Applications*. Tech. rep. CERN, 2015.

[187] A. Caratelli et al. "OT-μDTC, a test bench for testing CMS Outer Tracker Phase-2 module prototypes". In: *PoS* EPS-HEP2019 (2020), p. 170. DOI: 10.22323/1.364.0170.

[188] Y. Liu et al. "EUDAQ2—A flexible data acquisition software framework for common test beams". In: *Journal of Instrumentation* 14.10 (Oct. 2019), P10033. DOI: 10.1088/1748-0221/14/10/P10033.

[189] *EUDAQ2 Software Repository*. Accessed: 2025-02-03. URL: https://github.com/eudaq/eudaq.

[190] *Corryvreckan Software Repository*. Accessed: 2025-03-03. URL: https://gitlab.cern.ch/corryvreckan/corryvreckan.

[191] *Corryvreckan branch for PS test beam*. Accessed: 2025-03-03. URL: https://gitlab.cern.ch/drastorg/corryvreckan/-/tree/daniil/ps-testbeam-nov23.

[192] C. Kleinwort. "General broken lines as advanced track fitting method". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 673 (2012), pp. 107–110. DOI: 10.1016/j.nima.2012.01.024.

[193] K. Pearson. "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 50.302 (1900), pp. 157–175. DOI: 10.1080/14786440009463897.

[194] F. James and M. Roos. "MINUIT - Function Minimization and Error Analysis". In: *CERN program library* D506 (1994).

[195] S. Spannagel and H. Jansen. *GBL Track Resolution Calculator v2.0*. 2016. DOI: 10.5281/zenodo.48795. URL: https://github.com/simonspa/resolution-simulator.

[196] 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.

[197] S. Lachnit and G. Vignola. *Private communication – EM interference in the test beam hall*. DESY. 2025.

[198] J. Großmann. "Silicon sensor prototypes and module concepts for the CMS tracker for the high luminosity large hadron collider". PhD thesis. TU Wien, 2018. DOI: 10.34726/HSS.2018.30028.

[199] D. Schell. "Development of a Macro-Pixel sensor for the Phase-2 Upgrade of the CMS experiment". PhD thesis. Karlsruhe Institute of Technology (KIT), 2019.

[200] CMS Tracker authors. "Semi-automated precision assembly of silicon detector modules". CMS Detector Note. Draft.

[201] CMS Tracker authors. "Test beam performance of an irradiated PS module for the Phase-2 CMS Outer Tracker upgrade". CMS Detector Note DN-2024/021. Draft.

[202] D. Rastorguev et al. "Timing performance of a digital SiPM prototype measured with a picosecond injection laser". In: *JINST* 20 (June 2025), p. C06035. DOI: 10.1088/1748-0221/20/06/C06035.

List of acronyms

μDTC	Micro Data, Trigger and Control
2PA	two-photon absorption
2S	strip-strip
AC	alternating current
ALICE	A Large Ion Collider Experiment
ASIC	application-specific integrated circuit
ATLAS	A Toroidal LHC ApparatuS
BPIX	pixel barrel
CBC	CMS Binary Chip
CERN	European Organization for Nuclear Research
CIC	Concentrator Integrated Circuit
CMOS	complementary metal-oxide-semiconductor
CMS	Compact Muon Solenoid
CPU	central processing unit
CSC	cathode strip chamber
CVD	continious vapor deposition
DAC	digital-to-analog converter
DC	direct current
DESY	Deutsches Elektronen-Synchrotron
DLL	delay-locked loop
DORIS	Double-Ring Storage Facility
dSiPM	digital silicon photomultiplier
DT	drift tubes
DUT	device under test
EB	ECAL barrel
ECAL	electromagnetic calorimeter
EE	ECAL endcaps
FASER	ForwArd Search ExpeRiment
FEH	front-end hybrid

FPGA	field-programmable gate array
FPIX	forward pixel detector
FZ	float zone
GBL	General Broken Lines
GEM	gas electron multiplier
GUI	graphical user interface
HB	HCAL barrel
HCAL	hadronic calorimeter
HE	HCAL endcaps
HERA	Hadron-Elektron Ringanlage
HGCAL	high-granularity calorimeter
HL-LHC	High-Luminosity Large Hadron Collider
HLT	High-Level Trigger
HO	HCAL outer
IP	interaction point
IR	infrared
ITS2	inner tracking system
L1T	Level-1 Trigger
LGAD	low-gain avalanche detector
LHC	Large Hadron Collider
LHCb	LHC beauty
LiDAR	light detection and ranging
LpGBT	Low-power Gigabit Transceiver
LS	long shutdown
m.i.p.	minimum-ionizing particle
MAPS	monolithic active pixel sensor
MaPSa	macro-pixel subassembly
MC	Monte Carlo
MOSFET	metal-oxide-semiconductor field effect transistor
MPA	Macro-pixel ASIC
MWPC	multi-wire proportional chamber
ND	neutral density
NIEL	non-ionizing energy loss
OPERA	Oscillation Project with Emulsion-tRacking Apparatus
OT	Outer Tracker
PDK	process development kit
PETRA	Positron–Electron Tandem Ring Accelerator
POH	power hybrid

PS	pixel-strip
PS-p	PS pixel sensor
PS-s	PS strip sensor
QDC	charge-to-digital converter
RF	radio frequency
RICH	ring-imaging Cherenkov
ROC	readout chip
ROH	readout hybrid
RPC	resistive plate chamber
SEH	service hybrid
SiPM	silicon photomultiplier
SM	Standard Model
SPAD	single-photon avalanche diode
SSA	Short-strip ASIC
TB2S	2S tracker barrel
TBPS	PS tracker barrel
TCAD	technology computer-aided design
TCT	Transient Current Technique
TDC	time-to-digital converter
TEC	tracker endcaps
TEDD	tracker endcap double disk
TIB	tracker inner barrel
TID	tracker inner disks
TLU	Trigger Logic Unit
TOB	tracker outer barrel
ToF	time-of-flight
ToF-PET	time-of-flight positron emission tomography
ToT	time-over-threshold
TRT	Transition Radiation Tracker
VTRx+	Versatile Transceiver plus
YETS	year-end technical shutdown

Acknowledgments

This work would not have been possible without great people who were with me during my journey. I am extremely grateful to those who kept me going regardless of any circumstances:

- My supervisor Dr. Paul Schütze, who was immensely supportive in all aspects. I appreciate that I was given a good deal of scientific freedom and that I received thoughtful feedback. This made me feel that my work is valuable and helped me conduct quality research.
- My second supervisor Prof. Dr. Katerina Lipka, whose strong leadership and vast academic experience ensured that I was always covered when it mattered most.
- My mentor Dr. Simon Spannagel, who was ready to advise me anytime on any work and personal matters with best of his expertise and understanding.
- The Tracker Upgrade team in DESY and worldwide, who created a great environment to be in: serious and professional, which made me proud of participating in such an amazing project as the CMS detector, but also welcoming and supportive. Particular thanks to Dr. Andreas Nürnberg, who guided me in many aspects related to the Tracker project.
- The Tangerine group and Young Tangerines, who are great researchers and an awesome lively community. I really liked being a part of it during my work on the lasers-related projects. The scientific and fun events set up by these people (particularly, Silicon Journal Club) brightened up the routine.
- Students of the CMS group, especially Aenne A., who was an exclusively kind and caring office mate, and Stepan Z. and Jacopo M., who provided moral and chemical support (in a form of great coffee).
- Everyone who helped me refine this manuscript. Special credit to my good friend Alexey K. from Lebedev Physical Institute for his selfless help in proofreading this and for critical discussions on modern trends in scientific writing.
- My family: my father Dmitriy, who taught me curiosity; my mother Svetlana, who taught me to not take things *too* seriously; my brother Nikita, who inspired me to be open to changes; and my partner Dr. Lisa S., who always believed in me.

Thanks to you and to many, many other colleagues, friends and strangers!