TECHNISCHE
UNIVERSITÄT
WIEN

DISSERTATION

Silicon Detector Technologies for Future Particle Collider Experiments

Ausgeführt am CERN zum Zwecke der Erlangung des akademischen Grades eines Doktors der technischen Wissenschaften

unter der Leitung von

Univ. Doz. Dipl.-Ing. Dr. Manfred Krammer
(Atominstitut)

und

Dr. Lucie Linssen
(CERN)

durch

Florian Michael Pitters, MSc.
00907462

eingereicht an der Technischen Universität Wien
Fakultät für Physik

Wien, 2019

Unterschrift Betreuer

Unterschrift Student

Abstract

Future collider experiments are typically motivated by the search for new particles and precision measurements of physical observables. The goal is to find deviations from existing theories. In practice, this requires advances towards higher luminosities at higher collision energies, finally resulting in a more difficult experimental environment. At the same time, more precise measurements are required.

One such collider is the high luminosity upgrade of the LHC (HL-LHC), that is planned to start operation in 2026. Compared to LHC, the number of simultaneous collisions (pile-up) will increase by roughly a factor four and the radiation exposure of the detectors will be an order of magnitude larger than for LHC.

Another proposed collider is the Compact Linear Collider (CLIC) that would provide electron-positron collisions at a centre-of-mass energy of up to 3 TeV. As the collision partners are fundamental particles, the initial state and energy are well known. This allows for precision measurements of Higgs and Top quark observables, including a wide variety of measurements that are particularly sensitive to physics beyond the standard model. Here, the detector is required to deliver an unprecedented precision on the measurement of jet energy, impact parameter and particle momentum.

For both colliders and their associated experiments, silicon detectors play an essential role in meeting the detection requirements. For CLIC, the desired momentum resolution together with the flavour tagging performance will require a vertex detector with a spatial resolution of around 3 µm and a time resolution better than 5 ns, all at a material budget of only 0.2% X_0 per layer. This can be achieved by using thin silicon pixel sensors of around 50 µm thickness. The design of the CLIC calorimeter is driven by the need of an unprecedented jet energy resolution. To achieve this, CLIC will employ a highly granular sampling calorimeter optimised for Particle Flow techniques. Silicon pad sensors will be used as active material in the electromagnetic calorimeter due to their compactness, operational stability, excellent signal-to-noise ratio and good segmentation properties. More recently, the CMS collaboration has decided to adopt a similar concept in the upgrade of their endcap calorimeters for the HL-LHC phase, a project commonly called HGCAL. Here, the reasoning to use silicon as active calorimeter material is however broader. It is additionally driven by the need for radiation hardness and precise time stamping for pile-up mitigation.

In this work, the role of silicon detectors in the CLIC vertex detector as well as the calorimeters for CLIC and HGCAL, is explored. The focus is put on the performance of thin detector substrates, issues of characterisation and calibration in connection with the high granularity, as well as the

possible time resolution that can be reached.

The first part explores the performance of thin silicon pixel detectors that are considered for the CLIC vertex detector. The Timepix3 ASIC is used as a test vehicle to study several aspects of thin silicon pixel sensors and to determine the specifications for the CLIC frontend electronics. A calibration method suited for a channel density of $O(10^4/cm^2)$ and addressing time-over-threshold and time-of-arrival readouts with a non-linear behaviour is investigated. Laboratory measurements and beam test results are presented and a special focus is put on the time resolution of ASIC and sensor. The second part covers the electrical characterisation of large area silicon pad sensors for fine-grained calorimeters. It describes the design and commissioning of a system for testing the voltage dependence of electrical current and capacitance of each pad. The system consists of an active multiplexer matrix and a passive probe card that handles only the contacts to the sensor. Results for a wide range of prototype silicon pad sensors are presented. The third part focusses on the sub-nanosecond timing capabilities the same sensors have when operated in a calorimeter environment. A prototype sensor is equipped with a dedicated readout and used to sample electromagnetic showers. It is shown that, on a module level, time resolutions of less than 15 ps can be achieved. Finally, the outlook on the time resolutions achievable at the system level is discussed.

Zusammenfassung

Neue Beschleunigerexperimente in der Hochenergiephysik sind meist motiviert von der Suche nach neuen Teilchen sowie Präzisionsmessungen physikalischer Observablen um existierende Theorien auf die Probe zu stellen und neue Physik jenseits des Standardmodells zu entdecken. Diese Suche stellt ständig wachsende Anforderungen an Energie und Luminosität der Kollisionsexperimente sowie an die Genauigkeit der Detektoren und Messmethoden.

Ein solcher Beschleuniger ist der High Luminosity Large Hadron Collider (HL-LHC), welcher 2026 den Betrieb aufnehmen soll. Im Vergleich zu seinem Vorgänger, dem LHC, wird sich Zahl der Kollisionen pro Umlauf vervierfachen und die Strahlenbelastung an den Detektoren verzehnfachen. Ein weiterer geplanter Beschleuniger ist der Compact Linear Collider (CLIC), ein Linearbeschleuniger für Kollisionen von Elektronen und Positronen bei einer Schwerpunktsenergie von bis zu 3 TeV. CLIC würde sehr präzise Messungen der Eigenschaften des Higgs Teilchens und des Top Quarks erlauben. Außerdem ermöglicht CLIC eine ganze Reihe von hochsensiblen Messungen für Phänomene, die von Theorien jenseits des Standardmodells vorhergesagt werden. Im zugehörigen Detektor sind dazu noch nie zuvor erreichte Auflösungen der Energie hadronischer Jets, des Stoßparameters und des Impulses erforderlich.

In den Experimenten beider Beschleuniger spielen Siliziumdetektoren eine entscheidende Rolle. Für CLIC wird ein Vertex Detektor mit 3 µm Orts- and 5 ns Zeitauflösung benötigt. Um den Anforderungen an die Auflösung von Impuls- und Stoßparameter zu genügen, darf eine Detektorlage gleichzeitig nicht dicker als $0.2\%\,X_0$ sein. Eine Möglichkeit dies zu erreichen sind sehr dünne Silizium Sensoren von nur 50 µm Dicke. Vom CLIC Kalorimeter wird eine noch nie erreichte Energieauflösung für hadronische Jets verlangt. Um diese zu erreichen, soll ein hochgranulares, für Particle Flow Algorithmen optimiertes Sampling Kalorimeter zum Einsatz kommen. Silizium Pad Sensoren werden dabei als aktive Deketorelemente benutzt, unter anderem wegen ihres ausgezeichnetem Signal zu Rausch Verhältnisses, ihrer Kompaktheit, Stabilität im Betrieb und der Möglichkeit, sie relativ einfach zu segmentieren. Die CMS Kollaboration hat sich nun dafür entschieden, dasselbe Kalorimeterkonzept auch in ihrem Upgrade für HL-LHC einzusetzen. Für CMS stehen dabei jedoch die Strahlenhärte und Zeitauflösung der Sensoren im Vordergrund.

Die vorliegende Arbeit beschäftigt sich mit der Rolle von Silizium Detektoren für CLIC und HGCAL. Der Schwerpunkt liegt dabei auf den Eigenschaften von dünnen Siliziumsubstraten, Fragen der Kalibrierung und Charakterisierung im Zusammenhang mit der hohen Granularität sowie der Zeitauflösung dieser Sensoren.

Der erste Teil untersucht die Eigenschaften von besonders dünnen Silizium Pixel Detektoren für den CLIC Vertex Detektor. Dabei wird der Timepix3 ASIC genutzt, um verschiedene Ei-

genschaften dieser Sensoren und deren Ausleseelektronik zu bestimmen. Im Zuge dieser Arbeit wurde eine Methode zur Kalibrierung von nichtlinearen Auslesesystemen mit mehreren zehntausend Kanälen pro cm^2 entwickelt und durch Labormessungen und Messungen mit Teilchenstrahlen verifiziert. Ein besonderer Fokus liegt dabei auf der Zeitauflösung. Der zweiten Teil behandelt die elektrische Charakterisierung großflächiger Silizium Pad Sensoren in hochgranularen Kalorimetern für CLIC und HGCAL. Dabei werden Strom und Kapazität über ein System vermessen, das aus einer aktiven Multiplexer Matrix und einer passiven Probekarte besteht, welche den Sensor kontaktiert. Das Design und die Kommissionierung dieses Systems werden besprochen und Resultate von Prototypsensoren für HGCAL gezeigt. Im dritten Teil wird die Zeitauflösung derselben Sensoren in einem Kalorimeter analysiert. Dazu wurde ein Prototypsensor mit schneller Elektronik ausgerüstet, um die intrinsische Auflösung des Sensors zu testen. In Messungen von elektromagnetischen Schauern mit einem einzelnen Sensor werden 15 ps erreicht. Schlussendlich werden Implikationen und weitere Herausforderungen für große Systeme analysiert.

Acknowledgements

First and foremost, I would like to thank my official and unofficial supervisors Lucie, Eva and Dominik. For the immense knowledge that I have gained from you, for the door that was never closed, for the opportunities I was given and for the trust that was placed in me. I also owe a lot to everyone in the EP-LCD group. The support and social bond I experienced in this group was exceptional, on a professional and a personal level. This is how a scientific working group should and can work. I would also like to thank Manfred for taking over the duty as University supervisor at TU Wien and his advice on progress and structure of this work.

For specific contributions to this work, I have to give my thanks (in order of appearance of their contribution throughout this thesis):

- To Jerome for his help with the Timepix3 assemblies as well as his advice and insight into wafer probing and silicon handling.
- To Xavi for his many explanations on the inner workings of the Timepix3 ASIC.
- To Adrian and Nilou, for their help with the SPIDR readout system.
- To Simon, Daniel and Morag for the Corryvreckan reconstruction tool as well as to Andreas for his help with the EUTelescope framework.
- To Fernando for his help with the probe station mechanics.
- To Andreas for the great tools of HexPlot and HexDAQ.
- To Szymon for his brilliance without which the ARRAY system would not have been possible. Also for the first version of the SPICE simulation model and his help with the capacitance measurements in general.
- To Adi for his help and advice on the timing module and the DRS4 reconstruction.
- To Thorben for providing the reconstructed samples for Skiroc2CMS.
- To Artur for the many discussions on the HGCAL system timing and his brilliant idea for the ToA calibration.

My gratitude is also to the HGCAL sensor and test beam working groups for the many discussions and advice. I am also grateful to the Austrian Doctoral Student Programme at CERN which enabled me to go to CERN in the first place and allowed me a significant amount of freedom in choosing my topic.

A special thanks goes to Simon, for the countless laughs and curses and all the (mostly) useful advice whenever it was required. I am also immensely thankful to my parents and family for their support in everything I do. My biggest thank however goes to my wife Johanna for her patience, endurance and support during the writing of this thesis and beyond.

List of publications that contain results shown in this thesis

- **2017** F. Pitters. 'The CLIC Detector Concept'. In: *CERN Proceedings* 1, pp. 37. DOI: 10.23727/CERN-Proceedings-2017-001.37.
- **2018** F. Pitters. 'The CMS High-Granularity Calorimeter for Operation at the High-Luminosity LHC'. In: *Springer Proceedings in Physics* 213. DOI: 10.1007/978-981-13-1316-5.
- **2018** The CMS Collaboration. 'The Phase-2 Upgrade of the CMS Endcap Calorimeter'. In: *CERN-LHCC-2017-023*. URL: https://cds.cern.ch/record/ 2293646.
- **2018** N. Akchurin et al. 'First beam tests of prototype silicon modules for the CMS High Granularity Endcap Calorimeter'. In: *Journal of Instrumentation* 13, P10023. DOI: 10.1088/1748-0221/13/10/P10023.
- **2018** F. Pitters. 'Time and Energy Calibration of Timepix3 Assemblies with Thin Silicon Sensors'. In: *CLICdp Note* 2018-008. URL: https://cds.cern.ch/record/2649493.
- **2019** F. Pitters et al. 'Time Resolution Studies with Timepix3 Assemblies with Thin Silicon Pixel Sensors'. In: *Journal of Instrumentation* 14, P05022. DOI: 10.1088/1748-0221/14/05/p05022.
- **2019** E. Brondolin et al. 'ARRAY: An Open Source, Modular and Integrated System for Characterisation of Large Area Silicon Pad Sensors'. In: *Nucl. Instr. Meth.* A 940, pp. 168. DOI: 10.1016/j.nima.2019.06.007.
- **2019** D. Dannheim et al. 'Detector Technologies for CLIC'. In: *CERN Yellow Reports: Monographs* 1. DOI: 10.23731/CYRM-2019-001.

Contents

Chapter 1

Introduction

In 2012, the ATLAS[1] and CMS[2] experiments at the LHC[3] have both reported the observation of a particle consistent with the Higgs boson of the standard model with a significance of more than $5\,\sigma$ [1, 2]. The Higgs boson was the last particle predicted by the standard model of particle physics – a model that unifies the description of three of the four fundamental forces of nature and has withstood more than 40 years of scientific scrutiny, including measurements with a relative precision of 10^{-15} [3, the magnetic moment of the electron]. Yet, despite its great success, there are phenomena that cannot be explained within the standard model. The most notable examples are neutrino oscillations, matter-antimatter asymmetry, dark matter and dark energy. There are also elements of the standard model that are not predicted by the model but enter as free parameters. Some of these parameters, e.g. the Higgs mass, have to be fine-tuned to meet the experimental data. This tuning is deemed unnatural and, while not a problem per se, is theoretically unsatisfying and hints to a more fundamental theory. Many physicists thus had hoped that the LHC would deliver results that deviate from the standard model predictions, e.g. by direct or indirect evidence for the existence of supersymmetric particles. On the contrary, so far the LHC results have confirmed the standard model. For example, the observation of tetra- and pentaquark states by LHCb[4] [4, 5] or the results on the branching ratio of $B_s^0 \to \mu^+\mu^-$ by CMS and LCHb [6, 7] agree with remarkable accuracy with the predictions of the standard model.

In absence of clear experimental guidance from the LHC, many new collider experiments have been proposed in the last couple of years. What they have in common is the aim for higher luminosities, more precise measurement instruments and higher collision energies than previous colliders of the same type. Typically, this results in a more difficult experimental environment due to an increase in pile-up and beam-related background. At the same time, more precise measurements are required.

One such collider is the high luminosity upgrade of the LHC (HL-LHC) [8] that is planned to start operation in 2026. By increasing the number of particles per bunch as well as adding additional

[1] A Toroidal LHC Apparatus (ATLAS).
[2] Compact Muon Solenoid (CMS).
[3] Large Hadron Collider (LHC).
[4] Large Hadron Collider beauty (LCHb).

beam focusing elements, the average number of collisions per bunch crossing, commonly called *pile-up*, will increase to about 140. For the current LHC experiments such as CMS, this results in a very challenging experimental environment as the tracks and jets from these collisions have to be separated. Additionally, the plan is to collect more than $3000\,\mathrm{fb}^{-1}$ over 10 years of operation which is more than an order of magnitude larger than the statistics LHC has collected at the time of writing. To survive this level of radiation exposure, a detector with an unprecedented radiation hardness will be required.

Another proposed collider is the Compact Linear Collider (CLIC) [9] that would provide electron-positron collisions at a centre-of-mass energy of up to 3 TeV. As the collision partners are fundamental particles, the initial state and energy of the scattering particles in CLIC are well known. This allows for precision measurements of Higgs and top quark observables and a wide variety of additional measurements that are particularly sensitive to physics beyond the standard model (BSM physics) and could confirm or discard many existing theoretical models. Here, the detector is required to deliver an unprecedented precision on the measurement of jet energy, impact parameter and particle momentum.

For both colliders and their associated experiments, silicon detectors play an essential role in meeting the detector requirements. For CLIC, the vertex detector requires a spatial resolution of 3 µm and a time resolution better than 5 ns, together with a maximum material budget of 0.2% X_0 per layer. To fulfil these requirements, silicon pixel detectors will be employed. Considering the low material budget which includes cables, support and cooling, at CLIC the maximum thickness allowed for sensor and readout together is 100 to 150 µm. This is significantly less than in existing experiments. For the electromagnetic calorimeter, CLIC will employ a highly granular sampling calorimeter made of silicon pad sensors and tungsten absorbers. This will enable the use of Particle Flow Analysis (PFA), a method to improve the jet energy resolution. The goal is to reach a resolution better than 3.5% for jet energies above above 100 GeV. For the HL-LHC phase, the CMS experiment has decided to adapt the same calorimeter concept in the upgrade of their endcap calorimeters, commonly called High Granularity Calorimeter (HGCAL). Here, the reasoning to use silicon as active calorimeter material is however broader. The use of thin silicon detectors operated at high bias voltages allows to mitigate the high radiation damage. For the rejection of pile-up, the strategy is to utilise the superb time resolution of silicon together with a high granularity.

In this work, the role of silicon detectors in the CLIC vertex detector and in the CLIC and HGCAL calorimeters is explored. A focus is put on the performance of thin detector substrates, characterisation and calibration methods compatible with the high granularity as well as the time resolution that can be reached in the different environments. Chapter 2 discusses the basic characteristics of silicon sensors as a detection material and introduces further considerations to be taken into account. Chapter 3 gives a brief reminder on collider experiments, taking CMS as an example. It also describes the proposed detector for CLIC as well as the CMS HGCAL upgrade with a focus on the role silicon detectors will play in these detectors. Chapter 4 explores the performance of thin planar silicon pixel detectors that are considered for the CLIC vertex detector. A calibration method suitable for a channel density of $O(10^4/\mathrm{cm}^2)$ that addresses the non-linear behaviour of time-over-threshold (ToT) and time-of-arrival (ToA) readouts is investigated. Here, the Timepix3 ASIC is used as a test vehicle. Chapter 5 covers the electrical characterisation of large area silicon pad sensors in prototyping and production phases for the fine-grained calorimeters of CMS HGCAL and CLIC. It describes the design and commissioning of a probe-card based testing system. Results for a wide range of prototype silicon pad sensors

are presented. Part 6 focusses on the excellent timing capabilities the same sensors have when operated in a calorimeter environment. Testbeam results with custom electronics are presented, showing that intrinsic time resolutions of less than 15 ps can be achieved on a single sensor level. Finally, implications for a full system and preliminary results with the Skiroc2CMS readout ASIC are discussed.

Some results in the presented work are shown as box plots. For these plots, the median of the distribution is represented by a solid line, the mean by a dashed line. The box boundary represents the inter-quartile range, i.e. the middle 50% of the distribution (which corresponds to $1.349\,\sigma$). The whiskers are at 2.5 times the inter-quartile range and contain $5.396\,\sigma$ or 98.95% of all data points. For more details on this form of representation the reader is referred to [10].

Chapter 2

Properties of Silicon Detectors

Silicon is widely used as a detector material in many radiation detection applications in particle, nuclear and atomic physics. The large number of freed charge carriers per incident radiation allows for an excellent energy resolutions even if the deposited energy is very low. As such, silicon detectors are used as imaging devices for various forms of spectroscopy, including visible light (e.g. in consumer cameras), alpha particles (e.g. for radon tracking) and X-rays (e.g. in medical devices). For particle tracking applications, it allows for very thin detectors that nevertheless provide an excellent signal-to-noise ratio. The possibility of micro-patterning the material with lithographic methods allows for unmatched spatial resolutions. Compared to other semiconductor materials that offer similar characteristics, silicon is the main material used in the micro-chip industry and therefore has a huge advantage in costs of and experience in the manufacturing.

This chapter discusses the basic characteristics of silicon as a detector material, its integration in a detector environment as well as the interactions that particles undergo if they traverse a silicon detector. It further discusses some considerations on timing measurements achievable with silicon detectors. Most information is considered textbook knowledge and is not referenced separately. The textbooks from Spieler [11], Leo [12], Knoll [13] as well as Kolanoski and Wermes [14] are used as main references for this chapter.

2.1 Basic Characteristics of Silicon Detectors

Silicon is a semiconductor from group IV of the periodic system with a band gap energy of $E_g = 1.12\,\text{eV}$. Electrons in the crystal lattice either occupy a localised state of individual atoms or a delocalised state of the lattice itself. The set of latter states is called conduction band. An electron occupying such a state is no longer bound to a fixed lattice location and can move freely through the crystal. The probability that a certain electronic state of energy E at a certain temperature T is occupied by an electron is given by the Fermi-Dirac distribution

$$f_e(E) = \frac{1}{e^{(E-E_F)/k_B T} + 1}, \tag{2.1}$$

where k_B is the Boltzmann constant and E_F the Fermi energy, the point at which the probability of a state being occupied is 50%.

An electron (e) can be promoted to the conduction band either via thermal excitation or external energy depositions. If this happens, a residual vacancy is left in the crystal lattice that can be described as a quasi-particle, commonly called hole (h). The average energy needed to create one electron-hole pair is about $W \sim 3.65\,\text{eV}$, though the exact value depends on the specific nature of the energy deposition [15]. By applying an electric field, electrons and holes can be spatially separated. This principle is the foundation of silicon (or any other semiconductor) as detector material.

In pure silicon crystals, the intrinsic charge carrier density is about $10^{10}\,\text{cm}^{-3}$ at room temperature. This is significantly larger than the number of charge carriers created by e.g. ionisation from a charged particle crossing the same volume. To reduce the intrinsic charge carrier density and to make silicon a viable detector material, so-called dopants are introduced.

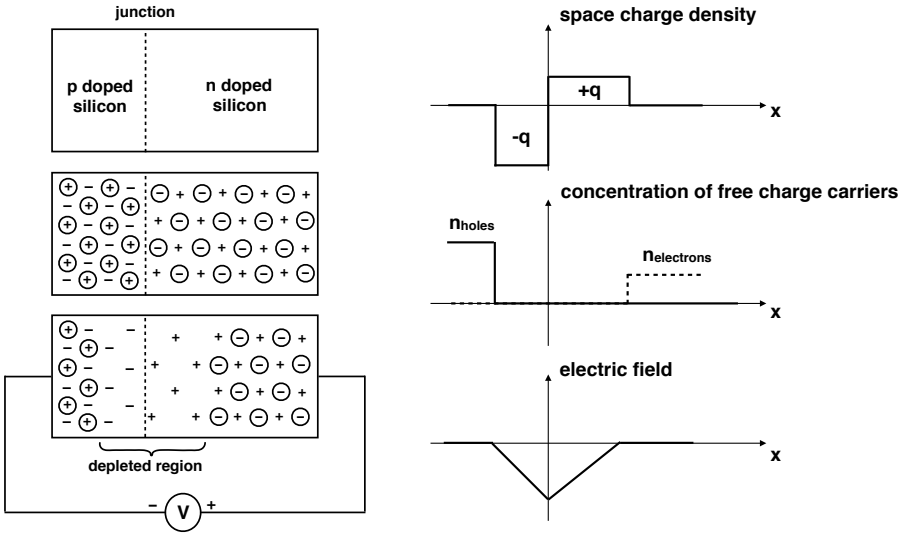
2.1.1 From a Crystal to a PN Junction

Dopants are foreign atoms from either group III (e.g. boron) or group V (e.g. phosphorous) that are added to the silicon crystal lattice. Due to the different number of valence electrons of the dopants, additional energy levels are introduced to the band gap. In the case of a group V dopant, the additional electron can not be accommodated anymore in the regular covalent bond structure. It introduces a loosely bound state in the band gap, typically just below the edge of the conduction band. The new state is referred to as a donor state as it can be easily promoted to the conduction band. Effectively, this doping method, called n-type doping, creates an abundance of quasi-free electrons. The electrons are then called majority charge carriers and holes are called minority charge carriers. In the case of a group III dopant, one electron is missing from the regular covalent bond structure and a state in the band gap just above the edge of the valence band is introduced. Here, one speaks of p-type doping and acceptor states. The majority charge carriers in such a material are holes.

If a p- and an n-type doped material come into contact, they form a so-called pn-junction. The acceptor states in the p-type part are quickly filled by the electrons that previously occupied the donor states in the n-type part. This flow of electrons results in a negatively charged p-type region and a positively charged n-type region, creating a potential difference V_j across the junction (see Fig. 2.1). In the equilibrium state, the contact region is then free of intrinsic charge carrier.

2.1.2 From a PN Junction to a Detector

The size of the depleted volume is determined by the doping concentration and limited to at most a few micrometers. To expand the depleted region, an external voltage V_{bias} is applied to contact electrodes in reverse direction: the positive pole of the voltage source is connected to the negative pole of the junction. The depleted region then expands proportional to the square root of the sum of the applied bias voltage and internal junction voltage. For the following discussion, an effective voltage $V_{eff} = V_{bias} + V_j$ is introduced. As V_j is typically well below 1 V, in most cases $V_{eff} \approx V_{bias}$ is a valid simplification.



(a) (b)

Figure 2.1: (a) Adjoining regions of p- and n-type doped silicon form a pn-junction. Mobile charge carriers are circled compared to the static charges of the nuclei. Once the pn-junction is established, the mobile charge carriers recombine in the contact region. (b) The resulting space charge density, free charge carriers concentration and electric field. Adapted from [11, Fig. 1.14].

Most detector structures use a highly asymmetric junction that consists of a lightly doped silicon substrate of one doping type (e.g. p-type) that is a few tens to hundred micrometer thick and is implanted with a shallow, strongly doped layer of the other type (e.g. n-type). The thick substrate is conventionally denoted bulk and the full sensor is either denoted as the bulk material (e.g. p-type) or as implant-on-bulk (e.g. n-on-p). For such structures, the extension w of the depletion region with reversely applied bias voltage is given by

$$w = \sqrt{2\epsilon_0\epsilon_r\mu\rho V_{\text{eff}}}, \tag{2.2}$$

with the relative permittivity ϵ_r and vacuum permittivity ϵ_0. The specific resistivity ρ in units of [Ωcm] is given by

$$\rho = \frac{1}{\mu e N}. \tag{2.3}$$

Here, N is the effective doping concentration, μ the mobility of the majority charge carriers of the bulk material and e is the elementary charge. If the applied voltage is increased, the depleted region extends until the depleted width equals the total bulk thickness t. This state is then called full depletion and the corresponding voltage is denoted full depletion voltage V_{fd}. If the area A of the implant is significantly larger than the thickness t, the depleted region is in a good approximation a parallel plate capacitor with

$$C = \epsilon_0\epsilon_r\frac{A}{w} = A\sqrt{\frac{\epsilon_0\epsilon_r}{2\mu\rho V_{\text{eff}}}}. \tag{2.4}$$

With increasing voltage, the capacitance decreases until V_{fd} is reached.

The electric field $E(x)$ in the sensor shows a linear profile with the maximum field at the junction. Before full depletion, it behaves like

$$E(x) = \frac{2V_{\text{eff}}}{w}\left(1 - \frac{x}{w}\right). \tag{2.5}$$

If the applied voltage is increased beyond V_{fd}, the electrical field increases with an additional constant contribution

$$E(x) = \frac{2V_{\text{fd}}}{t}\left(1 - \frac{x}{t}\right) + \frac{2(V_{\text{eff}} - V_{\text{fd}})}{t}. \tag{2.6}$$

When free charge carriers are created in the depleted volume, they are accelerated by the electric field and drift towards the corresponding electrode. Their movement is determined by a superposition of two forces: thermal diffusion and drift in the electric field. In combination, the charge cloud will drift towards the electrode while expanding laterally with a Gaussian profile. The drift velocity v as a function of the electric field E is given by

$$\vec{v} = \mu\vec{E} \tag{2.7}$$

with the charge carrier mobility

$$\mu = \frac{e}{k_B T}D. \tag{2.8}$$

Here, D is the diffusion coefficient, that is a material property. It depends on the temperature, doping concentration and the electric field. At low electric fields, the mobility is approximately constant and the drift velocity rises linearly with the electric field. At 300 K, μ is 1350 cm^2/Vs for electrons and 450 cm^2/Vs for holes [11, pg. 69]. This results in a lower resistivity for p-type material and higher drift velocities for electrons. For higher fields, the increasing number of

collisions limits the drift velocity. At room temperature, the drift velocity of electrons saturates at fields of about $5{\cdot}10^4\,\mathrm{V/cm}$ to a value of about $10^7\,\mathrm{cm/s}$ which is roughly the same as the thermal velocity that governs diffusion [11]. At fields above $10^5\,\mathrm{V/cm}$, impact ionisation by the drift electrons becomes relevant and results in multiplication of the signal. At sufficiently high fields, a sustained avalanche of charge carriers occurs that is commonly called breakdown. While some detector types rely on this mode of operation, it is an unwanted phenomenon in all structures discussed in this work.

If a pair of free charge carriers is created and drifts toward the corresponding electrodes, the moving charges will induce a current given by Ramo's theorem

$$i = q \cdot \vec{v} \cdot \vec{E}_W. \tag{2.9}$$

The variable $\vec{E}_\mathrm{W}$ is called weighting field that is determined by removing all charges q and applying a unit potential to the collection electrode while grounding to all others. For a non-segmented planar geometry, it behaves as $E_\mathrm{W} \propto 1/x$. The total current flow is the sum of the induced current on both electrodes and consists of a faster component from the electron current and a slower component from the hole current. Typical charge collection times are of the order of a few to a few tens of ns.

In the reversed biased sensor, thermal excitation results in a movement of charge carriers even in the absence of external energy deposition. This so-called leakage current contributes to the noise of the detector and has several components, generally separated into volume and surface currents. In the depleted volume, thermal excitations cause the so-called volume generation current. It depends on the band gap and the impurities of the crystal and grows approximately with $\sqrt{V_\mathrm{bias}}$. In the non-depleted volume, free charge carriers have a finite probability to diffuse into the depleted volume before they recombine. This contribution is called diffusion current. Other sources of current produced in the detector include e.g. currents due to resistive paths between the frontside structures on the sensor (see following section). The volume generation current scales with temperature [11] as

$$\frac{I(T_2)}{I(T_1)} = \left(\frac{T_2}{T_1}\right)^2 \exp\left[-\frac{E_\mathrm{eff}}{2k_B}\left(\frac{T_1 - T_2}{T_1 T_2}\right)\right], \tag{2.10}$$

where E_eff is about $1.21\,\mathrm{eV}$ [16]. The surface current varies with temperature as well but also with humidity.

2.1.3 Basic Detector Structures

Fig. 2.2 shows the basic structure of a segmented planar p-on-n silicon sensor. The substrate is usually of high resistivity (several kΩcm) and cut from an ingot grown via a float zone process [17]. It is uniformly doped with a relatively low concentration. At the backside, an additional highly doped implant is used to create an ohmic contact to a metal, typically aluminium or gold. The additional implant is necessary to avoid a Schottky barrier [18]. On the frontside, the implant is segmented. This allows to collect spatially resolved information about the energy deposition of the incident radiation. Typically, the segmentation is done in one dimension via strips or in two dimensions via pads. If the pads are very small, e.g. less than $1\,\mathrm{mm}^2$ in area, they are referred to as pixels. For a segmented detector, the weighting field E_W becomes non-symmetric in the transversal direction and charge can also be induced on the neighbouring pixels. Between

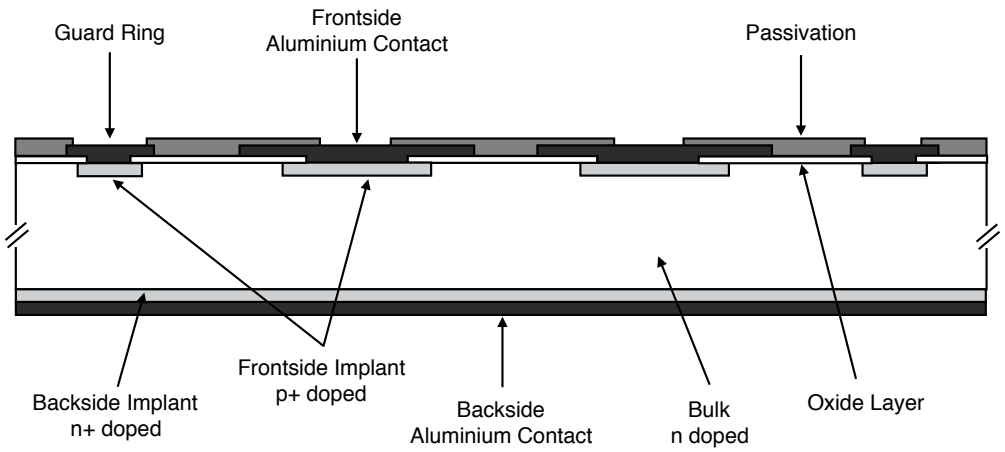


Figure 2.2: A basic segmented planar p-on-n silicon sensor. Reproduced from [11, Fig. 2.22].

the implants, a layer of silicon oxide (SiO_2) truncates the edge of the silicon lattice and provides electrical isolation. On top of the implant, a metal layer is deposited. Typically, a small overhang of the metal compared to the implant is used to move the peak electrical field into the SiO_2, which has a larger breakdown field than pure silicon. In order to protect and isolate the frontside it is chemically passivated. Openings in the passivation layer allow to establish an electrical contact to the metal. The edge of the silicon sensor is diced to remove excess material. The dicing process inevitably introduces a large amount of defects at the edge of the silicon lattice. To absorb the resulting leakage current, so-called guard ring structures are used. These extra implants surround the sensor and collect all current from the edge region. This avoids an increase in noise for the inner pads. However, it also creates an insensitive area.

2.1.4 Effects of Irradiation

The effects of radiation are generally separated in two groups: bulk and surface damages.

Bulk damage is mainly caused by heavy particles, such as protons or neutrons, in nuclear collision processes that can dislocate lattice atoms and create additional energy states in the band gap. Depending on the nature of the defect, this has several consequences. Drifting charge carriers can get *trapped* in the meta-stable states and are lost for the signal acquisition. This reduces the charge collection efficiency. The additional levels also change the effective doping concentration and increase the leakage current. A common working hypothesis is the so-called NIEL hypothesis which states that the radiation damage is linearly proportional to the non-ionising energy loss which in turn is proportional to the displacement energy of the lattice. While this assumption ignores several types of nuclear reactions, it is a useful approximation to compare radiation effects of different particles. The irradiation dose is then given in fluencies of 1 MeV neutron equivalent per cm^2, denoted as $[n_{eq}/cm^2]$.

Surface damage is mostly caused by electromagnetically interacting particles in electronic ionisation processes in the silicon oxide. In SiO_2, the mobility of electrons is about an order of magnitude larger than for ions. This causes a net increase of positive charges in the oxide which,

at the interface to the silicon bulk, results in a net increase of negative charges. In n-on-p sensors, this reduces the inter-pad resistivity. To counteract this effect, additional p levels are introduced between the n-doped implants. These are called p-spray or p-stops.

Reducing the negative effects of irradiation is often the main concern for detectors operating at the HL-LHC. In Sec. 3.3.3 and 5.1 further measures taken by HGCAL to limit the radiation effects are discussed.

2.2 Readout and Detector Integration

To build the silicon sensor into a full detector, additional elements are needed. The sensor output is interfaced to the frontend electronics which are usually integrated in an Application Specific Integrated Circuit (ASIC). In the frontend, the signal is amplified, shaped and digitised and only a bitstream is sent for further processing. This bitstream is handled by the backend electronics which also provides power, control and other infrastructure to the frontend and sensor. Typically, several sensor and frontend electronics are built into modules that provide mechanical support and share backend electronics as well as additional infrastructure e.g. for cooling.

2.2.1 Signal Readout Techniques

Fig. 2.3 depicts a typical readout chain consisting of the sensor followed by an amplification stage with a feedback loop, a shaping stage and a digitising stage. In general, if sensor and frontend electronics are physically separated, it is called a hybrid detector. If they are integrated on the same silicon substrate, it is called a monolithic detector.

The sensor itself can be approximated via a capacitor and a current source in parallel. The current induced on the collection electrode is typically too small to be measured directly. Hence, amplification is required. A common amplification structure is the charge sensitive amplifier (CSA) that allows to amplify the charge deposited in the sensor to a proportional voltage signal and control the gain of the amplification via the value of the feedback capacitance. Such a structure is shown in the second panel of Fig. 2.3. As the current develops in the sensor, a feedback capacitor is charged and a voltage develops across its poles. To discharge the feedback capacitance and return the voltage to the baseline, a second current path is introduced to the feedback loop. Typically, this is realised via a resistive element (as in Fig. 2.3) or a constant current source (e.g. in Timepix3, see Chapter 4).

Often, additional shaping elements are used to differentiate and integrate the signal. Each stage introduces a characteristic frequency dependence represented by the band width f_{BW} and the characteristic time constant τ, related by

$$\tau = \frac{1}{2\pi f_{\mathrm{BW}}}, \tag{2.11}$$

that defines the output signal shape in time. For many applications, the most relevant quantity is the rise time, i.e. the time the signal takes to rise from 10% to 90% of its maximal amplitude. For most input signals and shaping elements, the rise time t_{rise}, the characteristic time τ and

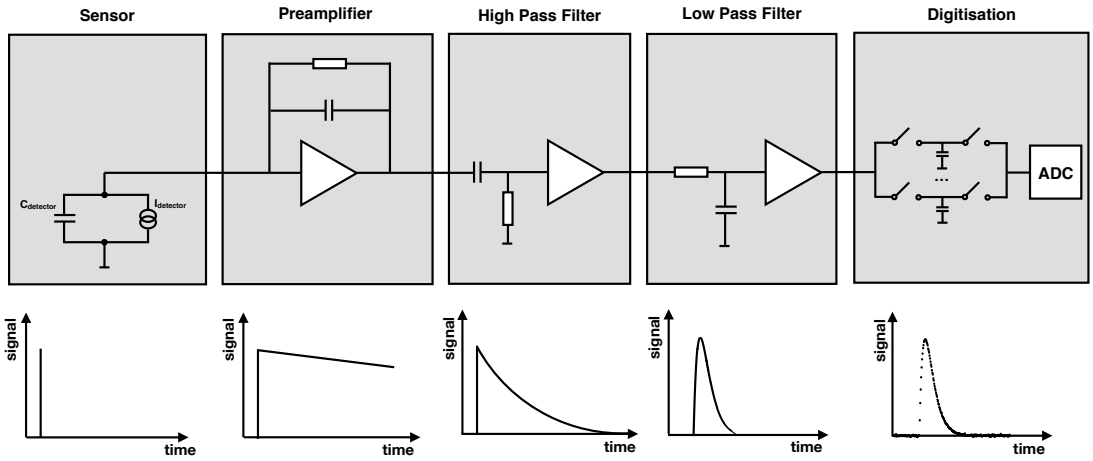


Figure 2.3: A typical frontend readout chain consisting of the sensor followed by an amplification stage with a feedback loop (here with resistive discharge), a shaping stage (here a 2-stage CR-RC element) and a digitising stage (here a sampling ADC). Adapted from [11, Fig. 1.5].

the band width f_{BW} are approximately related via

$$t_{\mathrm{rise}} \approx 2.2\tau = \frac{0.35}{f_{\mathrm{BW}}}. \tag{2.12}$$

Each element of the readout chain introduces a time constant. In a good approximation they are additive,

$$t_{\mathrm{rise}} = \sqrt{\sum_i t_{\mathrm{rise,i}}^2}. \tag{2.13}$$

where i represents the different sources. The choice of the characteristic time constant used for a system depends on the signal rate, the detector capacitance and the measurement requirements for a particular application. In most systems, an optimal value for the time constant can be found that minimises the noise at a given input capacitance.

After shaping, the signal is digitised. Typically, arrival time and pulse height of the signal are the quantities of interest. In the simplest case, a comparator is used to determine if a signal is above a certain threshold. It outputs a single bit, 0 or 1, for the pulse height. The arrival time can be obtained by comparison of the time of the threshold crossing with an external clock. This is a common choice in high rate tracking applications at LHC in which one only requires information on if, where and in which bunch crossing a particle has traversed the detector. An extension of this method is to count for how long the signal exceeds the threshold to obtain a more precise amplitude information. Such a system is used in Chapter 4. Another common implementation is digitisation of the full waveform via sampling with a frequency f_{sampling} well above the band width f_{BW}. During data acquisition, each voltage sample is stored in a separate capacitor which is sequentially overwritten until the acquisition stops. At the end of the acquisition, a common analog-to-digital converter (ADC) measures the voltage of each capacitor and stores the discretely digitised waveform. Such an implementation is sketched in the last panel of Fig. 2.3 and is used for example in Chapter 6. Compared to the previous examples, such a readout requires many more bits per event along with a significantly higher power budget and/or lower rates.

After the signal is digitised in the pixel, it is sent to the periphery of the ASIC that holds functionality that is common to all pixels and provides the interface to the backend electronics. For more information on the individual parts of common frontend structures the reader is referred to [11, Chapter 2.9 and 4] and [19, Chapter 3].

2.2.2 Bonding and Module Integration

To interconnect the sensor output with the ASIC input in hybrid detectors, different techniques are used, depending on the channel density. For the low density connections required in silicon pad sensors, the most common technique is wire bonding. In this technique, a machine places a wire with high precision and melts the tip shortly before contacting to the metal surface of the contact pad. The same technique is also used to interface the ASIC with the rest of the module. For the high density connections required in silicon pixel detectors, bump bonding is used. Here, tiny structures of a solderable material are deposited on each pixel of either sensor or ASIC. The two sides are brought into contact in a so-called flip-chip process. Depending on the material used, an optimised cycle of heat and pressure is needed to ensure a solid electrical connection between the sensor and the ASIC pixels. Often, a special surface treatment called Under Bump Metallisation (UBM) is required before deposition of the bump structures. An overview of different bump bonding techniques and the achievable channel density can be found in [20].

2.3 Particle Interaction in Silicon

Particles of different type and energy are subject to different interactions in different materials. In the following, only the dominant interactions of particles and energy ranges relevant to this work are briefly discussed. For a more complete discussion, the reader is referred to [12, chapter 2] and [13, chapter 2] which serve as reference for most of this chapter. For the section on hadronic interactions, [21, chapter 2.3] is used as reference.

2.3.1 Interactions of Minimum Ionising Particles

For heavy charged particles with velocity $\beta = v/c$ and Lorentz factor γ, the energy loss is dominated by a large number of elastic Coulomb interactions with electrons of the target material. Above $\beta\gamma > 0.1$, the mean energy loss is described by the Bethe Bloch curve. Minimum ionising particles (MIPs) are a class of particles whose energy loss is close to the minimum of that curve. Prominent examples are high-energy muons, pions or protons. The maximum energy transfer per elastic interaction with a target electron is kinematically limited and small compared to the particle's incident energy. The energy transfer is subject to statistical fluctuations, with small transfers more probable than larger ones. When traversing a thin layer of material, this results in an asymmetric distribution with a tail towards larger energy depositions. This distribution is approximated by the Landau model and is defined by its most probable value and its width. For very thin layers and highly relativistic particles, the model however deviates from experimental results. Bichsel [22] provides a phenomenological parametrisation of the most probable energy loss Δp (in [keV]) in a thin layer of silicon with thickness t (in [µm]). For particles with charge

Table 2.1: Most probable energy loss of minimum ionising particles in thin silicon substrates. Energy values are taken from [22]. Charge values are calculated assuming 3.65 eV per electron.

Thickness	Most Probable Value		
[µm]	[keV]	[ke-]	[e-/µm]
50	11.94	3.27	65.4
100	26.34	7.22	72.2
120	32.16	8.81	73.4
150	40.75	11.16	74.4
200	55.27	15.14	75.7
300	84.89	23.26	77.5

$\pm 1e$ and $\beta\gamma > 500$ this yields

$$\Delta p = t \cdot (100.6 + 35.35 \ln t) \qquad 13 < t < 110 \tag{2.14}$$

$$\Delta p = t \cdot (190 + 16.3 \ln t) \qquad 110 < t < 3000. \tag{2.15}$$

The uncertainty on this parametrisation is quoted as less than ±1%. Tab. 2.1 lists the most probable energy loss for all substrate thicknesses used in this work.

2.3.2 Interactions of Low Energy Photons

For photons in the energy range of 1 keV to 100 keV, the dominant interaction mechanism is photo absorption. The cross section for photons in that energy range scales with the atomic number Z approximately as

$$\sigma_{\text{photo}} \approx \frac{Z^5}{E_\gamma}. \tag{2.16}$$

The number of primary photons decreases exponentially with distance x,

$$I(x) = I_0 \exp\Big(-\frac{x}{\lambda_{\text{photo}}}\Big), \tag{2.17}$$

where λ_{photo} is called the attenuation length that depends on energy and material.

2.3.3 Interactions of High-Energy Photons, Electrons and Positrons

High-energy photons, electrons and positrons can initiate a so-called *electromagnetic shower* in the target material, meaning that a single primary particle creates a cascade of descendants. Two processes play a dominant role. High-energy electrons and positrons lose energy through bremsstrahlung, which is the emission of a photon when the particle is deflected in an electromagnetic field. Photons, in turn, can produce an electron-positron pair in the presence of an electromagnetic field. The combination of both effects results in a cascade, which continues until the energy of the pair-produced electrons and positrons drop below a critical energy E_{crit}, at

which impact ionisation becomes the dominant channel for energy loss. The value of E_{crit} is material dependent. For solids, a good approximation [21] is

$$E_{\text{crit}} \approx \frac{610\,\text{MeV}}{Z + 1.24}. \tag{2.18}$$

Here, Z is the proton number of the material. The two most important parameters to characterise the electromagnetic shower are the radiation length X_0 and the Molière radius ρ_M. The Molière radius describes the lateral size of the shower. The lateral expansion is Gaussian and $\pm 1\sigma$ of the energy deposition is contained within

$$\rho_M = X_0 \frac{21\,\text{MeV}}{E_{\text{crit}}\,[\text{MeV}]}. \tag{2.19}$$

The radiation length defines the longitudinal expansion of the shower. It is defined as the energy at which an incident electron or positron has lost all but $1/e$ of its initial energy.

$$E(x) = E_0 \exp\Big(-\frac{x}{X_0}\Big). \tag{2.20}$$

A list of values for X_0 and ρ_M for all materials relevant to this work is given in Tab. 2.2.

2.3.4 Interactions of High-Energy Hadrons

Charged high-energy hadrons like pions or protons behave as minimum ionising particles when interacting with electrons of the target medium but they can also undergo nuclear reactions. The most relevant type of reaction for high-energy hadrons are inelastic nuclear reactions that result in excited nuclei in the target material. These can then de-excite via various channels but most importantly via evaporation or production of secondary particles. As the secondary particles are usually also hadrons with sufficient energy to produce tertiary particles, a *hadronic shower* develops. Within the hadronic cascade, the production of particles with a predominantly electromagnetic decay channel, most importantly $\pi^0 \to \gamma\gamma$, create electromagnetic cascades. Once the maximal transferable energy of the shower particles drops below the pion production threshold - the lightest hadron - the hadronic cascade stops.

The longitudinal development of hadronic showers is quantified by the nuclear interaction length λ_{int} that is defined via the probability of a particle undergoing a nuclear interaction after traversing a distance x of a given material

$$P(x) = \exp\Big(-\frac{x}{\lambda_{\text{int}}}\Big). \tag{2.21}$$

In Tab. 2.2, values for λ_{int} are listed for several materials relevant to this work. It should be noted that the values for λ_{int} are significantly larger than for X_0. In lateral direction, roughly 95% of all energy is deposited within one λ_{int}. Due to neutron scattering events however, energy deposition occurs also far from the shower core. Hadronic showers are therefore significantly larger in lateral as well as longitudinal expansion compared to electromagnetic showers.

Hadronic interactions can deposit a significant fraction of the energy in ways that are 'invisible' to the detector. The most important example is energy that is deposited as nuclear binding energy. As a consequence, the amount of visible energy in the detector from an incident 100 GeV pion is

Table 2.2: Radiation length X_0, Molière radius ρ_M and nuclear interaction length λ_{int} for the materials relevant to this work. Values taken from [21].

Material	X_0 [cm]	ρ_M [cm]	λ_{INT} [cm]
Si	9.36	4.8	45.5
Fe	1.76	1.69	16.8
Pb	0.56	1.60	17.0
W	0.35	0.93	9.6

smaller than from a 100 GeV electron. This feature is called non-compensation. In a homogenous material, the difference can be significant and varies with the target material but also with the incident particle type and energy. In combination with the mixture of hadronic and electromagnetic cascades and the relatively small hadron multiplicities, the amount of visible energy to the detector fluctuates significantly more from event to event than for pure electromagnetic showers.

2.4 Considerations on the Time Resolution in Silicon Detectors

This section discusses the different uncertainty contributions on the time measurement with silicon diodes. A similar discussion for the case of MIP timing with LGADs can be found in [23]. The current pulse from the sensor is a mixture of signal and noise and is typically shaped, amplified and digitised. All these steps introduce an uncertainty on the time measurement.

2.4.1 Contributions from Jitter

The timestamp of a pulse is typically taken from the rising edge of the pulse. As the pulse is a mixture of signal and noise, the uncertainty on the measurement point depends on the slew rate dV/dt and the noise level N of the system,

$$\sigma_{\text{jitter}} = \frac{N}{dV/dt}. \tag{2.22}$$

Assuming that the rise time is constant, this can be expressed via the signal-to-noise ratio S/N and the total rise time t_{s} of the signal,

$$\sigma_{\text{jitter}} \approx \frac{t_{\text{s}}}{S/N}. \tag{2.23}$$

It should be mentioned here that the noise and rise time are not independent. Often, the optimum time resolution is found if the rise time of the amplifier matches the intrinsic rise time of the signal.

2.4.2 Contributions from Digitisation

The digitisation process is typically done by a TDC or waveform sampling. In the case of a TDC with binning b, the resolution is given by the quantisation noise,

$$\sigma_{\text{binning}} = \frac{b}{\sqrt{12}}. \tag{2.24}$$

Additional uncertainties may be introduced if e.g. the bins are not of equal size. For a waveform digitiser, the limited sampling rate introduces an uncertainty that, according to [24], can be approximated via the number of sampling points on the rising edge n_{samples} and the S/N of the pulse,

$$\sigma_{\text{sampling}} \approx \frac{1}{\sqrt{n_{\text{samples}}}} \frac{t_{\text{s}}}{S/N}. \tag{2.25}$$

2.4.3 Contributions from Signal Formation

Besides the readout electronics, the signal development itself is not instantaneous and presents a source of uncertainty. The energy deposition is often non-uniform throughout the sensor. This is particularly relevant to signals created by minimum ionising particles that follow a Landau shape in their energy loss. For a signal created from many quasi simultaneous particles within a shower, it becomes negligible as the energy deposition approaches a Gaussian distribution. Once the charges are created, they couple to the readout electrodes. Here, non-uniformities in the weighting field and drift velocity affect the coupling strength and contribute to the time resolution. These contributions can be reduced by using large cell sizes that result in more uniform weighting fields and the use of high electric fields that saturate the drift velocity for most of the sensor thickness. Charge sharing or drift time effects from inclined tracks might also influence the time resolution.

2.4.4 Contributions from Other Sources

The above discussion includes a set of common contributions to the time resolution. Other sources of uncertainty include e.g. imperfect calibration of the clock delay, channel-to-channel variations in the clock periods and mis-synchronisation between channels. One specific case of imperfect calibration that can contribute to the time resolution is timewalk, a phenomenon discussed in more detail in Sec. 4.1.

Chapter 3

Silicon Detectors in the CLIC Detector and CMS HGCAL

The first use of silicon sensors in high-energy particle physics is generally attributed to the NA11 experiment at the SPS around 1980 [25]. For cross-section measurements of charmed particles, a spatial resolution of 10 µm was required to identify secondary vertices from the decay. This was achieved with planar silicon micro-strip sensors with a pitch of 20 µm. Since then, silicon sensors have a longstanding tradition as vertex and later tracking detectors in high-energy particle physics experiments.

In this chapter, the role of silicon sensors in present and future collider experiments is discussed. The detector layouts of the CLIC detector and CMS are presented and essential differences are highlighted.

Parts of this chapter have also been published by the author in [26] and [27] and are quoted verbatim.

3.1 Current Detectors in High-Energy Physics

Modern particle detectors in high-energy physics are usually based on a sequence of detectors surrounding the central collision point. Exploiting the different interaction mechanisms of different particles, each layer is optimised for a specific task. This typically includes a multi-stage tracking detector to reconstruct trajectories of charged particles in a magnetic field, a calorimeter for energy measurements, muon detectors and additional particle identification detectors. The underlying goal is to detect, identify and measure all particles emerging from the collision in order to reconstruct the primary scattering process. In the following, this is illustrated based on the example of the CMS detector.

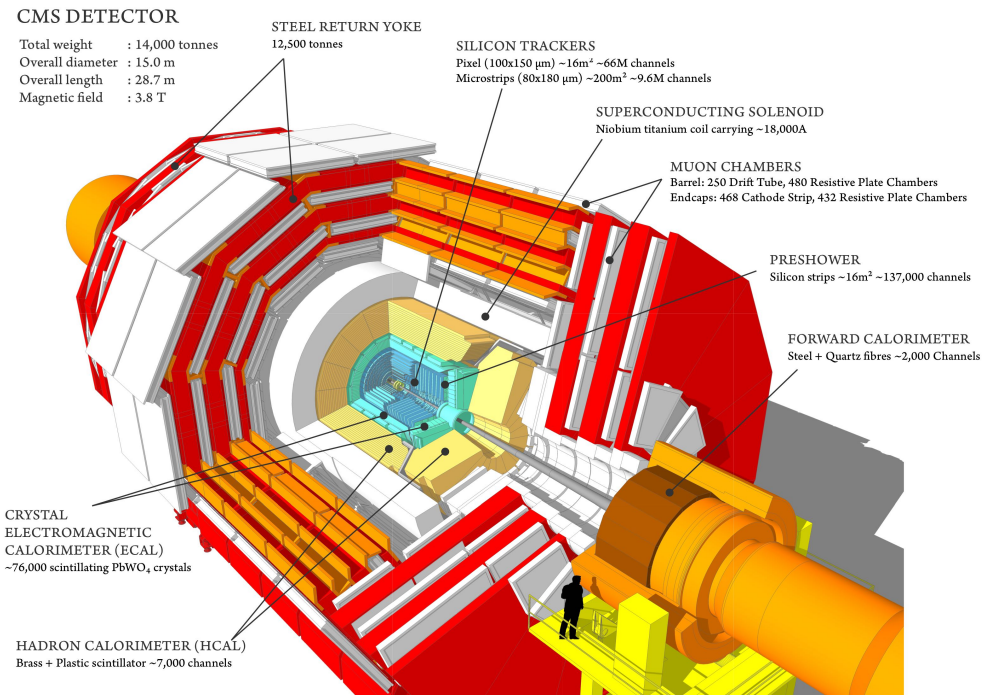


Figure 3.1: Overview of the CMS detector. Taken from [29].

3.1.1 The CMS Experiment

To date, in the LHC, bunches of two proton beams of 6.5 TeV cross every 25 ns. In each bunch crossing, approximately 40 quasi-simultaneous inelastic proton-proton collisions take place. Per collision, several final state particles and consequent decay products emerge that traverse the detector volume. In the event reconstruction they have to be separated to gain insight into the underlying physics. While CMS was conceptually conceived as a multi-purpose detector with a broad physics programme, one of the main motivations was the search for the Higgs boson. The most promising measurement channels were thought to be $H \to 4l$ and $H \to \gamma\gamma$ [28]. CMS is therefore optimised for an excellent energy and momentum resolution for electromagnetic particles as well as for good muon identification. An overview of the CMS detector is shown in Fig. 3.1. In the following, the main measurements and sub-detectors are briefly discussed.

The Momentum Measurement

The momentum measurement is performed using a magnetic spectrometer that consists of a multi-layer tracking detector and a magnetic field. Charged particles emerging from the collision are bent by the magnetic field. From the trajectory, the momentum can be reconstructed. For a solenoidal magnetic field with field lines in the beam direction, the resolution of the transverse momentum component $\mathrm{p_T}$ can be approximately parametrised as

$$\frac{\sigma_{\mathrm{p}_T}}{p_T^2} = \sqrt{a^2_{\mathrm{p}_T} + b^2_{\mathrm{p}_T}/(p^2 \sin^3\theta)}. \tag{3.1}$$

The parameter a_{p_T} represents the contribution from the curvature measurement while parameter b_{p_T} represents contributions from multiple coulomb scattering, with [14, adapted from Eqs. 9.50 and 9.56]

$$a_{\mathrm{p}_T} \propto \frac{\sigma_{\text{spatial}}}{BL^2}\sqrt{\frac{1}{N+4}} \qquad \text{and} \qquad b_{\mathrm{p}_T} \propto \frac{1}{BL\beta}\sqrt{\frac{L}{X_0}}. \tag{3.2}$$

where L is the tracker radius, B the magnetic field, N the number of layers (or hits) and X_0 the radiation length of the detector material. It is therefore advantageous to have a light tracker with many layers and a large radius in a strong magnetic field. In CMS, the magnetic field is generated by a 3.8 T superconducting solenoid that is located outside of the calorimeters. A return yoke made of steel confines the magnetic field to the detector volume and adds mechanical stability. The tracker is implemented in two parts: the inner tracker and the muon tracker. The inner tracker consist of multiple layers of silicon pixel and silicon strip detectors. The total tracker radius is 1.1 m. Together, they compose nearly 75 million readout channels. As this detector is located close to the interaction point, it is also exposed to large radiation doses. The muon tracker is located outside the solenoid, interleaved in the return yoke. It consists of several layers of resistive plate chambers (RPCs) and drift tubes (DT) in the barrel and cathode strip chambers (CSC) in the endcaps. These allow to identify muons and improve the measurement of their momentum.

The Impact Parameter Measurement

The tracker also has the task to reconstruct the spatial origin of a track, specifically whether it correlates with the primary collision point or a secondary vertex from a short-lived decay product. The impact parameter d_0 is used as a measure to determine secondary vertices. The resolution in $r\phi$ can be approximately parametrised as

$$\sigma_{\mathrm{d}_0} = \sqrt{a_{\mathrm{d}_0}^2 + b_{\mathrm{d}_0}^2 [\mathrm{GeV/c}]^2/(p^2 \sin^3\theta)}. \tag{3.3}$$

Similar to the momentum measurement, the parameter a_{d_0} represents the contribution from the measurement points while parameter b_{d_0} represents contributions from multiple coulomb scattering, with [14, adapted from Eqs. 9.67 and 9.71]

$$a_{\mathrm{d}_0} \propto \frac{\sigma_{\text{spatial}}}{\sqrt{N}} \qquad \text{and} \qquad b_{\mathrm{d}_0} \propto r_B\sqrt{\frac{x}{X_0}}. \tag{3.4}$$

Here, x is the thickness of the traversed material and r_B the radius of the beam pipe. While this is a useful parametrisation, it is not exact and depends on the geometry that is used. For CLIC, the angular dependence has been shown to deviate from this parametrisation, especially at high momenta [30, Fig. 27]. Compared to the momentum measurement, not all detector layers have equal weight on the uncertainty and the resolution is mostly determined by the innermost tracking layers. This part of the tracker is therefore often called vertex detector.

The Energy Measurement

The energy is typically measured in a calorimeter which fully absorbs the particles and measures the deposited energy. In a classic sampling calorimeter, the energy resolution can be parametrised

as

$$\frac{\sigma_E}{E} = \sqrt{\left(\frac{a_\mathrm{E}}{E}\right)^2 + \left(\frac{b_\mathrm{E}}{\sqrt{E}}\right)^2 + c_\mathrm{E}^2}. \tag{3.5}$$

Here, parameter a_E represents contributions from noise, b_E represents contributions from stochastic variations and c_E represents constant contributions from e.g. shower leakage and calibration errors.

In almost all high-energy physics detectors including CMS, the calorimeter is split in two separate parts, one for particles that interact only electromagnetically (e.g. photons and electrons) and one for particles that also interact hadronically (e.g. protons and pions). In CMS, the electromagnetic calorimeter (ECAL) consists of about 70000 $PbWO_2$ scintillating crystals read out by silicon avalanche photodiodes (APDs) in the barrel and vacuum phototriodes (VPTs) in the endcaps. Each crystal is 2.2 by 2.2 cm^2 at the front face and 23 cm deep. The hadronic calorimeter (HCAL) is a sampling calorimeter made of brass and plastic scintillating tiles read out via wavelength-shifting fibres and hybrid photodiodes (HPDs). The tile sizes change with η and vary from about 20 to 200 cm^2. Additional preshower detectors and forward calorimeters supplement the system. The tracker and calorimeter are located inside the solenoid. As will be discussed in the next chapter, this is also a requirement for particle flow techniques.

Particle Identification

To obtain a global event description, all particles have to be identified and measured. In many experiments, additional detectors for e.g. time-of-flight or Cherenkov radiation are used to obtain further information. CMS does not employ such detectors. Rather, particle ID is achieved by combining information from all sub-detectors. Here, the granularity of the detectors is an essential tool. For example, the granularity of the ECAL allows for γ/π^0 separation via their lateral shower shapes.

Triggering

Due to the huge data flow of some several tens of Mbyte every 25 ns, not every event can be stored to disk. First, a hardware trigger reduces the event rate from 40 MHz to 100 kHz based on information from the calorimeters and muon tracker. These are then transferred to an on-site server farm, where the events are analysed in software. This so-called High Level Trigger reduces the event rate to a few hundred Hz for storage.

3.1.2 Future Detectors

For CMS and other LHC experiments, the main challenges compared to previous experiments are the high-energy of the particles, the high radiation exposure and the large number of quasi-simultaneous collisions that need to be separated. Silicon detectors are mostly used for the vertex and tracking detectors due to their excellent spatial resolution paired with high rate capabilities, integration aspects, operational stability and relatively good radiation hardness.

With these characteristics, silicon sensors allow for integrated and compact tracker designs with low material budgets.

At this point in time, several new colliders for the post-LHC phase have been proposed. There are two large categories: lepton colliders and hadron colliders. Hadron colliders collide composite particles, typically protons, and as such the primary scattering process is between quarks and gluons. This implies that the energy and initial state of the scattering particles is not well defined. Lepton colliders, on the other hand, collide fundamental particles, typically electrons and positrons, and as such the scattering particles have a well defined initial state, which in turn allows for precision measurements of many physics observables. This poses very challenging requirements on the detector system to match the precision, e.g. in terms of jet energy and momentum resolution. Radiation damage is typically not an issue. For hadron colliders however, radiation damage together with pile-up are the main challenges.

For future colliders, silicon tracking detectors are likely to retain their important role. In future lepton colliders such as CLIC, silicon sensors have also been proposed as active material in sampling calorimeters. For the high luminosity upgrade of the LHC (HL-LHC), a proton-proton collider, the CMS experiment follows the same strategy with their calorimeter endcap upgrade. As will be discussed in the following two sections, the use of silicon in such calorimeters is driven by different reasoning.

3.2 The Compact Linear Collider

The Compact Linear Collider, CLIC, is a proposed concept for a future lepton collider. The study, hosted by CERN, aims to provide e^+e^- collisions at a centre-of-mass energy of up to $\sqrt{s} = 3\,\text{TeV}$ in the post-LHC era.

CLIC offers a unique sensitivity to particles produced in electroweak interactions. The physics programme includes precision measurement of Higgs [31] and top quark [32] properties as well as direct and indirect searches of physics beyond the Standard Model [33]. In view of the wide range of physics opportunities and to optimise the overall physics yield, it is foreseen to build CLIC in three stages of $\sqrt{s} = 380\,\text{GeV}$, $\sqrt{s} = 1.5\,\text{TeV}$ and $\sqrt{s} = 3\,\text{TeV}$ which will provide integrated luminosities of $1000\,\text{fb}^{-1}$, $2500\,\text{fb}^{-1}$ and $5000\,\text{fb}^{-1}$, respectively [9]. All stages foresee $\pm 80\%$ electron polarisation and no positron polarisation. The first stage is aimed at Higgs and top quark physics. Model independent measurements of the Higgs branching ratios and width will be enabled via $e^+e^- \to ZH$ and $e^+e^- \to H\nu_e\overline{\nu}_e$ production. A top quark pair production threshold scan around 350 GeV will result in a total uncertainty of about 50 MeV on the top quark mass, dominated by current theory uncertainties. The second stage and third stage give access to measurements of Higgs self-coupling via double Higgs events from $e^+e^- \to HH\nu_e\overline{\nu}_e$ and $e^+e^- \to ZHH$. All stages will provide a range of measurements that are directly or indirectly sensitive to BSM physics. Together, they will provide a precise probe of the standard model effective field theory with a mass reach well beyond the centre-of-mass energy of the collider.

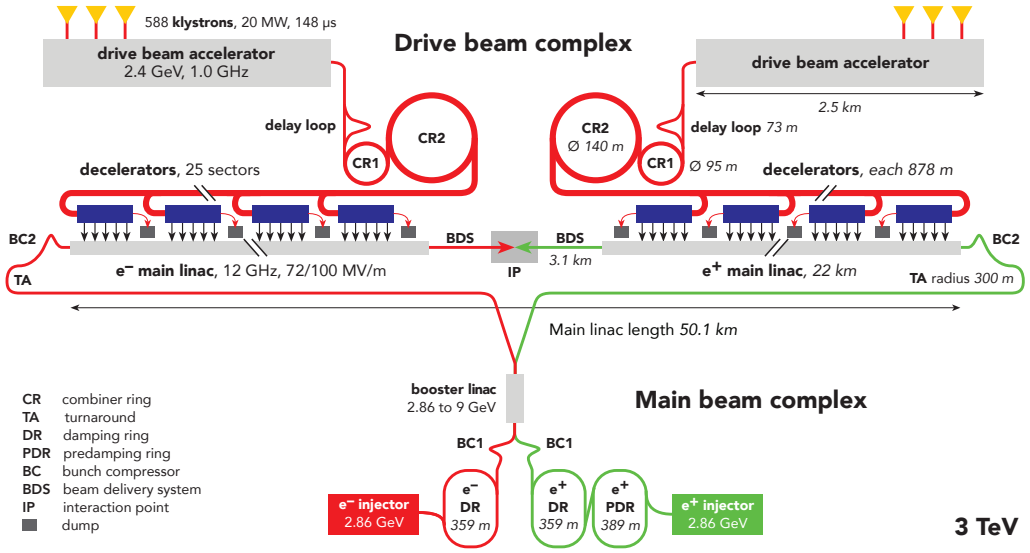


Figure 3.2: Two-beam acceleration scheme for CLIC, used to accelerate electrons and positrons to 3 TeV. A low-energy but high-intensity drive beam is first bunched and then decelerated to power the cavities for the main beam at the desired frequency of 12 GHz. This allows for accelerating gradients of 100 MV/m. Figure is taken from [9].

3.2.1 The Accelerator

For a linear collider to be able to accelerate particles to multi-TeV energies at reasonable site lengths, the accelerating structures must operate at very high electrical field gradients. This excludes the use of superconducting RF structures because their maximum gradient is intrinsically limited by the critical field of the used material. Normal conducting cavities, however, have been shown to hold accelerating gradients of 145 MV/m [9] with low breakdown rates when operated at a frequency of several GHz. For CLIC, the goal is to obtain gradients of 100 MV/m at a frequency of 12 GHz for the 3 TeV stage.

To power the accelerating cavities efficiently at this frequency, CLIC is based on a novel two-beam acceleration scheme, shown in Fig. 3.2 for the 3 TeV stage. The idea is to use a high-intensity but rather low-energy electron beam, the so-called drive beam, and restructure it into 12 GHz bunches via a delay loop and two combiner rings. This beam is then decelerated in dedicated cavities and the extracted 12 GHz power is transferred via wave guides to the accelerating cavities of the main electron and positron beams. The resulting beam structure shows a bunch spacing of 0.5 ns with 312 bunches making up one bunch train. The bunch train repetition rate is 50 Hz.

To achieve the desired instantaneous luminosity of about $6 \cdot 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, the beam sizes at the interaction point are focused to $\sigma_x \approx 40$ nm and $\sigma_y \approx 1$ nm at 3 TeV. The strong focusing, together with the beam structure of the two-beam acceleration scheme, creates very high charge densities in the collision region that cause beam-beam interactions. As a result, the energy of individual electrons and positrons is reduced. Collisions therefore take place over a wider range of energies. At 3 TeV, 35% of all collisions are within 1% of the nominal $\sqrt{s}$ value. This distribution, called

the luminosity spectrum, must be measured and de-convoluted in every physics study.

3.2.2 Requirements and Challenges

The design requirements for the CLIC detector are mostly driven by the desired precision of the physics measurements. The track momentum resolution determines the feasibility and precision of many physics studies, e.g. via the reconstruction of the di-muon invariant mass in HZ events. The aim is set to $\sigma_{p_T}/p_T^2 \leq 2 \cdot 10^{-5}$ per GeV/c. Another crucial parameter is the jet energy resolution. A value of $\sigma_E/E \leq 3.5\%$ for jet energies above 100 GeV is required for the separation of W and Z candidates in hadronic decays. Moreover, efficient identification of secondary vertices for flavour tagging of heavy quark states is needed. The derived requirement is a transverse impact parameter resolution of $\sigma_{d_0} = 5\,\mu m \oplus 15\,\mu m$ GeV/c/(p $\sin^{3/2}\theta$).

In addition to the physics-driven detector requirements, the beam-beam interactions set requirements for background suppression. Background particles from beam-beam interactions are mostly incoherent e^+e^- pairs and low p_T hadronic jets produced at low angles. To mitigate their influence on measurements of physics observables, a time resolution of 1 ns for the calorimeter hits and 5 ns for the tracking hits and a high granularity throughout the detector are required. Furthermore, a large geometrical coverage in the forward region down to 10 mrad is required for electrons and photons. The time structure of the beam allows for a triggerless readout of the full detector and power pulsing of the electronics.

3.2.3 CLIC Detector Overview

The CLIC detector concept started originally from the ILD and SiD concepts for the ILC [34, 35] and was adapted to CLIC specific requirements and the higher collision energy of 3 TeV. The resulting detector model, CLICdet, is shown in Fig. 3.3. The innermost part of the CLIC detector is a low-mass vertex detector followed by an all-silicon tracker. The detector is optimised for particle flow analysis. Therefore, the calorimeters must be placed inside the magnet and as close to the tracker as possible. The electromagnetic calorimeter has a depth of $23\,X_0$ and the hadronic calorimeter is $7.5\,\lambda_{int}$ deep. A superconducting solenoid that produces a 4 T magnetic field and an iron return yoke interleaved with detectors for muon identification are located on the outside. The forward region is equipped with two additional calorimeters for extended geometrical coverage (BeamCal) and luminosity measurements (LumiCal).

3.2.4 Tracking and Vertexing at CLIC

The requirements for the vertex detector are determined mainly by the desired momentum resolution and flavour tagging capabilities as well as the need for efficient background suppression. To achieve the aims outlined in the previous section, the goal is to reach a single-point resolution of $\sigma_{xy} \approx 3\,\mu m$ and a time resolution better than 5 ns. Besides the spatial resolution, the granularity is also driven by the maximum cell occupancy at the 3 TeV stage, which is required to be at most 3% per bunch train. To achieve this, the pixel pitch should not exceed 25 µm. Another challenge comes from the low material budget. The goal of $0.2\%X_0$ per layer translates to an equivalent of roughly 200 µm of silicon for sensor, readout, cooling, support and cabling. The

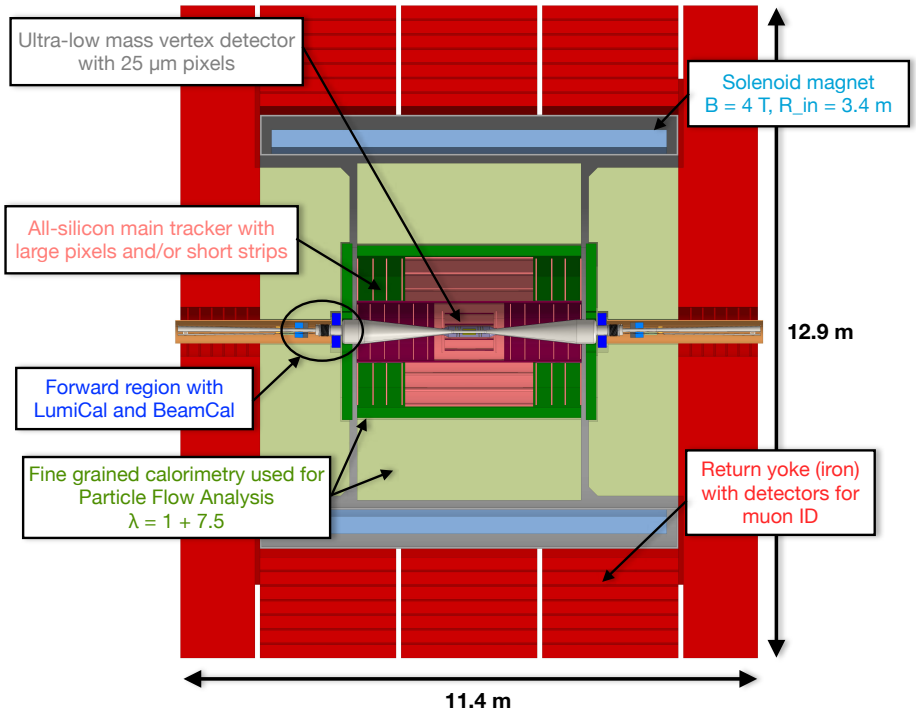


Figure 3.3: Cross section of CLIC detector model.

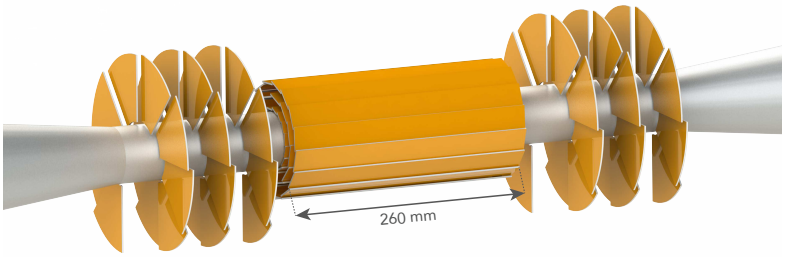


Figure 3.4: Geometry of CLICdet vertex detector. A barrel design with spiral endcaps, together with power pulsing of the electronics, allows for forced gas flow cooling. This is necessary to fulfil the strict material budget.

geometry of the vertex detector is shown in Fig. 3.4. It is designed in three double layers and has an overall length of 560 mm. The innermost barrel layer is located at a radius of 31 mm from the interaction point.

To meet the low mass requirements and avoid the need for liquid cooling, the layout of the vertex detector is optimised for low power dissipation. This is achieved via power pulsing of the electronics: Taking advantage of the pulsed beam structure, most of the electronics is powered down after every bunch train and only switched back on at a few µs before the next one. Together with an optimised spiral geometry in the endcaps, this allows for forced air flow cooling.

For the tracker, a single-point resolution of $\sigma_{r\varphi} \approx 7\,\mu\text{m}$ in the $r\varphi$ plane and a time resolution of better than 5 ns are needed to achieve the desired momentum resolution and background suppression. The total material budget shall not exceed 1.5% X_0 per layer. An all-silicon tracker with a lightweight carbon fibre support structure is envisaged to meet these requirements. In the $r\varphi$ plane, a 50 µm pitch is needed to reach the required single-point resolution. To stay below the occupancy limits, a maximum cell pitch between 1 mm and 10 mm in the z direction is needed, depending on the exact position inside the detector.

3.2.5 Calorimetry at CLIC

The calorimeter requirements are dominated by the need for a jet energy resolution of approximately 3.5% for jet energies above 100 GeV. With conventional sampling calorimetry, this is very difficult to achieve.

Limitations in the Hadronic Energy Resolution

The approach of measuring the energy of hadronic particles by measuring their energy loss in a given material is intrinsically limited because their shower development carries electromagnetic (e) and hadronic (h) components which interact differently in the material and the ratio of e/h, even at the same initial energy, differs from event to event (see also Sec. 2.3.4). While a clever combination and arrangement of different materials can achieve a mean ratio of $\text{e/h} = 1$,

i.e. a compensating calorimeter, on an event-by-event basis the fluctuations may result in a ratio different from 1. To improve beyond this intrinsic limitation, two approaches are generally considered: measure the e/h ratio even by event or reduce the usage of the hadronic calorimeter in the energy reconstruction as much as possible. CLICdet principally relies on the second approach that is exploited by a technique called Particle Flow Analysis (PFA).

The Particle Flow Paradigm

In an average jet, about 60% of the energy is carried by charged particles of which most are hadrons, about 30% by photons and 10% by neutral hadrons [36]. The approach of Particle Flow Analysis is to measure the momentum of the 60% charged particles in the more precise tracking detectors (via their momentum) rather than in the calorimeter. The energy of photons and neutral hadrons are still measured in the calorimeters. Following this approach, the limitations from hadronic calorimetry affect only about 10% of the measured energy. To avoid double counting, the individual energy depositions of all particles have to be resolved separately and correctly assigned. This is achieved by using calorimeters with a high lateral and longitudinal granularity as well as a design that aids the separation of electromagnetic and hadronic particles. For example, a small Molière radius ρ_M results in narrow electromagnetic showers and helps separating γ from π^0 showers. Similarly, a large ratio of λ_{int}/X_0 helps to separate hadronic and electromagnetic showers in the longitudinal direction. The solenoid is located outside the calorimeters to facilitating a correct assignment of the calorimeter clusters to tracks originating from the same particle.

In such a detector, the jet energy resolution is dominated by wrong association of energy depositions rather than the intrinsic energy resolution of the calorimeters. Effectively, this approach changes the problem of summing up energy depositions into a problem of pattern recognition. Dedicated software to reconstruct the particle flow is necessary for the analysis. For CLIC, PandoraPFA is used [36, 37].

The CLICdet Calorimeter

The granularity choices for CLICdet are 5 by 5 mm^2 in the electromagnetic calorimeter and 30 by 30 mm^2 in the hadronic calorimeter. This is driven by the effective lateral dimensions of electromagnetic and hadronic showers and how well they can be separated. It has been shown that smaller cell sizes do not improve the energy resolution in a significant way [38]. A time resolution of about 1 ns at the cell level is foreseen to suppress energy deposits from beam induced background.

The ECAL uses 40 layers of silicon pad sensors as active material and tungsten (W) as absorber. Tungsten has a large ratio of λ_{int}/X_0 and a very small Molière radius ρ_M (see Tab. 2.2). The calorimeter is however not made of pure tungsten but is interleaved with the active material, mechanical support and cabling. To keep the favourable characteristics of tungsten, it is important to keep the distance between two absorber layers small. This is facilitated by a low power consumption of the electronics (aided by power pulsing) that allows for using passive conductive cooling within the layers, thereby reducing the thickness of the active layers overall. For the latter, silicon pad sensors of 500 µm thickness are foreseen. Silicon as active material also offers the required granularity, operational stability, a high sampling fraction and an excellent

signal-to-noise ratio. However, both silicon and tungsten are expensive.

For the HCAL, the choice is 60 layers of 3 mm thick plastic scintillating tiles with silicon photomultiplier (SiPM) readout and steel absorbers. Steel is significantly cheaper than tungsten and it has been shown that for a hadronic calorimeter nearly the same performance can be reached [39].

Particle Flow in Future Detectors

All future lepton colliders have at least one detector design based on a calorimeter optimised for PFA similar to CLICdet. Recently, CMS has followed the same strategy with a similar design for its endcap calorimeter upgrade at HL-LHC. As this will be the first full scale calorimeter of this kind, it is an excellent learning opportunity for CLIC.

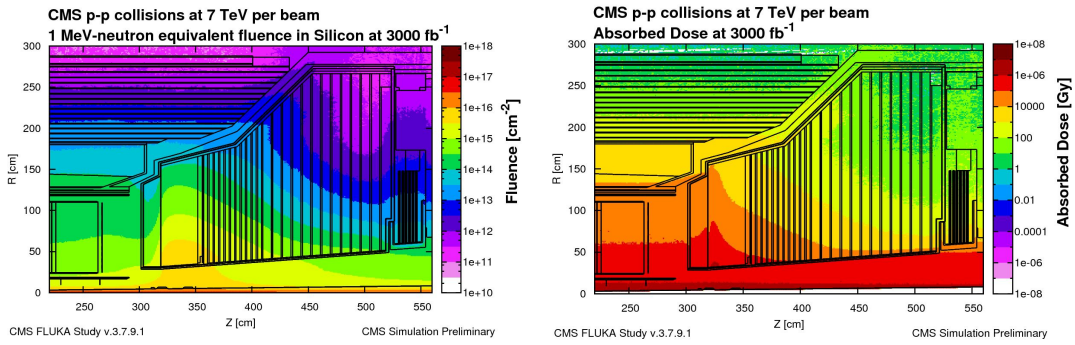
3.3 The CMS HGCAL Upgrade

From 2026 onwards, LHC will have concluded a major luminosity upgrade. In the baseline scenario, the instantaneous luminosity will increase to about $5 \times 10^{34}\,\mathrm{cm^{-2}s^{-1}}$. This is achieved by a redesign of the magnetic focusing system and the addition of crab cavities to rotate the beam just before the collision as well as a general increase in beam current [8]. The upgraded accelerator called the High-Luminosity LHC (HL-LHC), will run for 10 years to deliver an integrated luminosity of $3000\,\mathrm{fb^{-1}}$, about 10 times the luminosity of the LHC. The average pile-up will increase to 140 collisions per bunch crossing. Additionally, the so-called ultimate scenario foresees an instantaneous luminosity of $7.5 \times 10^{34}\,\mathrm{cm^{-2}s^{-1}}$, an integrated luminosity of about $4000\,\mathrm{fb^{-1}}$ and an average pile-up of 200. The current CMS detector was designed for operation at pile-up 25 and up to $500\,\mathrm{fb^{-1}}$ [28]. To cope with the new environment and retain a good physics performance over the full running period, several upgrades to CMS are planned [40].

3.3.1 Requirements and Challenges

The main challenges of a detector operation at HL-LHC are similar to the ones at LHC but even more demanding. Fig. 3.5 shows the expected hadron fluences and total ionising dose as a function of r and z in the CMS detector endcap. In the innermost regions, the detector has to withstand $10^{16}\,\mathrm{n_{eq}/cm^2}$ and 150 MRad. Under these conditions, the current endcap calorimeter would degrade very quickly in performance [40] and is therefore completely replaced by the new calorimeter endcap CE, also referred to as HGCAL (High Granularity Calorimeter).

The second challenge is the pile-up that will increase to an average 140 (200) per bunch crossing. The main strategy to distinguish pile-up events from each other is high spatial granularity and precise timing information. For the collision characteristics at HL-LHC, CMS assumes a Gaussian profile in time with a width of 180 to 200 ps [42]. The spatial bunch crossing profile with 45 mm RMS in longitudinal direction will result in a peak vertex density per bunch crossing of $1.2\,\mathrm{mm^{-1}}$ ($1.8\,\mathrm{mm^{-1}}$) compared to $0.3\,\mathrm{mm^{-1}}$ for the LHC. If a hit time resolution of about 30 ps is available, then timing cuts in the event reconstruction can effectively reduce the vertex density to LHC



(a) (b)

Figure 3.5: The expected integrated hadron fluences for the CMS endcap are shown in (a) and the total ionising dose in (b). The flux and dose are varying with r and z, allowing for different technology choices depending on the exact location. Figures are from [41].

levels. For charged particles, this resolution will be provided by a new separate MIP timing detector. For photons, the ECAL is used.

3.3.2 Overview of HGCAL

The CMS HGCAL consists of an electromagnetic part called CE-E and two hadronic parts called CE-H. The system is shown in Fig. 3.6. The electromagnetic part will be $25\,X_0$ deep and will consists of 28 layers of silicon pad sensors as active elements with lead in a stainless steel envelope as main absorber material. The two hadronic parts are in total $8.5\,\lambda_{\text{int}}$ deep with 24 layers and steel absorbers. As active elements, silicon will be used in the high radiation regions. In low radiation regions, scintillating tiles with SiPM readout are used. The full system will be maintained at -30°C using evaporative CO_2 cooling to limit the leakage current of the silicon sensors.

The design is targeted to achieve high performance for physics objects reconstructed in the presence of high levels of pile-up. The high transverse and longitudinal granularity of the detector, together with its timing performance, allows for the separation of signal and pile-up events within the calorimeter. Particle Flow Analysis in 5 dimensions (space, energy and time) will be an essential tool to achieve this.

3.3.3 Active Elements

One of the most relevant parameters for the detector performance is the signal-to-noise ratio, S/N. For silicon, it has been shown that the signal loss due to irradiation is decreased in thinner sensors and when operating at increased bias voltages [43]. This is shown in Fig. 3.7. The reason is the higher electric field that reduces the trapping efficiency of charges moving through the sensor. Because the increase in the electric field after full depletion is constant throughout the bulk of the sensor, using over-depleted thin sensors is more effective then under-depleted

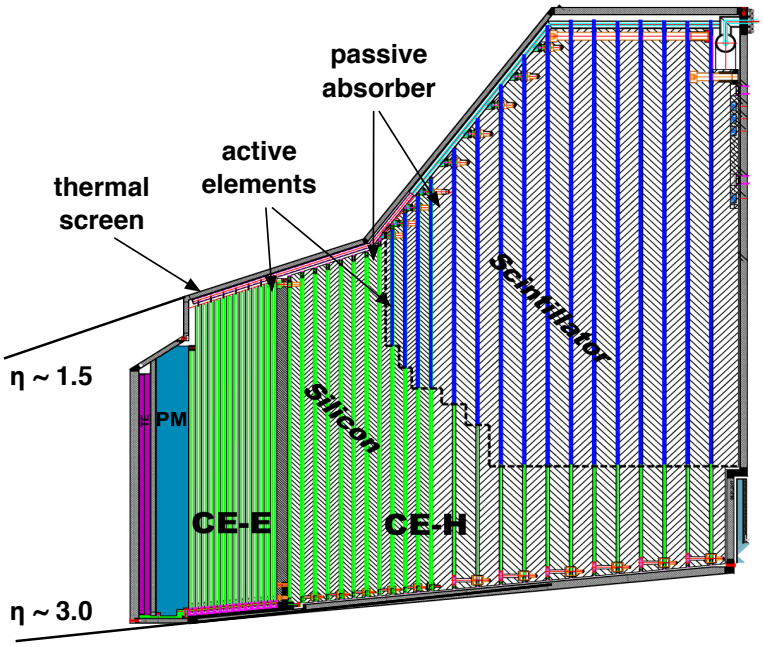


Figure 3.6: Overview of the HGCAL endcap. The Figure is adapted from [41, Fig. 1.6].

sensors of equal active thickness. The increased noise contribution from the leakage current can be mitigated by cooling. Additionally, silicon sensors can deliver time resolutions well below the required 30 ps when operated in a calorimeter environment. This is shown in Chapter 6 of this work.

The use of 8 inch silicon sensor wafers with hexagonal geometry is foreseen to reduce cost as a larger fraction of the circular wafer can be used. The active thickness will be adapted to the expected radiation dose and will vary between 120, 200 and 300 µm. A granularity of 0.5 cm^2 for the 120 µm and about 1.2 cm^2 for 200 and 300 µm thick sensors will be used.

The scintillating tiles will be arranged in a SiPM-on-tile geometry, where the SiPM is located in a small cutout on the backside of the tile, which aids uniform light collection and integration with the readout PCB. The size of the tiles will vary throughout the detector from 4 cm^2 to 32 cm^2, increasing with the radius r while covering a constant fraction in ϕ. As for CLIC, 3 mm thick tiles are foreseen.

The detector will be calibrated with minimum ionising particles (MIPs). It is therefore essential that even after 3000 fb^{-1} the S/N for MIPs is sufficiently high. The transition between silicon pad sensors and the less radiation hard plastic scintillators (and SiPMs) is determined by the point at which the S/N for the plastic scintillators will drop below 5 after 3000 fb^{-1}. For the silicon sensors, it is foreseen to have a few cells with smaller area than the regular ones on each sensor. The smaller area at unchanged thickness will reduce the noise contributions from capacitance and leakage current in these cells, such that they are still sensitive to single MIPs after 3000 fb^{-1}.

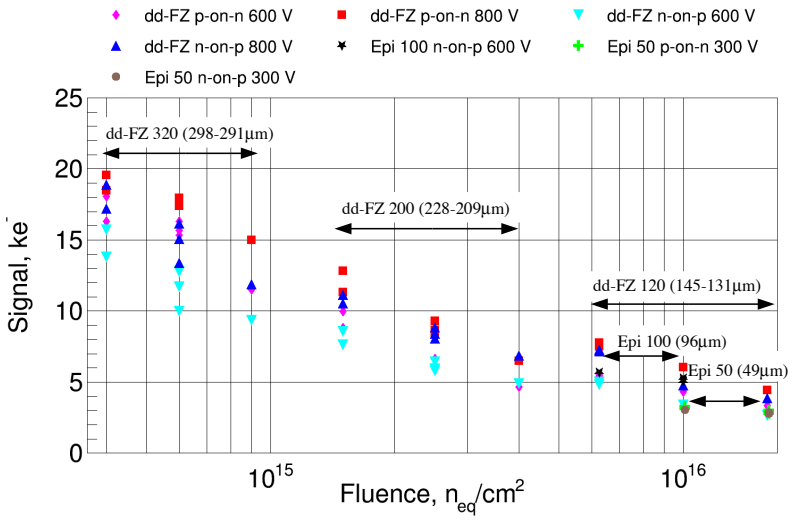


Figure 3.7: The most probable signal of a MIP in silicon diodes for different neutron fluences. Thinner sensors and operation at higher bias voltages mitigate the signal loss. Figure is taken from [41].

3.3.4 Module Integration

Silicon modules start with a metallic baseplate (copper tungsten in CE-E and pure copper in CE-H), that has a polyamide gold-plated foil glued to it. The silicon sensor is then glued onto that foil. The readout PCB, which hosts the front-end ASICs, is in turn glued onto the sensor and wire-bonds reaching through holes in the PCB connect to the sensor contact pads. The baseplate acts as interface to the cooling system and as additional absorber material.

3.3.5 Readout Electronics

The driving requirements for the front-end readout ASIC are the large dynamic range of 0.4 fC to 10 pC (15 bits) of deposited charge, the noise level below 2000 e- at 40 pF load capacitance, timing information with better than 30 ps accuracy and radiation hardness up to 150 MRad. The goal for the power budget amounts to 10 mW/channel for the analog part. To meet these requirements, a chip based on the OMEGA ASIC family [44] is being developed. The baseline option includes a traditional sampling gain stage and a time-over-threshold stage, as well as a time-of-arrival path. The ASIC will be fabricated in TSMC 130 nm CMOS technology.

3.3.6 Comparison to the CLICdet Calorimeter

The main target of HGCAL is to achieve a similar physics performance as in the current detector at the increased pile-up environment and to sustain that performance up to 3000 fb^{-1} integrated luminosity. In comparison to CLIC, the choice of silicon is therefore mainly driven by the

radiation hardness, the granularity that can be achieved and the timing capabilities. Due to the high radiation exposure, approximately 10^4 to 10^5 times higher than at lepton colliders, a significantly more powerful cooling system is required. The cooling pipes are located inside the active calorimeter volume. This increases the spacing between adjacent calorimeter layers from 4 mm envisaged for CLIC to about 20 mm for HGCAL. The advantage of using tungsten as absorber material is then attenuated and lead presents a more cost effective solution at similar performance.

Chapter 4

Calibration and Performance Studies of Thin Silicon Pixel Detectors

As explored in Sec. 3.2, the CLIC vertex detector will require a spatial resolution of about 3 µm and a time resolution of approximately 5 ns at a material budget of 0.2% radiation length per layer [9]. These constraints in the material budget require the combination of sensor and ASIC to be implemented in a total of 100 to 150 µm of silicon, with the remaining budget needed for cables and support. In case of a hybrid assembly, i.e. sensor and ASIC are implemented on separate substrates and connected e.g. via bump bonds, this most likely results in a 50 µm thin sensor bonded to a 50 µm thin readout ASIC. Thinning the ASIC to 50 µm is expected to have to have negligible impact on the performance because all functionality is implemented in the first 10 to 20 µm of the substrate while the remaining substrate provides only mechanical stability.[1] The performance implications of planar sensors of 50 µm to 100 µm however, have to be investigated. Here, the Timepix3 ASIC is used as a test vehicle to study these sensors.

Timepix3 [45] is a 65k channel multi-purpose readout ASIC from the Medipix family [46, 47] for hybrid pixel detectors. Its versatility, high rate capabilities and excellent performance makes it a suitable test vehicle to study thin silicon sensors, even though its geometry of 55 µm by 55 µm pixels will not enable a 3 µm spatial resolution. It can measure time and energy simultaneously by employing time-of-arrival (ToA) and time-over-threshold (ToT) techniques. Both techniques are systematically affected by timewalk which results in a non-linear dependence of the energy and time response on the signal amplitude.

This chapter investigates the performance of thin silicon sensors as well as the calibration method suited for an ASIC with $O(10^4)$ channels, and more generally, highly granular detectors. Five assemblies of Timepix3 ASICs bump bonded to silicon pixel sensors of different thicknesses between 50 µm and 150 µm are calibrated and tested in particle beams. In Sec. 4.1, the experimental setup of the Timepix3 ASICs, the sensors and the further readout are described as well as the method-

[1] Implications from the additional mechanical stress and engineering aspects such as ASIC handling and assembly need to be studied however.

ology that is used for the timing studies. Sec. 4.2 studies the general characteristics of the ASIC such as noise and threshold dispersion. Sec. 4.3 describes the ASIC's calibration by a combination of test pulse injection and radioactive source or beam measurements. Sec. 4.4 shows the performance of thin silicon sensors and the impact of the calibration.

Significant parts of this chapter have also been published by the author in [48] and [49] and are quoted verbatim. Some figures are also shown in [50].

4.1 Experimental Setup and Methodology

4.1.1 The Timepix3 ASIC

The Timepix3 ASIC comprises a matrix of 256 by 256 square pixels with 55 µm pitch. A simplified pixel layout of the ASIC is shown in Fig. 4.1. Each pixel consists of a charge sensitive amplifier (CSA) with a feedback capacitor for amplification and shaping of the input signal. If there is a voltage difference between the input and output node of the CSA, the capacitor is actively discharged via a current controlled by a Krummenacher feedback network [51]. The discharge current is limited by the network, resulting in a constant discharge. The CSA output is connected to a single threshold discriminator with the possibility of a 4 bit local threshold adjustment. The global threshold is controlled by two DACs (fine and coarse) that are combined and linearised. The least significant bit (LSB) of the fine DAC corresponds to 0.5 mV. The discriminator output is then connected to the in-pixel digital logic, where the signal is processed.

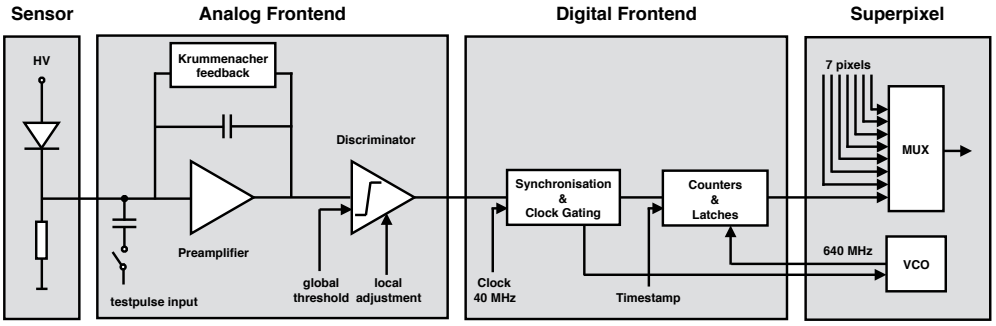


Figure 4.1: Schematic of the Timepix3 readout circuit. Reproduced from [45, fig. 3] with small adaptations.

In the main operating mode used in the presented work, the ASIC employs a 10 bit time-over-threshold (ToT) and a 14 bit time-of-arrival (ToA) binary counter running at a 40 MHz clock for simultaneous measurements of energy and time. This mode is called ToT & ToA mode. An example operating diagram is shown in Fig. 4.2. If the input signal rises above the constant threshold value, the discriminator starts the ToT counter which then runs until the signal falls below the threshold again. The time the signal stays above threshold is a measure of the signal amplitude. At the time the signal crosses the threshold, the ToA value is taken from a global

Gray code counter that is clocked at 40 MHz. An additional local 4 bit counter is clocked at 640 MHz, but only until the next rising edge of the 40 MHz clock. Combining the counter values for the 40 MHz global clock and 640 MHz fast clock results in a more precise time and energy measurement. A more detailed description of the ASIC can be found in [52].

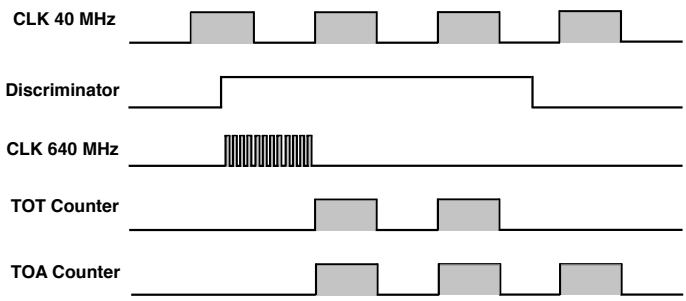


Figure 4.2: Operating principle of the ToT & ToA operating mode. Reproduced from [45, Fig. 4] with small adaptations.

Alternatively, the ASIC can be configured to operate in photon counting and integrated ToT (PC & iToT) mode in which the total number of hits above threshold and the integral ToT of those hits are recorded. In the presented work, this mode of operation is used for noise measurements and threshold scans.

Each block of 2x4 pixels makes up a super-pixel sharing a range of common infrastructure, most importantly the same voltage-controlled oscillator (VCO) that delivers the 640 MHz clock. Additionally, a set of 2x16 pixels share, among others, the clock buffer.

Data can be recorded in two different ways:

- In *frame-based* operation, data are recorded for a given amount of time (the shutter or frame time). Subsequently, all pixels are read out simultaneously. This mode introduces a global dead time and the shutter time has to be chosen carefully to avoid pile-up. In this work, this mode is used for noise measurements only.
- In *data driven* operation, each pixel is read out individually and immediately after an event was registered. This mode has no global dead time as long as the data rate stays below several tens of MHz. This is the major readout mode used in the presented work.

Time and Energy Measurement

Owing to the use of a constant value threshold, the time at which the signal crosses the threshold depends on the signal amplitude as the rise time of any signal is a system constant. This phenomenon is commonly called timewalk. The shift in time $f_{\mathrm{toa}}(x)$ with amplitude x can be modelled with three free parameters as

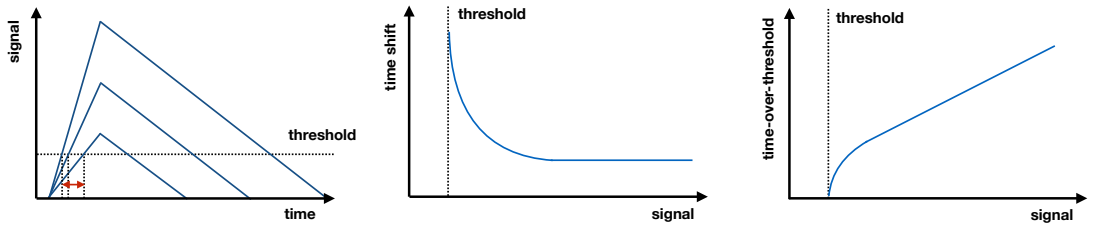
$$f_{\mathrm{toa}}(x) = \frac{c_{\mathrm{toa}}}{x - t_{\mathrm{toa}}} + d_{\mathrm{toa}}. \tag{4.1}$$

The three free parameters c_{toa}, t_{toa} and d_{toa}, correspond to the curvature, the asymptote and the offset of the response function.

As the energy measurement is based on time as well, it suffers from the same phenomenon. Assuming a linear discharge, the energy response $f_{\mathrm{tot}}(x)$ can be modelled with four free parameters as

$$f_{\mathrm{tot}}(x) = a_{\mathrm{tot}} \cdot x + b_{\mathrm{tot}} - \frac{c_{\mathrm{tot}}}{x - t_{\mathrm{tot}}}. \tag{4.2}$$

The free parameters a_{tot} and b_{tot} correspond to the gain and offset of the linear part while c_{tot} and t_{tot} again describe the curvature and asymptote of the non-linear part. The effect of timewalk on both ToA and ToT is sketched in Fig. 4.3.



(a) (b) (c)

Figure 4.3: Timewalk effect (a) resulting from signals with different amplitudes but constant rise time crossing a constant threshold on the ToA (b) and ToT (c) response.

Both function are motivated by trigonometric considerations described in App. B. It should be stressed that the assumptions that go into the derivation are only approximate and deviations are possible. As will be shown at a later point, the model fits sufficiently well for Timepix3.

The gain and dynamic range of this measurement can be controlled via the discharge current i_{krum} of the Krummenacher network. It can be set by a 7 bit DAC. This value is a trade-off between precision and temperature stability: A smaller discharge current can in principle increase the precision of the energy measurement but requires a more stringent temperature control. For the chosen value, a change in gain of about 3 % over 20 °C has been observed. The current also determines the dynamic range of the readout. For the presented work, this DAC value was set to 10 which corresponds to a total dynamic range of roughly 150k electrons. The temperature was controlled with a fan to about ±1 °C. If not stated otherwise, the operating temperature was 54 °C.

Test Pulse Injection Input

Each pixel of the ASIC has a possibility for test pulse injection (see Fig. 4.1). Here, the test charge is injected by applying a step voltage to a test capacitor that is connected in parallel to the sensor input. The design value of this capacitor is 3 fF. The injection input can be activated and deactivated separately for each pixel. An ADC allows to measure the applied voltage. The pulse generator as well as the ADC are global to the ASIC and located at the periphery. Therefore, the voltage that is actually applied to the input capacitor at each pixel is not a priori known.

Table 4.1: List of assemblies and their nominal operating conditions used in the presented work.

Chip ID [-]	Sensor Thickness [µm]	Type [-]	Nom. Op. Voltage [V]	Nom. Threshold [LSB]
W19_C7	50	n-on-p	-15	1190
W19_F7	50	n-on-p	-15	1187
W19_G7	50	n-on-p	-15	1148
W5_E2	100	n-on-p	-20	1160
W5_F1	150	n-on-p	-30	1153

There is also the possibility to inject a test pulse directly into the digital frontend, bypassing the preamplifier and discriminator (digital test pulse).

4.1.2 The Sensor

The Timepix3 ASIC is bump-bonded to a planar silicon pixel sensor of the same pixel pitch as the ASIC. Five different n-on-p sensors with active edge processing and different thicknesses are investigated. Sensor thicknesses of 50 µm, 100 µm and 150 µm thickness are investigated. The list of assemblies is shown in Tab. 4.1. The naming scheme is based on the ASIC and consists of the wafer number (e.g. W19) followed by the position of the ASIC on the wafer (e.g. C7). The nominal operating voltage is set to be approximately 5 V above the depletion voltage. The nominal operating threshold was obtained from noise and baseline measurements explained in the following sections.

4.1.3 Other Elements of the Readout Chain

The assembly of ASIC and sensor is wire-bonded to a mezzanine PCB, the so-called chip board, that handles mainly the power distribution for the ASIC and the bias voltage supply for the sensor. The chip board is then connected via an FMC connector to the SPIDR readout board [53]. There, an FPGA handles data acquisition and slow control. One SPIDR board can hold two chip boards. A TDC channel with a precision of about 250 ps [53] is available for an external trigger input. The recorded data is transmitted via a 10 Gbit Ethernet connection to a PC.

4.1.4 X-ray Tube Setup

For the energy calibration, measurements in an x-ray tube have been conducted. Here, electrons from a thin filament are accelerated to 40 keV by an electrostatic field, focused and sent onto an external target material. The resulting fluorescent x-rays are measured with a Timepix3 assembly. The setup is shown in Fig. 4.4.

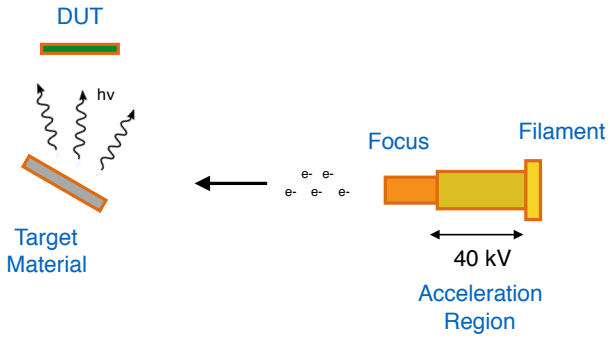


Figure 4.4: Schematic of the setup used for x-ray tube measurements.

4.1.5 Testbeam Setup

For the time calibration and final performance evaluation of the assemblies, beam tests have been conducted in the CERN SPS H6 beamline with 120 GeV incident pions. The Timepix3 assemblies have been measured in the CLICdp telescope which is shown in Fig. 4.5. The telescope employs six tracking planes of Timepix3 assemblies, each with a 300 µm thick sensor. Three planes are upstream and three are downstream of the device under test (DUT). In total, four SPIDR systems are used to read the seven assemblies. The telescope planes are tilted by 9 degrees to optimise charge sharing. This way, a pointing resolution at the DUT position of about 1.8 µm [54] is achieved. The track time resolution is about 1.1 ns, as shown in this work. Two organic scintillators read out by photomultiplier tubes (PMT) in the front and in the rear can be used to obtain a more precise reference timestamp. The telescope itself is not triggered but all recorded hits are read out and matched offline via their timestamps. Hits on the scintillators are only read out if they are recorded in coincidence with each other. They are then sampled by the TDC input of the second SPIDR board.

For the track reconstruction, the data stream is divided into slices of 25 µs. Clusters on the six planes that coincide within ±200 ns and ±200 µm to each other are associated with a track candidate and reconstructed with a straight line fit. If the fit passes a χ^2/ndof criterion of less than 3 it is considered a track. On average, eight tracks are found within the 25 µs.

Three timestamps can then be extracted from track, scintillators and DUT. The track timestamp is taken as the average of the six associated clusters on the six telescope planes, corrected for average time-of-flight and latencies. Per cluster, the timestamp of the pixel with the largest energy deposition is assigned as cluster timestamp. The DUT timestamp is obtained in the same way. Lastly, a coincidence scintillator timestamp is associated with a track if one can be found within a window of ±7 ns around the track timestamp.

The scintillator has a better time resolution than the telescope tracks and is therefore the preferred time reference. However, due to the low efficiency of the scintillators (about 75% per device) and the untriggered data stream, a finite window length is necessary to avoid association with a wrong track. The value of ±7 ns is motivated by Fig. 4.6 which shows the RMS of the time difference between track and DUT for different window lengths as well as the probability

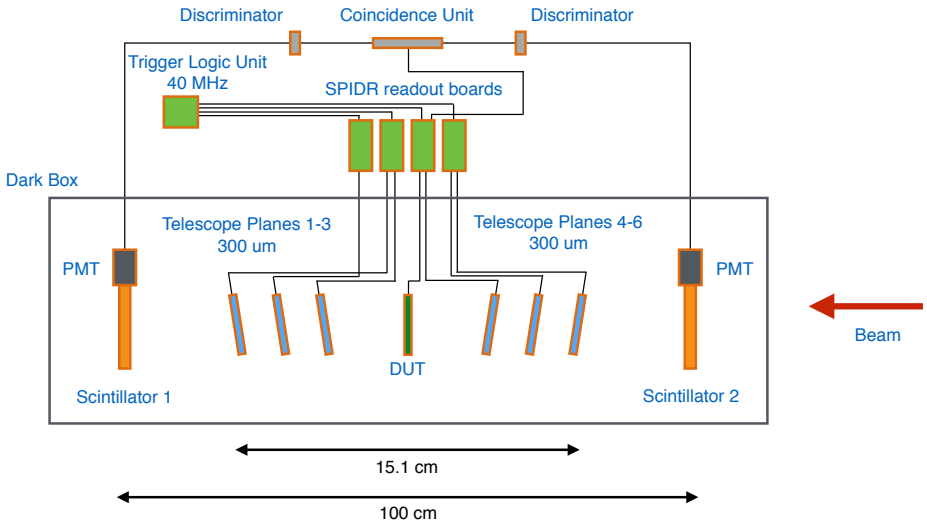


Figure 4.5: Schematic of the telescope setup and readout used for beam tests.

of finding a coincident scintillator hit within that window. Since the three timestamps are all independent, the length of the acceptance window used to search for a scintillator timestamp must not bias the measured RMS between track and DUT timestamps. A window of ±7 ns is the smallest value that fulfils this criterion. The probability of finding a coincident scintillator hit within this window is 61%, in agreement with previous lab measurements. At smaller windows, the probability and RMS of the residual $t_{track} - t_{dut}$ distribution starts to drop as the distribution becomes biased towards smaller time differences. At larger windows, the probability of finding a scintillator hit rises due to mis-association. The coincident scintillator timestamp is then used as a time reference and tracks are used only as position reference. Tracks without associated scintillator hits are discarded from the analysis.

4.2 General Characterisation of the Timepix3 ASIC

4.2.1 Threshold Variations and Equalisation

During the wafer production, process variations can lead to differences in behaviour across the ASIC. For an energy measurement based on time-over-threshold, pixel-to-pixel variations of the threshold are particularly relevant because they result in different turn-on points. Similarly, a time measurement based on time-of-arrival will suffer from an additional delay that varies from pixel to pixel and, if not corrected for, will result in an additional source of uncertainty on the time resolution. To counteract pixel-to-pixel variations in the effective threshold, the Timepix3 ASIC employs a local voltage correction to the global threshold. This 4 bit adjustment setting is called *trim DAC*. The goal of the threshold equalisation is to find the set of trim DACs that

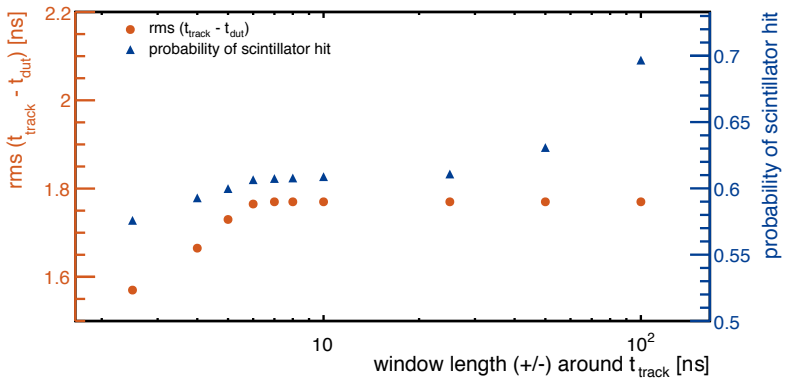


Figure 4.6: Influence of the time window used to search for a scintillator timestamp on the measured RMS between track and DUT timestamps.

minimises the dispersion of the effective threshold across the pixel matrix.

Due to gain dispersion, the equalisation should be done ideally at the operating threshold of the assembly. However, for practical reasons this is often not possible. In the presented work, the ASIC is equalised at the edge of the noise floor defined as the first threshold setting that records more than 10 noise hits in a time of 1 ms. The equalisation is then achieved by the following procedure.

1. Set the trim DAC setting for all pixels to the minimum value (0x0) and do a threshold scan for the noise edge. Repeat the procedure with the maximum trim DAC value (0xF). The result is two broad distributions representing the noise edge position in units of the global threshold DAC for all pixels in the matrix.
2. Extract the global average threshold at trim DAC 0x0 and 0xF as the mean values of the two distributions. Set a first equalisation target as midpoint between the two global threshold values. For each pixel, interpolate to find the trim DAC setting closest to that target.
3. Scan again for the noise edge and set the refined equalisation target as the mean of the new distribution. For each pixel, see if a better trim DAC setting can be found and scan again for the noise edge.
4. Iterate the last point until no change is observed anymore. Typically, one iteration is enough.

The measured noise edge dispersion before and after the equalisation is shown in Fig. 4.7. It is notable that the distribution becomes box-like after the equalisation, indicating that the precision of the process is limited by the binning of the trim DACs (4 bit). The RMS is typically reduced by a factor 6 to 7 to about 3 LSB.

During steps 2, 3 and 4, if no trim DAC setting can be found that brings the pixel within the range of one trim DAC step to the equalisation target, the pixel is identified as noisy and masked. Additional pixels might be masked manually at later stages. In total, between 33 and 46 pixels have been masked per assembly.

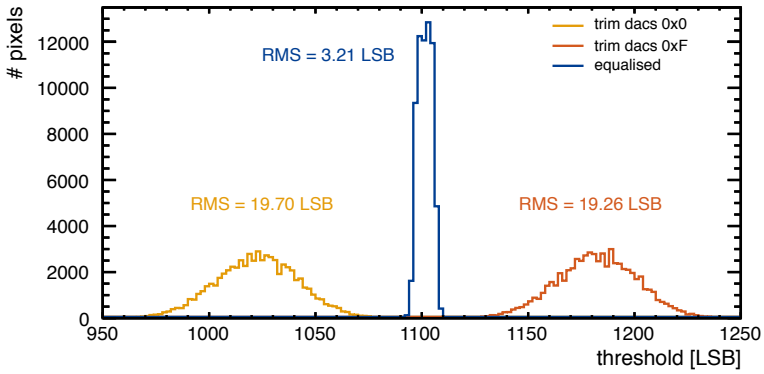


Figure 4.7: Illustration of the threshold equalisation. The distributions of the noise edge (see text) for trim DAC 0, trim DAC F and equalised trim DACs are shown. After equalisation, the RMS of the distribution is typically reduced by a factor 6 to 7. Assembly shown is W5_E2.

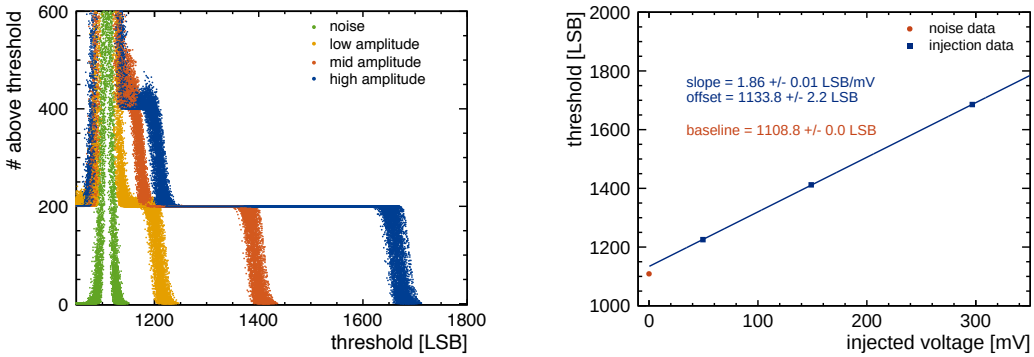
4.2.2 Threshold Gain Measurement

To establish the gain per threshold LSB, a test pulse with a fixed amplitude is injected into one pixel frontend N times. Then a threshold scan around the expected amplitude range is performed and the number of signals seen by the frontend are counted at each threshold. The result is the cumulative distribution function (CDF) of the normal distribution. This function, a modified error function, is given in Eq. 4.3 and is often called s-curve.

$$f_{\mathrm{cdf}}(x) = \frac{N}{2}\left[1 - \mathrm{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right] \tag{4.3}$$

The mean μ of this function corresponds to the test pulse signal and σ to the noise. This is repeated for three test pulse amplitudes. A linear fit is performed and the gain can be extracted. A noise scan determining the baseline is also performed but not included in the fit. The procedure is quite lengthy as only a subset of pixels can be injected with pulses at the same time to avoid crosstalk. Therefore, the gain calibration is only done for one diagonal of 256 pixels per chip.

Fig. 4.8(a) shows the resulting s-curves of 200 pulses injected in each of the 256 diagonal pixels for an n-on-p sensor. The polarity of the threshold DAC is such that a higher DAC value corresponds to a higher threshold for negative input signals. Therefore, at large threshold values no pulses are seen at the input as they are all below threshold. Lowering the threshold, more and more pulses are observed until all 200 injected pulses are seen at the input. At even lower thresholds, a second s-curve can be seen. This is the result of the undershoot of the injected step voltage. When decreasing the threshold even further, superimposes the injected pulses with the noise floor. Per pixel, each transition region is fitted with a CDF and the mean values of the curves are fitted with a linear function to extract the gain of the pixel in units of LSB/mV. This is shown in Fig. 4.8(b) for an example pixel. A dedicated threshold scan around the noise floor is also executed. As one can see from the fit in Fig. 4.8(b), the test pulse voltage has an offset with respect to the baseline. The same effect is seen in other test pulse measurements and discussed in more detail in Sec. 4.3. Here, only the mean values of the gain are used to convert the noise



(a) (b)

Figure 4.8: Illustration of the threshold gain measurement. 200 test pulses are injected into 256 pixels at 3 different test pulse amplitudes (a). The colours correspond to the different amplitudes. A noise scan of the baseline is also shown. For each pixel, the s-curves are fitted with a CDF function and a linear fit is performed to extract the gain per threshold step (b). Assembly shown is W5_E2.

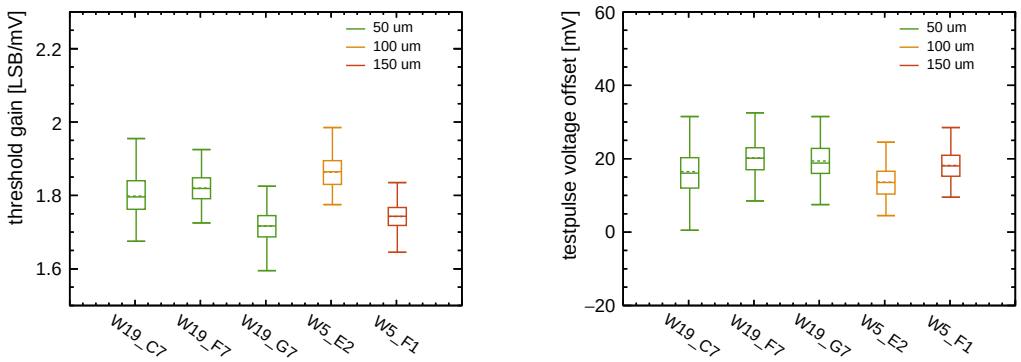
measurements described in the following section into a voltage. Fig. 4.9 shows the offset and gain for all assemblies.

4.2.3 Noise Performance

To measure the noise, the threshold is scanned across the noise floor and the resulting distribution is fitted with a Gaussian. The mean of the Gaussian is defined as the baseline and the standard deviation of the noise. Fig. 4.10 shows a map of the resulting values across the pixel matrix. While the baseline distribution appears random, the noise shows a clear pattern of 2x16 pixels which mirrors the digital building blocks of the ASIC, in which 4 super-pixels share a common clock buffer.

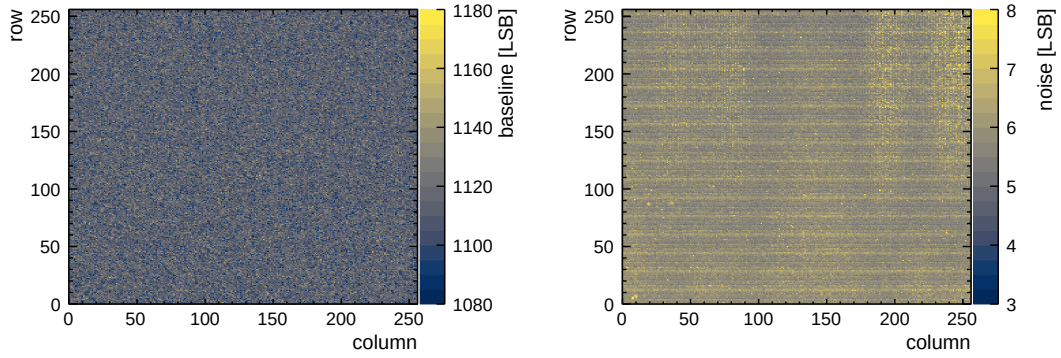
Fig. 4.11 shows the noise for different ASICs before and after assembly with a sensor.[2] The values are given in units of electrons. The derivation of this conversion factor is described only later in Sec. 4.3 but is already used here for reasons of clarity. A small amount of additional noise is introduced by the assembly process but no correlation with the sensor capacitance or the leakage current could be found. The noise appears to be dominated by the intrinsic parameters of the ASIC. This is also supported by the fact that the noise is not affected significantly by a change in the operating temperature, measured on the ASIC, as is shown in Fig. 4.12(a). The noise is however affected by enabling the global Gray counter as can be seen in Fig. 4.12(b). This is relevant as the noise measurement is done in PC & iToT mode while normal data taking is done in ToT & ToA mode. The latter requires the Gray counter while the former does not and deactivates it by default, therefore reducing the noise. For this measurement the Gray counter was enabled manually.

[2] Measurements before assembly are done at the wafer level with a dedicated probe card.



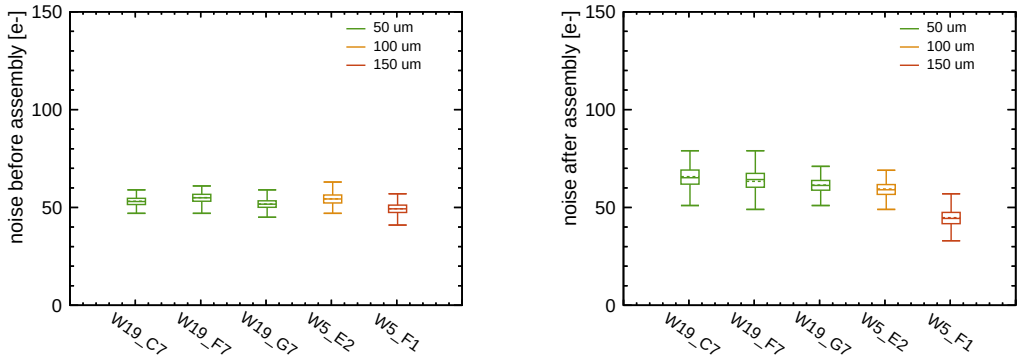
(a) (b)

Figure 4.9: Threshold gain (a) and offset (b) for all diagonal pixels as measured for the various assemblies. See Sec. A for the conventions used in this plot style.



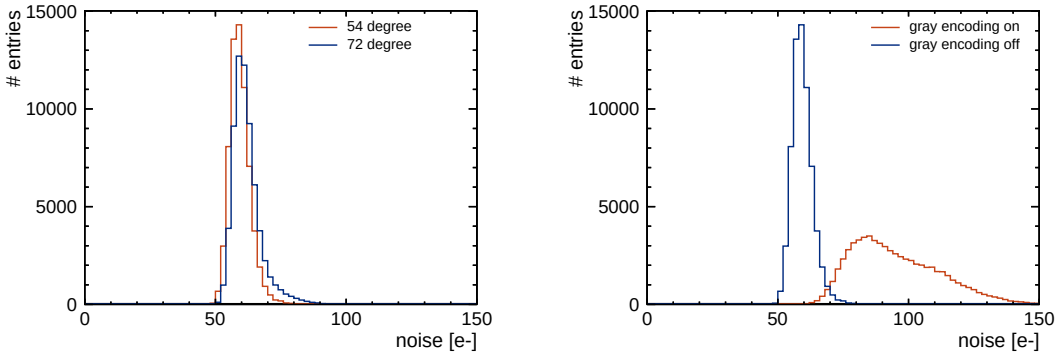
(a) (b)

Figure 4.10: Baseline (a) and noise distribution (b) after equalisation. Assembly shown is W5_E2.



(a) (b)

Figure 4.11: Noise for all pixels of various ASICs before (a) and after (b) assembly. The noise is affected by the assembly process but not significantly by the sensor capacitance.



(a) (b)

Figure 4.12: Noise distribution of all pixels for different operating temperatures (a) and with and without activated Gray encoding (b). Assembly shown is W5_E2.

Based on the measurements described in this section, the nominal threshold values in Tab. 4.1 are chosen to be $6\,\sigma$ from the average baseline, where σ is the quadratic sum of the RMS of the distribution of all baselines for all pixels and the average noise without Gray counter of all pixels.

4.3 Calibration

To reach the best possible performance, each pixel has to be calibrated independently. Due to the non-linear behaviour of the timewalk and the large number of pixels, this is done with a combination of test pulses and x-ray or beam data. While x-ray measurements might be readily available, measurements with particle beams needed for a complete calibration of the time response are rare and valuable. With the chosen approach, only about 100 hits from minimum ionising particles per pixel are required for a full time calibration. An alternative approach, using only beam data, is likely to require about two orders of magnitude more hits per pixel to reach the same precision in the non-linear part.

4.3.1 Energy Calibration

The energy calibration of the assemblies is done in two steps. First, test pulse injection is used to finely sample the non-linear f_{tot} function and obtain a relative calibration from time-over-threshold to voltage. Then, an absolute calibration to energy is extracted via sources of known energy.

Time-Over-Treshold to Voltage Calibration

The calibration of f_{tot} is done via test pulse injection in ToT & ToA mode. Step voltages of various amplitudes are injected via the input capacitor. To avoid crosstalk, only 1 pixel in a 16 by 16 pixel sub-matrix is injected at the same time. Per step, 100 test pulses are injected while mean and standard deviation of the response are recorded as well as the number of detected pulses. Only sample points that detected all 100 pulses are selected for the fit. From that list, the point with the lowest amplitude is removed to limit noise effects. For the stability of the fitting process, it has proven useful to add a threshold estimation. Therefore, the midpoint between the last amplitude that recorded 0 and the first amplitude that recorded 100 pulses is added as threshold value with an uncertainty along the x-axis of the residuals to those two points. The result is fitted with Eq. 4.2. An example fit can be seen in Fig. 4.13.

Later, the inverse function of Eq. 4.2,

$$f_{tot}^{-1}(x) = \frac{a_{tot}t_{tot} + x - b_{tot} + \sqrt{(b_{tot} + a_{tot}t_{tot} - x)^2 + 4a_{tot}c_{tot}}}{2a_{tot}}, \tag{4.4}$$

is used to convert ToT values to voltage. The extracted parameters of the ToT response for all assemblies are shown in Fig. 4.14. From Fig. 4.14(a) one can see that the spread of the ToT gain is systematically smaller for the three assemblies from wafer 19 than for the other assemblies from wafer 5. This is more prominent if one looks at the variation of this value across the matrix

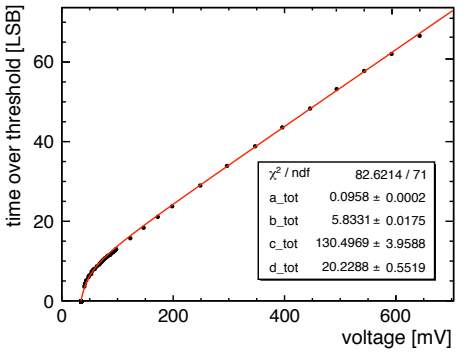


Figure 4.13: Single pixel example fit of the time-over-threshold response measured with test pulses. From assembly W19_C7.

Table 4.2: List of x-ray targets and their corresponding energy used for this study. Energies are weighted with intensities according to [55].

Target Element [-]	Kα Energy [keV]	Attenuation Length (Si,300K) [mm]
Fe	6.40	0.034
Cu	8.04	0.066
Pb	10.55 (Lα)	0.14
Zr	15.77	0.43
In	24.14	1.40

as shown in Fig. 4.15 for 2 examples. While the ASICs from wafer 5 show larger scale structures across the matrix, the ASICs from wafer 19 appear to show only random variations. This is due to a change in the power distribution of the discharge current that was implemented from wafer 13 onwards.

Voltage to Energy Calibration

For the absolute energy calibration, fluorescent x-rays from an electron tube are used with the setup described in Sec. 4.1.4. The x-rays enter from the sensor backside. The target materials used and their relevant energy lines are listed in Tab. 4.2. The values are averaged Kα lines, weighted by their intensity according to [55]. To limit the effect of sub-threshold energy depositions, only single pixel clusters are considered. All assemblies are calibrated by Fe and In targets. Other targets are used for verification only. Per pixel and target, 500 events are required for a per-pixel calibration. Pixels with fewer events are assigned a global average value.

The peaks of the known energy lines are then used to obtain an absolute energy calibration. This procedure is shown in Fig. 4.16 for the global spectra. One would assume a direct proportionality

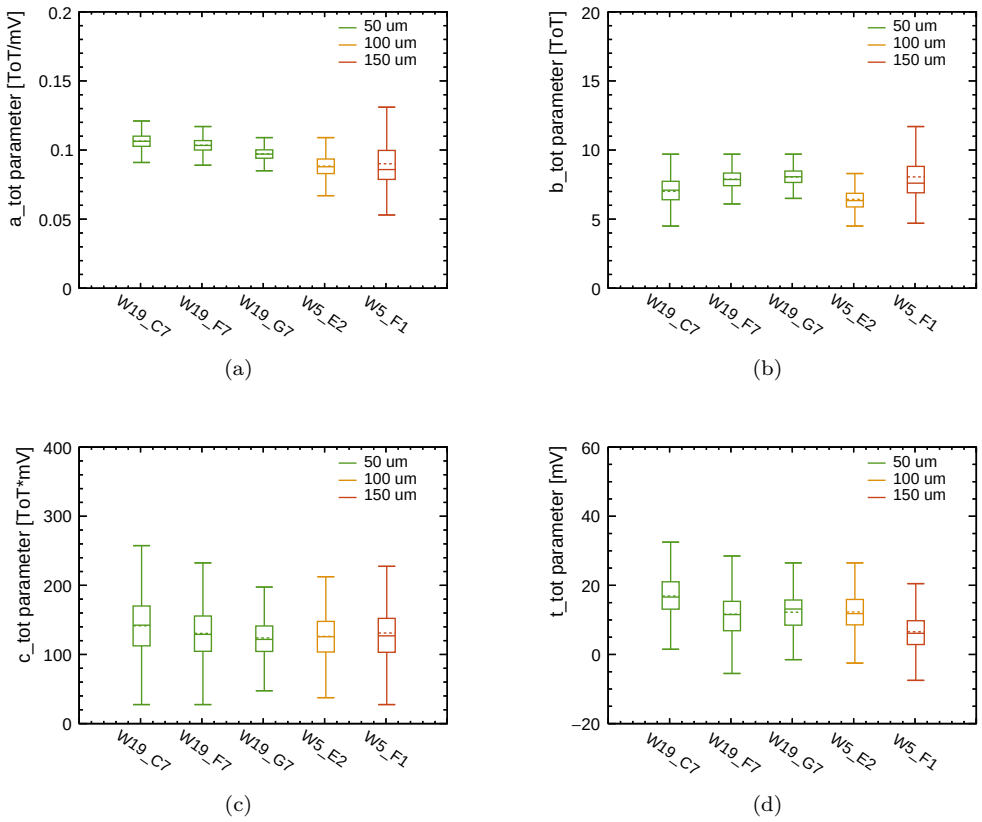


Figure 4.14: ToT calibration parameters for 5 different assemblies.

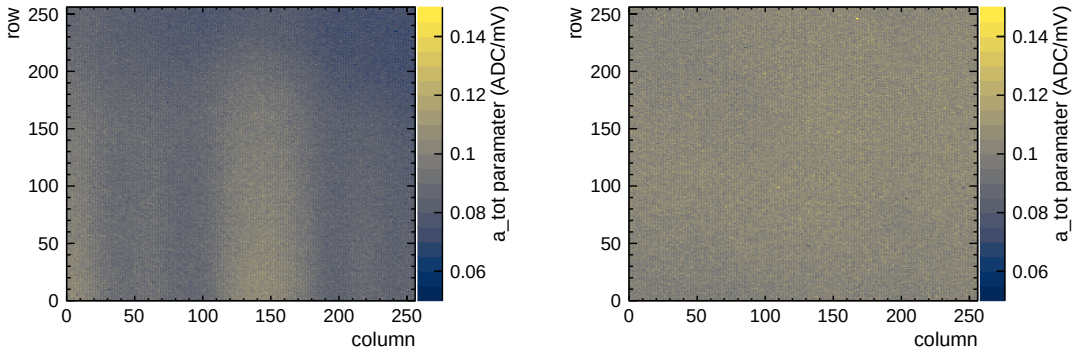


Figure 4.15: The gain in units of ADC per mV for an ASIC from wafer 5 (a) and from wafer 19 (b).

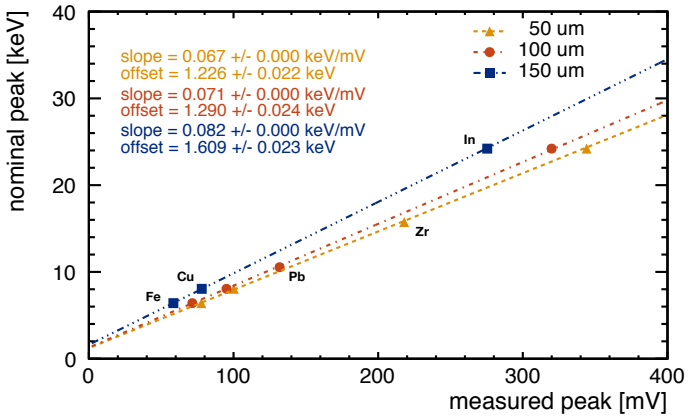
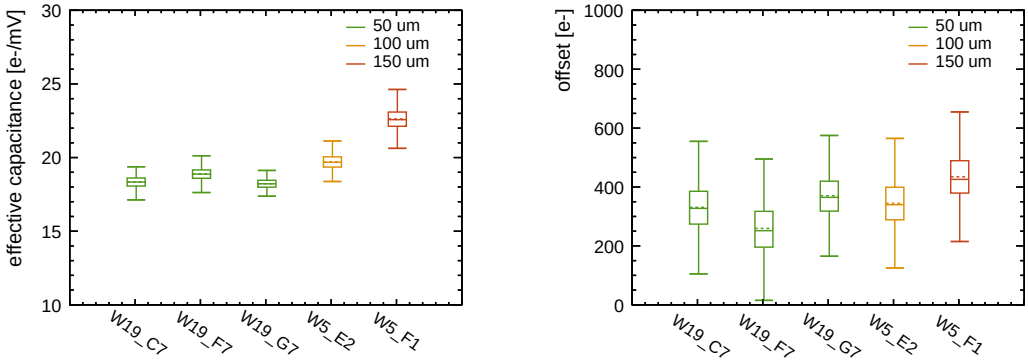


Figure 4.16: Nominal peak position in energy vs. measured peak position in voltage. The three assemblies show a different slope and an offset different from zero.

that is determined by the test pulse input capacitor. The observed behaviour is indeed linear but shows an offset. The same behaviour is also observed in the threshold gain calibration with test pulses described Sec. 4.2.2. From this observation follows, that two parameters are necessary for a full absolute calibration.

$$f_{\text{volt}}(x) = a_{\text{volt}}x + b_{\text{volt}} \tag{4.5}$$

The extracted parameters for all assemblies are shown in Fig. 4.17. The charge scale is obtained by dividing the energy by 3.65 eV [15], the average energy needed to create one free charge carrier pair in silicon.



(a) (b)

Figure 4.17: Voltage calibration parameters for all pixels and 5 different assemblies.

The a_{volt} parameter corresponds to a capacitance in units of [e-/mV] which in turn corresponds to [0.16 fF]. This value should be equal to the value of the capacitance used for the test pulse injection. The design value of 3 fF or 18.75 e-/mV is in good agreement with the values measured

for the 50 µm sensor but is about 20% smaller than the values measured for the 150 µm sensor. Whether this discrepancy results from an underestimation of the x-ray energy scale due to sub-threshold charge sharing or is of different origin will be subject of future studies. The b_{volt} parameter corresponds to an offset in [e-]. When interpreting the absolute value of these two parameters, one should keep in mind that the voltage is measured at the periphery of the ASIC and not at the pixel itself. Possible losses due to e.g. the signal routing or cross talk are not identifiable. Therefore, the voltage scale should only be seen as an intermediate step towards the final calibration scale which is in [e-] and measured with x-ray sources of known energy.

Verification of the Energy Calibration

With a full absolute calibration pixel-by-pixel, the global energy response is analysed again. The result is shown in Fig. 4.18. All peaks are now located at the expected position. It should be pointed out that the Fe and In peaks are used for the calibration while the Cu, Pb and Zr measurements are independent and only used for verification purposes.

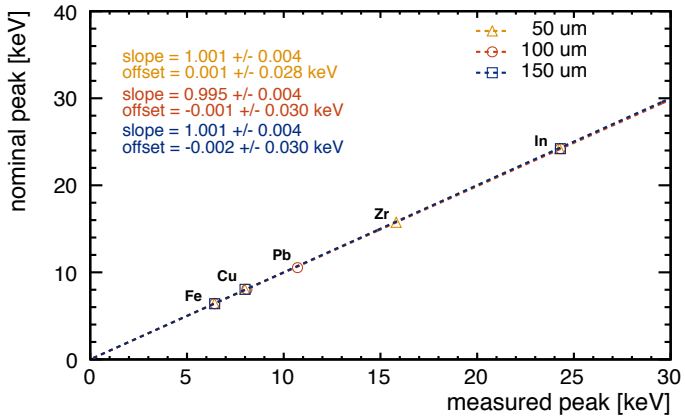
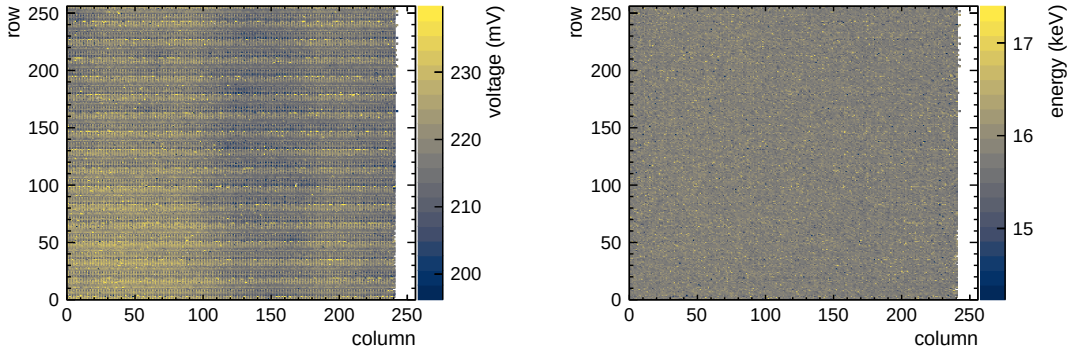


Figure 4.18: Nominal peak energy vs. measured peak position in energy. Fe and In points are used for calibration, while the Cu, Pb and Zr points are added for verification purposes. All slopes do now coincide and originate at zero.

Uniformity of Responses

Fig. 4.19 shows the most probable value of the Zr Kα line, located about midway between the two calibration points, for the absolute calibration and with relative calibration only. The uniformity is much improved when using the absolute calibration and the 16 x 2 pattern has disappeared. Across the matrix, the relative dispersion is reduced from 6.6% to 5.3%. Only pixels with more than 100 events within the region of interest are considered.

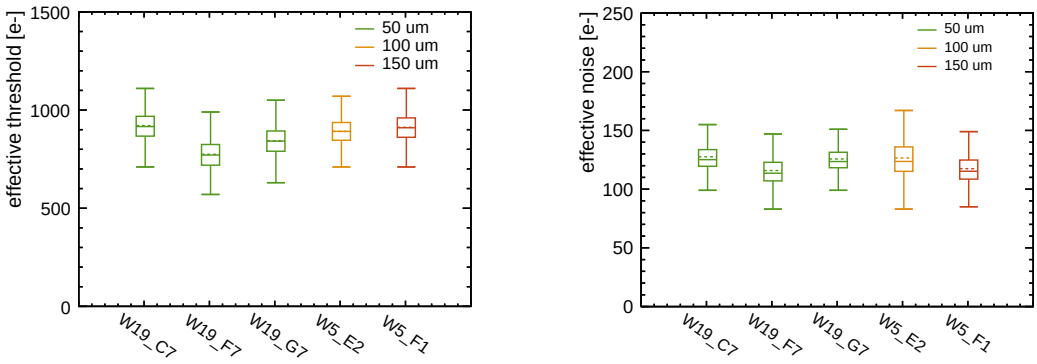


(a) (b)

Figure 4.19: Most probable value of the Zr Kα line before (a) and after (b) the absolute calibration. Assembly shown is W19_C7.

Effective Threshold and Noise

From the test pulse calibration one can also extract the effective threshold by fitting a CDF to the region in which the number of detected test pulses goes from 0 to 100, similarly as it is done in Sec. 4.2.2 for the threshold gain calibration. The result can be seen in Fig. 4.20(a). The average value across all assemblies is around 880 electrons. The effective noise with activated Gray encoding per assembly is shown in Fig. 4.20(b) in units of [e-]. The average value is around 90 electrons. The average noise and threshold values for all assemblies can be found in Tab. 4.3.



(a) (b)

Figure 4.20: Effective threshold and noise values for all pixels and 5 different assemblies.

Table 4.3: Average values for effective threshold and noise parameters for all assemblies.

Chip ID [-]	Threshold RMS [e-]	Avg. Noise Gray Code Off [e-]	Avg. Noise Gray Code On [e-]	Eff. Threshold [e-]
W19_C7	89	66	90	924
W19_F7	72	64	89	815
W19_G7	82	64	91	849
W5_E2	76	61	96	895
W5_F1	79	57	86	913

4.3.2 Time Calibration

Similar to the energy calibration, the time calibration is done in two steps by a combination of test pulses for the non-linear timewalk and beam data to calibrate the delay.

Timewalk Calibration

The timewalk is calibrated via injection of electrical test pulses of various amplitudes. First, a test pulse is injected directly into the in-pixel digital logic, bypassing the discriminator and thus avoiding timewalk. After a fixed delay, a second test pulse is injected into the analog frontend. The ToA values of both pulses are recorded and the time difference is calculated. A shift of the 640 MHz clock of one cycle per double column has to be taken into account. This procedure is repeated 100 times for various amplitudes for the second pulse and the dependence of time difference on the injected amplitude is extracted. The injection scheme is identical to the ToT calibration. The resulting curve is then fitted with Eq. 4.1. An example data curve and fit can be seen in Fig. 4.21.

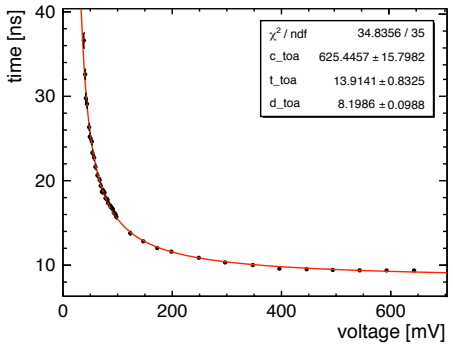


Figure 4.21: Example fit of the timewalk measured with test pulses for a single pixel. From assembly W19_C7.

The extracted parameters of the timewalk response for all assemblies are shown in Fig. 4.22. The ranges of the parameters have not been restricted and therefore also negative values for t_{toa} are possible. It should be noted that it is also possible to do the calibration directly with ToT data, in case the energy calibration is not required.

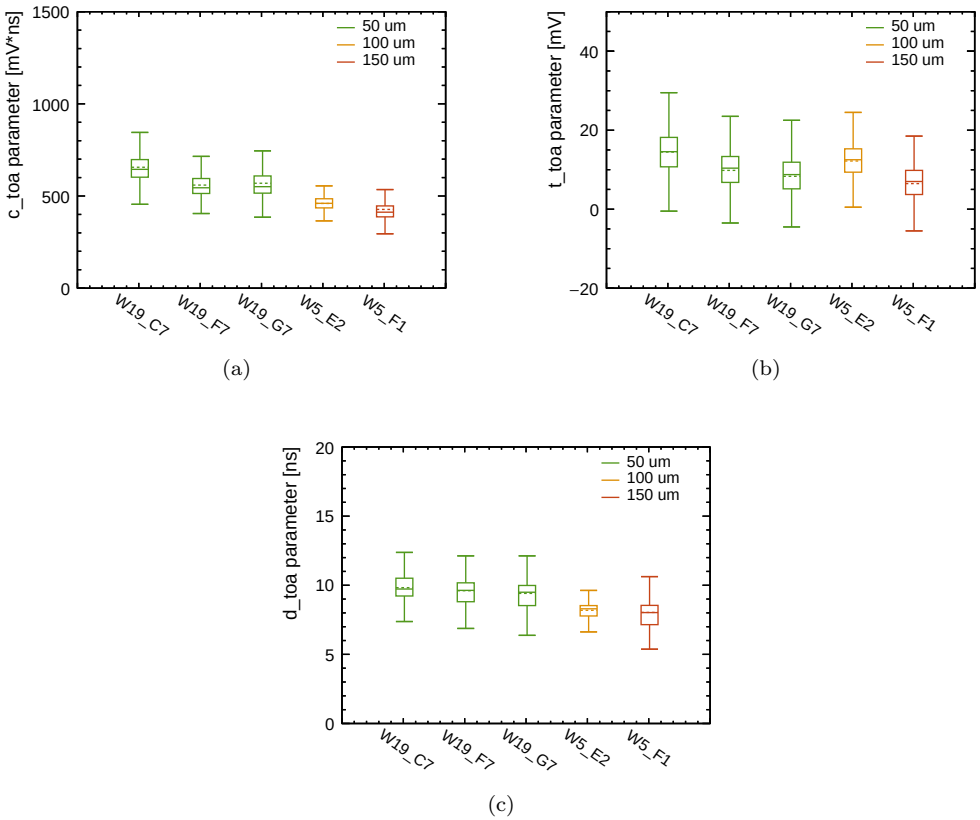


Figure 4.22: ToA calibration parameters for all pixels and 5 different assemblies.

Delay Calibration

The timewalk calibration also includes a delay with the d_{toa} parameter. However, as the test pulse injection uses the same clock as the ToA, any variation in the clock distribution between pixels can not be detected this way. To calibrate these offsets in time, beam tests have been conducted in the setup described in Sec. 4.1.5.

Fig. 4.23 shows the mean time difference between scintillator and DUT across the pixel matrix. These values are used to obtain a pixel by pixel offset calibration. Qualitatively, structures with a period of 16 double columns can be observed and a tendency of later timestamps from higher row numbers. The latter is expected as the clock propagation along one column is expected to take about 1.5 ns between the top and the bottom row. This number is roughly observed but

superposed with more dominant variations. The values from Fig. 4.23 are then used to correct for the offset in time introduced by the clock delay. For each assembly, this is done individually.

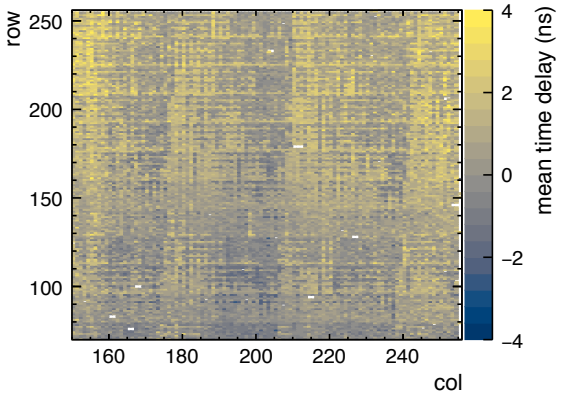


Figure 4.23: Mean time difference between DUT and scintillator across the pixel matrix.

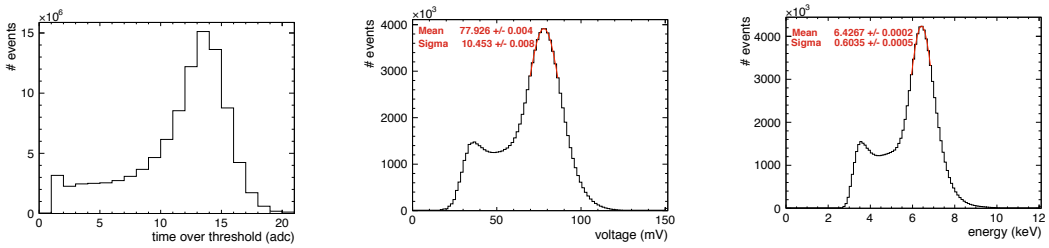
4.4 Performance of Thin Silicon Sensors

4.4.1 Energy Resolution and Accuracy

From a CLIC perspective, the energy calibration is mostly relevant to determine the effective threshold and more general, obtain physical units. For other use cases of Timepix3 however, e.g. spectroscopy applications, the energy resolution is essential.

Fig. 4.24 shows a Fe spectrum in time-over-threshold, voltage and energy. The time-over-threshold spectrum is non-linear and the pixel response is not yet equalised. No fit to the peak is attempted. Converting the spectrum to a voltage, linearises the x scale and results in a Gaussian peak. The absolute calibration improves the resolution and results in a more distinct threshold. This becomes more evident when looking at Fig. 4.25 which shows the resolution as function of x-ray energy. For all energies, the resolution is improved. At the 6.40 keV Fe line, a resolution of 9.3 % is achieved.

The most probable energy deposition for 120 GeV pions in units of [e-] is shown in Fig. 4.26 and compared to Bichsel theory (see Tab. 2.1). The measured values, obtained by calibration with x-rays, appear to overestimate the energy deposition of minimum ionising particles by about 600 e- for 100 µm and 1800 e- for 150 µm. For 50 µm the result matches within ± 50 e-. The most plausible explanation for these discrepancies are sub-threshold charge sharing effects which are expected to differ for continuous energy deposition of MIPs and the point-like photo absorption from x-rays. For clarification of the energy scale, future studies are required. For CLIC however, only the effective threshold is of interest and all three curves from Fig. 4.26 coincide well at low energies. The procedure and values given in Sec. 4.3.1 present therefore a good estimate of the effective threshold.



(a) (b) (c)

Figure 4.24: Recorded fluorescence spectrum of Fe with events from all individual pixels. The ToT spectrum (a) is non-linear, no peak fit is attempted. Calibration to voltage (b) and energy (c) corrects for the non-linearities and improves the resolution by equalising the response across the matrix. Shown assembly is W19_C7.

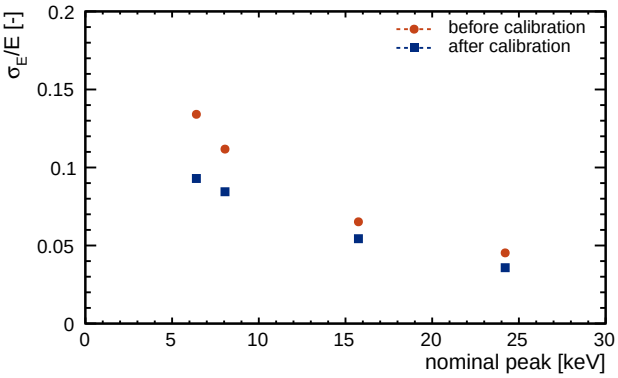


Figure 4.25: Energy resolution before and after absolute calibration for various x-ray energies. Shown assembly is W19_C7.

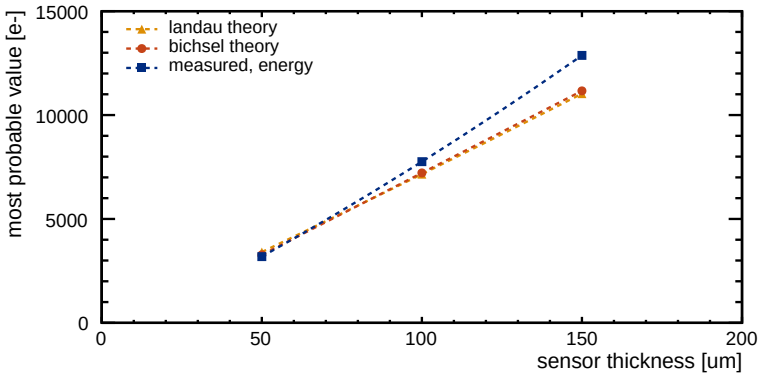


Figure 4.26: Most probable detected charge for beams of minimum ionising particles. The calibration appears to overestimate the energy scale at higher depositions but coincides well with theory at smaller depositions.

4.4.2 Clustersize

The cluster size is an interesting parameter because it indirectly determines the spatial resolution. If no charge sharing is present, i.e. only single pixel clusters are recorded, the spatial resolution will be limited to the binary resolution given by the pixel pitch.

$$\sigma_{\text{spatial,binary}} = \frac{\text{pitch}}{\sqrt{12}} \tag{4.6}$$

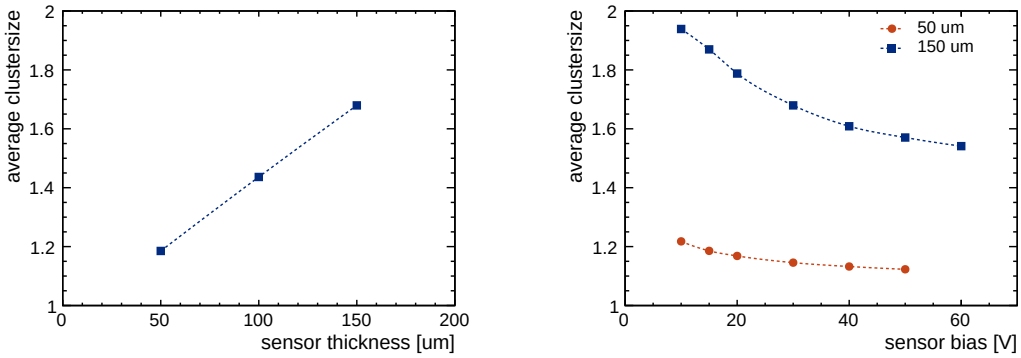
For multi pixel clusters however, the resolution can be improved by measuring the energy deposition of each pixel in a cluster and determining the centre of gravity.

Fig. 4.27(a) shows the average cluster size for MIPs for the 50 µm, 100 µm and 150 µm assemblies at their nominal operating voltage. For 50 µm, the multi pixel clusters are limited to about 20% and the improvement in spatial resolution from energy information is very limited. For 100 µm and 150 µm multi pixel clusters are more common. Fig. 4.27(b) shows the average cluster size for different bias voltages of the 50 µm and 150 µm thick sensors. It can be seen that the cluster size decreases with higher bias voltage due to the larger electric field resulting in a larger drift velocity and hence faster charge collection and reduced diffusion. As will be shown in the next two chapters, this improves the time resolution but degrades the spatial resolution.

4.4.3 Spatial Resolution

The spatial resolution has been investigated by another student and is documented in [54]. In the following, only a short summary of these results will be presented. While the analysis of the data shown in this subsection is not done by the author of this work, the calibration used for the analysis is the same as presented in Sec. 4.3.

Fig. 4.28(a) shows the extracted resolution after energy weighting via eta correction [56] for MIPs



(a) (b)

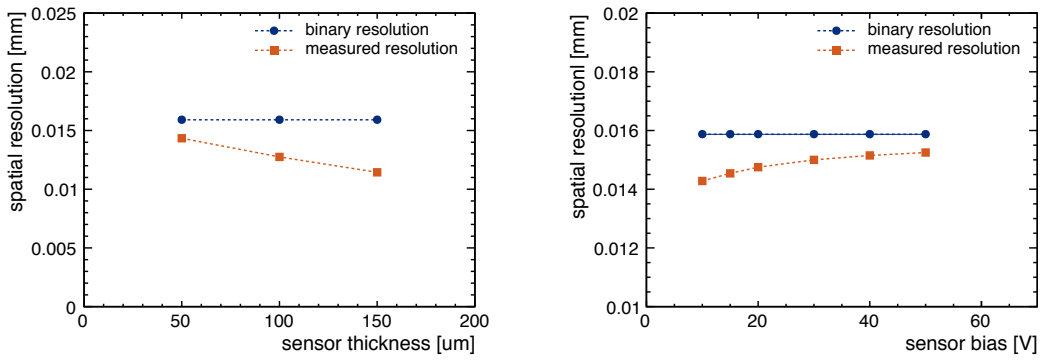
Figure 4.27: (a) The average cluster size for different sensor thicknesses operated at 5 V over depletion. (b) The average cluster size for different bias voltages. Higher bias voltages result in faster charge collection and less diffusion which in turn improves the time resolution but degrades the spatial resolution.

as well as the binary resolution with a 55 µm pixel pitch. As expected from the results shown in the previous chapter, only small improvements can be gained for thin planar sensors. Fig. 4.28(b) shows the result for a 50 µm thick sensor at different bias voltages. A better resolution can be achieved at lower bias voltages due to increased diffusion.

4.4.4 Time Resolution

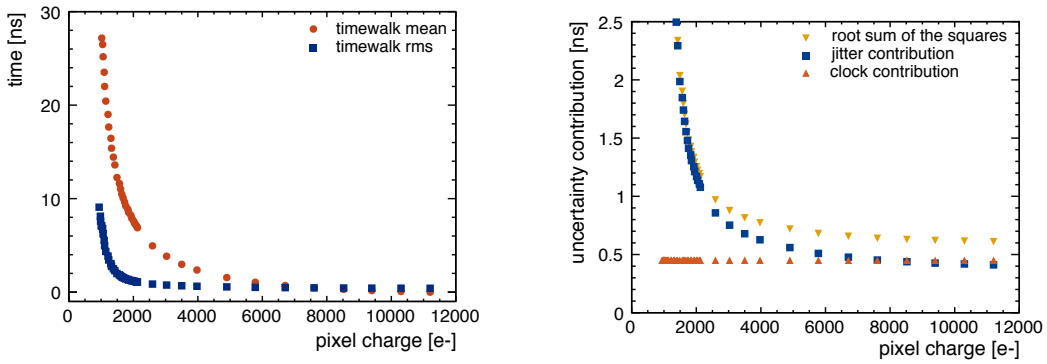
For thin silicon sensors, the effect of timewalk is significantly increased due to the small energy deposition of minimum ionising particles. This can be understood from Fig. 4.29(a) which shows the mean and RMS of the measured timewalk averaged over all pixels. The steepest gradient in the mean timewalk is about 5 ns over 200 e- and the most probable signal in 50 µm silicon (about 3.3 ke-) shows just below 5 ns of timewalk. While the mean timewalk can be corrected with the procedure layed out in Sec. 4.3.2, the RMS of the timewalk is an irreducible contribution to the time resolution. It decreases with increasing pulse height reaching a value of about 420 ps at 10 ke-. It should be pointed out that this is not the limit due to bin size of the fast ToA clock running at 640 MHz. The test pulse injection is not randomised and occurs always at the same clock phase. Rather, this is the jitter of the discriminator and the amplifier due to noise and a finite rise time. This value is in agreement with simulations of the ASIC design [57]. An additional contribution to the time resolution of Timepix3 of at least 0.45 ns from the fine ToA clock period of 1.56 ns divided by $\sqrt{12}$ is expected. The contributions from jitter and finite clock binning as well as the root sum of the squares are shown in Fig. 4.29(b). At 10 ke- the total uncertainty reaches a value of about 620 ps. Effects from the signal development inside the sensor itself as well as non-uniformities in the clock period and imperfect calibration may degrade this resolution.

The total time resolution including sensor contributions is measured in particle beams with the



(a) (b)

Figure 4.28: (a) Spatial resolution reachable following energy weighting for a 55 µm pixel pitch and different sensor thicknesses, compared to the binary resolution. (b) The resolution for a 50 µm thick sensor at different bias voltages. Measured data points are from [54, Fig. 7.10 and Fig. 7.11] with 1.8 µm track resolution unfolded.



(a) (b)

Figure 4.29: Mean and RMS of the measured time delay per pixel from 100 injected test pulses per amplitude (a). While the mean timewalk can be corrected, the RMS of the timewalk due to noise is irreducible and contributes to the final time resolution. (b) shows the expected contributions from jitter and binning of the fast ToA clock. For assembly W19_C7.

setup described in Sec. 4.1.5. From the three obtained timestamps of track, DUT and scintillator, three time differences can be calculated per track. The RMS values of the corresponding histograms then depend on the individual time resolutions σ.

$$\mathrm{RMS}_{t_{\mathrm{dut}}-t_{\mathrm{track}}} = \sqrt{\sigma^2_{\mathrm{dut}} + \sigma^2_{\mathrm{track}}} \tag{4.7}$$

$$\mathrm{RMS}_{t_{\mathrm{dut}}-t_{\mathrm{ref}}} = \sqrt{\sigma^2_{\mathrm{dut}} + \sigma^2_{\mathrm{ref}}} \tag{4.8}$$

$$\mathrm{RMS}_{t_{\mathrm{ref}}-t_{\mathrm{track}}} = \sqrt{\sigma^2_{\mathrm{ref}} + \sigma^2_{\mathrm{track}}} \tag{4.9}$$

Assuming they are not correlated, the time resolution of the scintillator σ_{ref} can be extracted.

$$\sigma_{\mathrm{ref}} = \sqrt{-\frac{1}{2}(\mathrm{RMS}^2_{t_{\mathrm{track}}-t_{\mathrm{dut}}} - \mathrm{RMS}^2_{t_{\mathrm{ref}}-t_{\mathrm{dut}}} - \mathrm{RMS}^2_{t_{\mathrm{ref}}-t_{\mathrm{track}}})} \tag{4.10}$$

Indeed, the three time differences in Eq. 4.10 are not completely independent because all three time measurements depend on the same clock supplied by the Trigger Logic Unit (TLU). However, due to the much finer binning of the TDC that measures the scintillator as well as the combination of multiple timestamps for the track, each with a statistically independent component from time-of-flight and the silicon drift, the correlation is assumed to be negligible. Eq. 4.10 is applied independently to 25 analysed runs, ranging over different sensor thicknesses, bias voltages and calibration procedures, and the average extracted resolution of the scintillator is

$$\sigma_{\mathrm{ref}} = 0.30 \pm 0.04 \text{ ns}.$$

In a similar way, the track time resolution, without any applied calibration, of the six telescope planes can be extracted as

$$\sigma_{\mathrm{track}} = 1.11 \pm 0.03 \text{ ns}.$$

The DUT resolution is defined as the RMS of the residual distribution between scintillator timestamp and DUT with the scintillator resolution subtracted in quadrature.

$$\sigma_{\mathrm{dut}} = \sqrt{\mathrm{RMS}^2_{t_{\mathrm{ref}}-t_{\mathrm{dut}}} - \sigma^2_{\mathrm{ref}}} \tag{4.11}$$

The uncertainty on the RMS is estimated as 0.03 ns from three different 50 µm assemblies measured at 15 V. The propagated uncertainty on the time resolution σ_{dut} is 0.04 ns.

Time Resolution versus Sensor Thickness

Fig. 4.30 shows the time resolution obtained in beam tests with 120 GeV pions for 3 different assemblies. A significant improvement in the resolution can be seen due to the timewalk correction for the 50 µm sensor. The improvement is less prominent for thicker sensors due to the larger energy deposits. After timewalk correction, all sensors yield approximately the same resolution at 5 V over-depletion.

Fig. 4.31 shows the extracted time difference between the scintillator and the DUT as a function of the signal amplitude (a) before and (b) after the applied calibration corrections for a 50 µm thick sensor at nominal bias voltage. The remaining asymmetry at the smallest charge deposits after applying corrections can be attributed to specific pixels in which the timewalk calibration process showed poor performance.

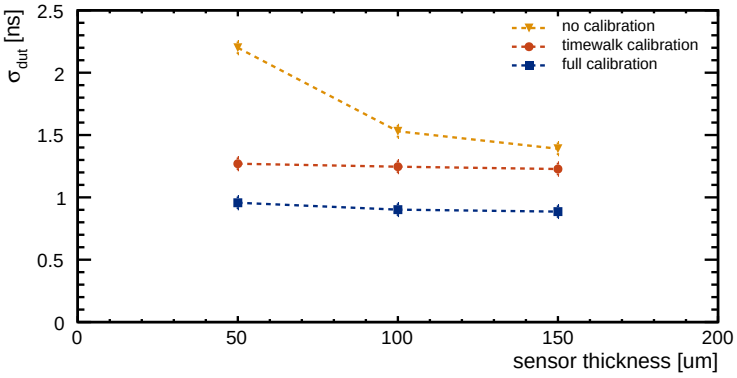
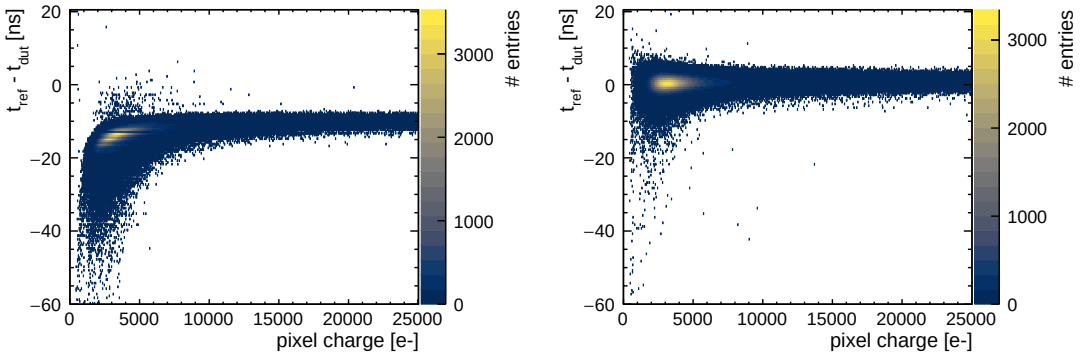


Figure 4.30: Time resolution with and without corrections for different sensor thicknesses measured in beams with 120 GeV pions. The reference timestamp is taken with an organic scintillator read out by a PMT. Its resolution of 0.3 ns is deconvoluted.



(a) (b)

Figure 4.31: Amplitude dependence (timewalk) of the time difference between scintillator and DUT before (a) and after (b) applied corrections for a 50 µm thick sensor.

Fig. 4.32 shows the RMS for the t_{ref} - t_{dut} distribution across the pixel matrix. Here, larger values every 16 rows can be correlated to larger noise values that were observed for these rows (see also Fig. 4.10).

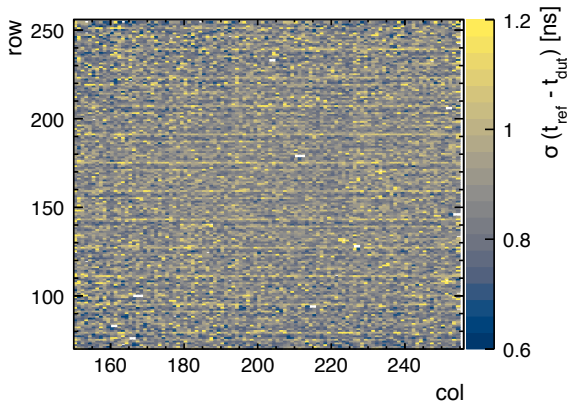


Figure 4.32: RMS of the time difference between DUT and scintillator across the pixel matrix.

Time Resolution versus Bias Voltage

The electric field also influences the time resolution. Fig. 4.33 shows the obtained time resolution for different bias voltages for the 50 µm and 150 µm sensor which deplete at around 10 V and 25 V respectively. The resolution however is improving beyond that point due to the stronger electric field at higher bias voltages. This in turn increases the drift velocity which results in a faster and more uniform signal creation on the CSA input and, owing to reduced diffusion, a trend towards larger pixel charges. At the same over-voltage, a slightly better resolution can also be observed for the 150 µm sensor which has a larger average charge deposition. Both observations are in good agreement with the expectations from Fig. 4.29(b). The best value for the time resolution reached in this study is

$$\sigma_{\mathrm{dut}} = 0.72 \pm 0.04 \text{ ns}$$

for a 150 µm thick sensor at 35 V over-depletion.

Fig. 4.34 shows the distributions of t_{ref} - t_{dut} for the highest bias voltages of two sensor thicknesses at the various calibration stages. The uncalibrated histogram is highly skewed due to timewalk and shows multiple peaks as pixel by pixel variations are not yet compensated. After applying both calibration steps, the multiple peak structures disappear and the histograms become Gaussian.

Concerning the uncertainty contribution to the signal formation inside the sensors, it can be concluded that for a pixel pitch of 55 µm and thin substrates, they are well below 0.72 ns. Further quantification would require a more precise time measurement than available with Timepix3.

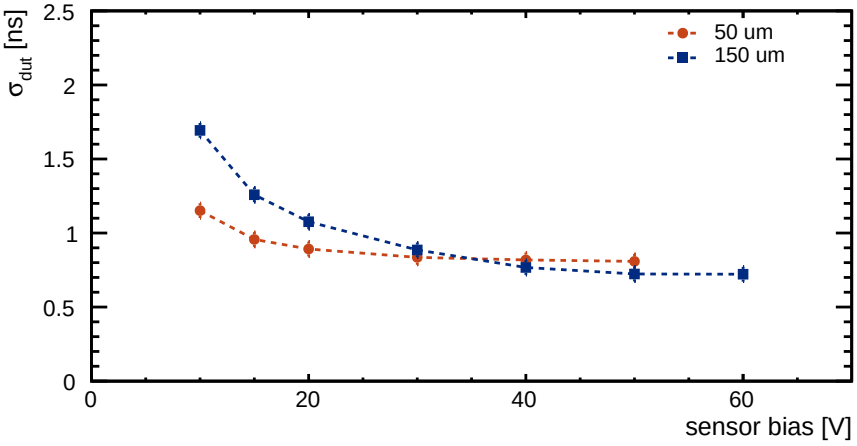
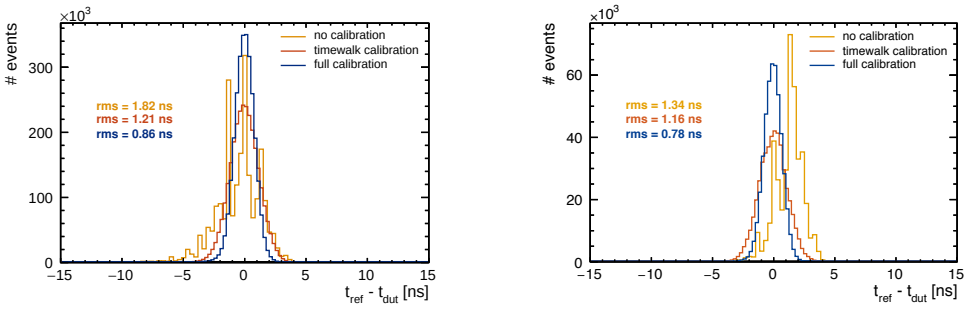


Figure 4.33: Time resolution at different bias voltage. An improvement beyond full depletion can be seen due to stronger electric fields.



(a) (b)

Figure 4.34: Example histogrammes for (a) a 50 µm thick sensor at 50 V bias voltage and (b) 150 µm thick sensor at 60 V bias voltage for different calibration stages. From the RMS values, 0.3 ns are unfolded to obtain the time resolution of the DUT.

4.5 Conclusions for the CLIC Vertex Detector and Beyond

The combined requirements for the CLIC vertex detector are difficult to meet. The fast timing requirements call for fully depleted sensors. However, to meet the material budget, thin sensors have to be used that due to the short drift length lack a significant diffused signal component.

For the spatial resolution, the results from Sec. 4.4.3 show that with 50 µm thin sensors, significant improvements beyond the binary resolution cannot easily be achieved. Therefore, in order to obtain a 3 µm spatial resolution as required for CLIC, one would need a pixel pitch of around 11 µm. It appears unrealistic that bump bonding at this pitch is technically feasible.[3] This means that if thin planar sensors are to be used for the CLIC vertex detector, new fine-pitch interconnect processes have to be found. Here, indium bumps could reach the required pitch (see e.g. [20]). However, the pixel pitch is not only limited by the interconnect density but also by the feature size in 65 nm technology. Adopting e.g. 14 nm technology that is already standard in consumer electronics for several years would most likely solve this problem. Alternatively, enhanced lateral drift sensors (ELAD) [58] try to increase charge sharing by shaping the electric field with additional implants.

For the time resolution, it was shown in Sec. 4.4.4 that 50 µm thin sensors comfortably meet the 5 ns timing requirement of the CLIC vertex detector, provided the timewalk is adequately corrected. For a CLIC specific ASIC with similar noise performance and discharge rate, a 5 bit energy measurement with a total dynamic range of about twice the most probable signal is sufficient to correct the timewalk to an uncertainty below 5 ns for all hits. The frequency of the fine ToA clock could also be reduced, if for example the power budget should demand it. For the Timepix3 ASIC and pixel readouts in general, another conclusion can be drawn from these results: To reach the design values in performance, a calibration method suitable for the high granularity is a key element. For the Timepix3 ASIC, a careful pixel-by-pixel calibration allows to reach close to the ASICs intrinsic limits in time resolution. As will be shown in Chapter 6, this is not always the case.

For the CLIC vertex detector, an alternative technology approach considered is the use of monolithic technologies in which the frontend electronics and the sensing diodes are implemented on the same substrate. The main challenge of these technologies is to shield the electronics from the bias voltage that is necessary to deplete the sensor. Often, this is only possible to a limited extend. As a result many of these technologies have a low electric field and a limited depleted volume. Compared to hybrid planar pixel assemblies, this typically improves the spatial resolution but deteriorates the time resolution. The results on all technology prototypes for the CLIC vertex detector are summarised in [50]. So far none of the above mentioned technologies have fully met the combined requirements of spatial resolution, time resolution and material budget.

[3] First trials with 25 µm have started with reasonable success.

Chapter 5

Electrical Characterisation of Silicon Pad Sensors

To reach the desired jet energy resolution, CLIC will require a highly granular and compact calorimeter as described in Sec. 3.2. The electromagnetic calorimeter of CLIC will employ roughly 2500 m^2 of silicon pad sensors with a pad area of 0.25 cm^2. For HGCAL, about 600 m^2 are foreseen with a slightly coarser granularity of 0.5 cm^2 and 1.2 cm^2.

In the R&D and prototyping phases of both projects, different silicon sensor design features have to be evaluated. The most interesting characteristics are typically the current vs. voltage (IV) and capacitance vs. voltage (CV) behaviour. From these measurements, important parameters for other parts of the full system can be derived, e.g. the leakage current that needs to be compensated by the power supplies or the capacitive load on the front-end electronics. Later, in the production phase, several thousand sensors have to undergo quality control to identify defective pads. Due to the large number of pads that have to be characterised, it is essential to use a fast and automated testing process.

In this chapter, the development and commissioning of a testing system designed for this purpose is described. It is composed of two independent components, an active switching matrix with 512 input channels that holds all controls, and a passive probe card that connects to the sensor. Section 5.1 reviews the silicon sensor design for HGCAL and highlights the various features of the prototype campaign. Section 5.2 gives a technical description of the system. Sections 5.3 and 5.4 show results on the system's performance and its limitations for current and capacitance measurements. Measurements of HGCAL prototype sensors are presented. Finally, section 5.5 summarises the results and draws conclusions for the CLIC calorimeter.

Some parts of this chapter have also been published by the author in [41] and [59] and are quoted verbatim.

5.1 Silicon Sensors for HGCal

The HGCAL silicon sensors are designed in a hexagonal geometry to maximise the used area on the circular wafers and therefore minimise cost. It is foreseen to use 8-inch wafers, in which case roughly 27000 sensors will be required. For prototyping, both 6- and 8-inch wafers are used. An annotated design of a prototype sensor is shown in Fig. 5.1.

The active sensor thickness is adapted to the expected radiation dose and varies between 120, 200 and 300 µm. The pad capacitance should not exceed 50 pF for all sensor thicknesses. Thinner sensors are therefore equipped with smaller pads. A granularity of 0.5 cm^2 for the 120 µm and 1.2 cm^2 for 200 and 300 µm thick sensors is used. For mounting purposes, the edges of the hexagonal sensors are truncated ('mouse bites').

One of the key aspects of the sensors is the high-voltage sustainability to mitigate radiation damage. The goal is to achieve a breakdown voltage above 1 kV. The metal-overhang and guard ring periphery are optimised for that purpose. A double guard ring structure is used, with the outer guard ring floating and the inner guard ring grounded, to smoothen the field gradient.

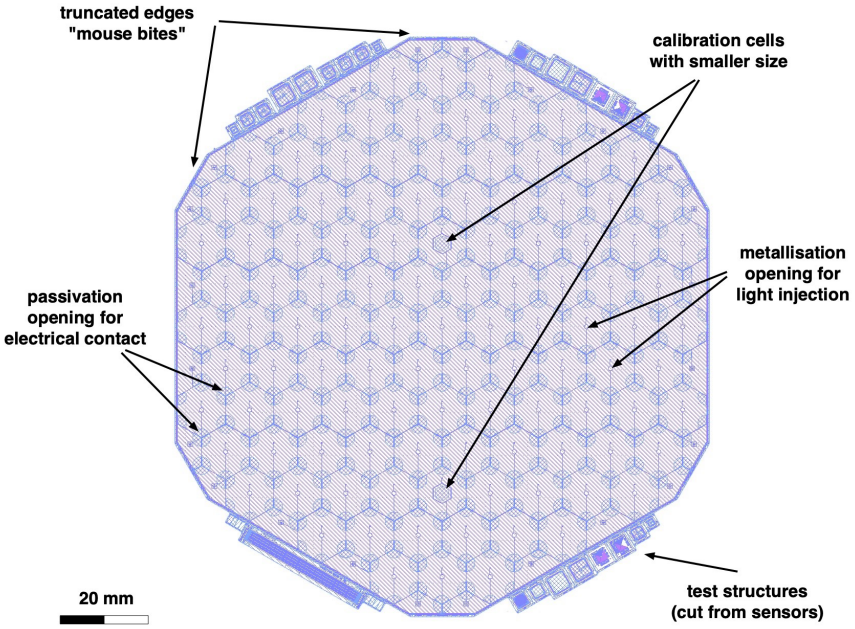


Figure 5.1: The design of a prototype HGCAL sensor.

As mentioned in Sec. 3.3.3, each sensor will be calibrated with MIPs. With irradiation, the leakage current will increase and eventually surpass the capacitance as the dominant noise contribution. To keep an acceptable S/N for MIPs up to an integrated luminosity of 3000 fb^{-1}, it is foreseen to use a few pads with smaller area than the regular ones on each sensor. The smaller area at unchanged thickness will reduce the noise contributions from capacitance and leakage current in these pads, so that they should still be sensitive to single MIPs up the highest

Table 5.1: List of sensor geometries used in this work. See text for explanations.

Common Name [-]	Size [in]	Thickness [µm]	No. Pads [-]	Polarity [-]	p-stop [-]	Backside
6in 135ch n-type	6	300	135	n-type	-	STD
6in 135ch p-type common	6	300	135	p-type	common	STD
6in 135ch p-type atoll	6	300	135	p-type	atoll	STD
6in 239ch n-type	6	120	239	n-type	-	DD
6in 239ch p-type common	6	120	239	p-type	common	DD
6in 239ch p-type atoll	6	120	239	p-type	atoll	DD
8in 271ch p-type atoll	8	300	271	p-type	atoll	thinning
8in 271ch p-type atoll	8	200	271	p-type	atoll	thinning
8in 271ch p-type atoll	8	120	271	p-type	atoll	epitaxial

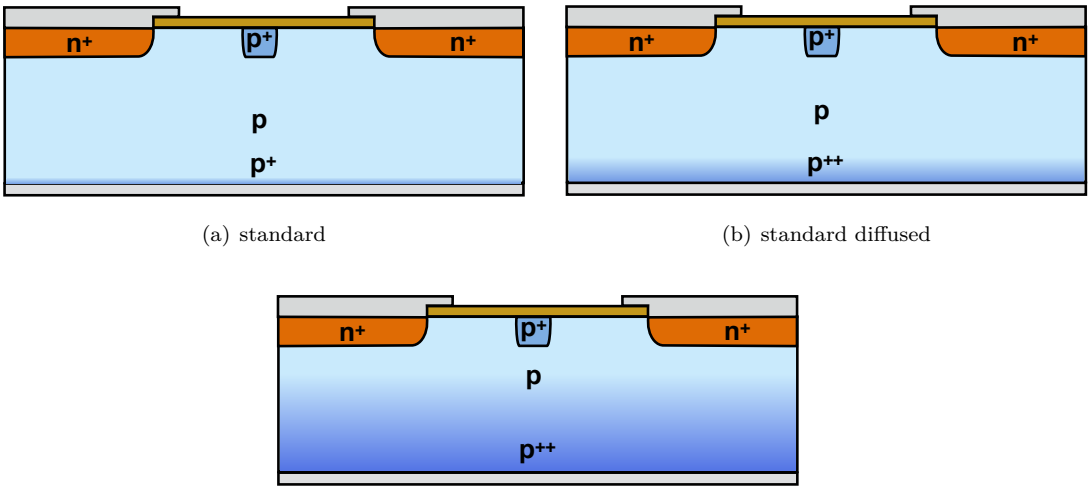
expected dose.

During the prototyping campaign of HGCal, several different sizes, polarities, thicknesses and geometries have been tested. An overview of all tested sensor types for this work can be found in Tab. 5.1.[1] The pad sizes for the prototype sensors are roughly 0.5 cm^2 for the 120 µm and 1.1 cm^2 for the 200 and 300 µm sensors but the exact values vary slightly. Two wafer sizes are used: 6 inch and 8 inch. While the 6 inch sensors are produced with a full wafer lithography, the 8 inch sensors of this first production series are produced by a stepping technique of smaller rectangles that are then repeated to cover the full wafer area. The latter creates some limitations on the sensor design due to symmetry. A full wafer lithography is foreseen for the series production of the 8 inch sensors as well.

5.1.1 Active Thickness

Several techniques to reach the active thickness are explored. For the 6 inch wafers, all sensors have a physical thickness of 320 µm. To reach the active thickness, a layer with high dopant concentration is introduced at the backside to create a robust backside implant. In the case of the 300 µm sensors, this is done by a process called *standard diffusion (STD)*, that creates a 20 µm thick backside implant. For the 120 µm sensors, the backside implant is much thicker and is reached by a different process called *deep diffusion (DD)*. The exact nature of both processes is kept confidential by the manufacturer. For the 8 inch wafers, the thickness is reached either by physical thinning of the backside followed by a standard backside processing or by epitaxial growth. In the latter, a layer of high resistive silicon is grown on a low resistive carrier wafer. The result is a thin active layer of high resistivity silicon. These processes are illustrated in Fig. 5.2.

[1] This is however not a complete list of prototype sensors tested for HGCAL.



(a) standard

(b) standard diffused

(c) deep diffused

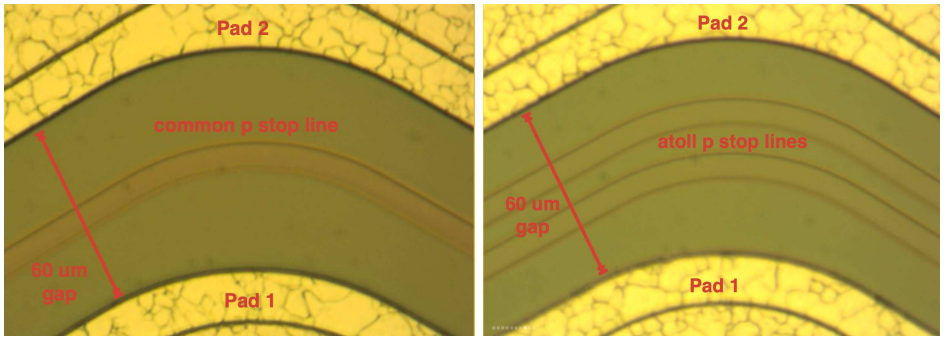
Figure 5.2: Different backside treatment illustrated on the example of p-type sensors. While standard sensors have a thin backside implant of about 1 µm (a), standard diffused (STD) sensors have a ≈ 20 µm thick backside implantation (b) and deep diffused (DD) sensors create a heavily reduced active thickness by using a deep doping gradient of up to 200 µm (c).

5.1.2 Polarity and Different Geometries

Both p- and n-type sensors are used during the prototyping phase. For p-type sensors, the frontside n-implants have to be isolated by additional p-implants due to charge-up effects on the Si-SiO_2 interface. For this, two p-stop geometries are explored: 'common' p-stops which consist of one connected grid that permeates the full sensor and 'atoll' p-stops which consist of independent implants around each pad. In both cases, the implants are electrically floating. The two geometry versions used are shown in Fig. 5.3.

The 6in 239ch design has an additional feature of so-called *jumper pads* that are shown in Fig. 5.4. These are extensions of the metal pad (and collection implant underneath) that could simplify the module design but possibly at the cost of high voltage instability and additional cross talk.

To explore the influence of the inter-pad parameters, different inter-pad gap distances are used. In the case of the 6in 135ch design, there are four quadrants, each with a different implant-to-implant gap distance (20, 40, 60 and 80 µm). For the 6in 239ch design, two halves of 30 and 50 µm gap distance are used and for the 8in designs there is only one gap distance of 50 µm, due to the production process. The layouts of the three sensor geometries are shown in Fig. 5.5.



(a) common (b) atoll

Figure 5.3: Microscope images of the different p-stop geometries used for p-type sensors. (a) The common p-stop geometry consists of a grid of p implants that links through the full sensors. (b) The atoll p-stop geometry has independent p implants around each sensor pad. In both cases, all p-stops are electrically floating.

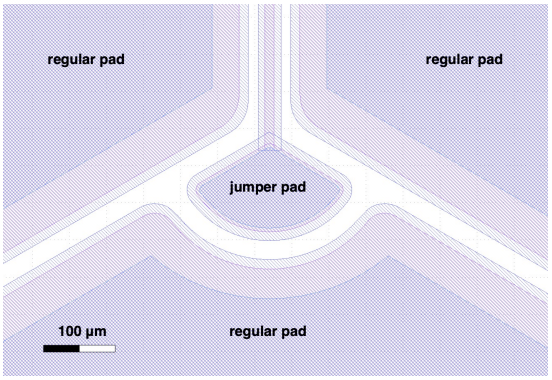
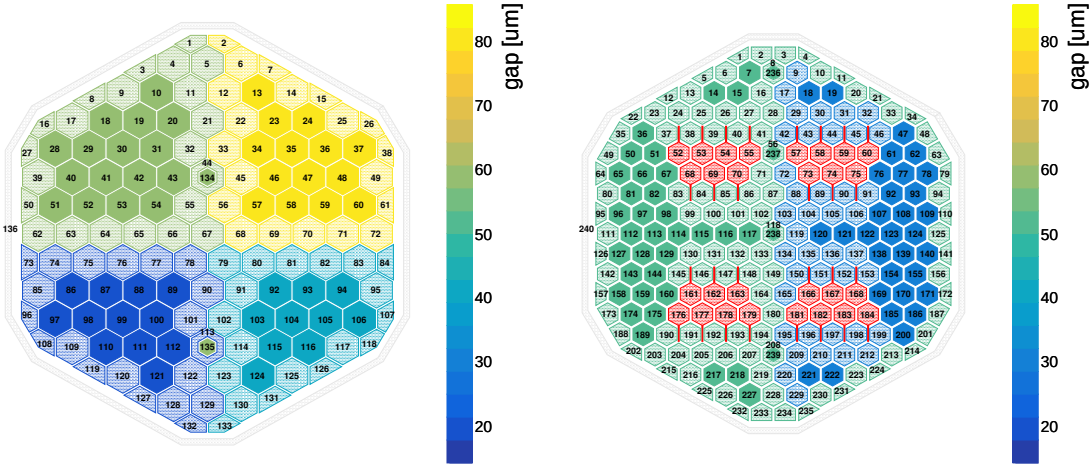
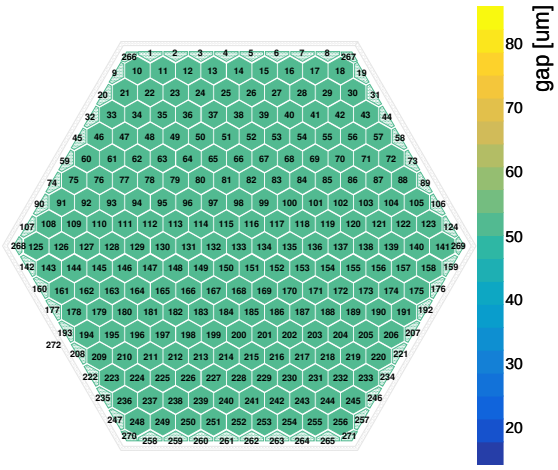


Figure 5.4: The jumper geometry used for selective pads on the 6in 239ch sensors.



(a) 6in 135ch design

(b) 6in 239ch design



(c) 8in 271ch design

Figure 5.5: Layout of the three sensor geometries tested in this work. Pads with a single interpad gap distance to all neighbours are in full colour while those with mixed gaps are shaded. Pads with jumper geometries are shown in red.

5.2 Measurement Setup

When characterising the sensors, it is essential that all pads are biased, similar to the operating conditions at the experiment. Floating pads, i.e. pads that are not biased properly, change the electric field configuration inside the sensor and therefore its properties. This is especially true for the guard ring area that surrounds the pad matrix. It is generally the most prone to breakdown effects caused by high-field regions. Due to the lack of a common bias structure, a probe-card based system that connects to all pads on the sensor simultaneously is required.

The testing setup should be capable of measuring multiple sensor layouts with up to 500 input channels, approximately corresponding to the number of 0.5 cm^2 pads on an 8-inch wafer. Currents ranging from hundreds of pA to 10 µA and capacitances up to 100 pF have to be measured accurately and in a reasonable measurement time. Ideally, the system would also be compact, cheap and easily portable to different laboratories. To fulfil these requirements, a solution with a multi-channel multiplexer and one set of precision instruments was selected. Commercially available instruments were selected for CV and IV measurements. For the switching system however, no suitable commercial solution was identified. While several commercial switching systems offer the required precision and long term stability, the number of channels is limited. An additional disadvantage of this type of system is related to the input connectivity requiring hundreds of high quality shielded cables to be connected to the probe card. Due to limitations of existing commercial systems, a dedicated switching matrix together with a probe card has been developed. It fulfils all requirements of performance, compactness and cost. The system is named ARRAY (switching mAtRix pRobe cArd sYstem).

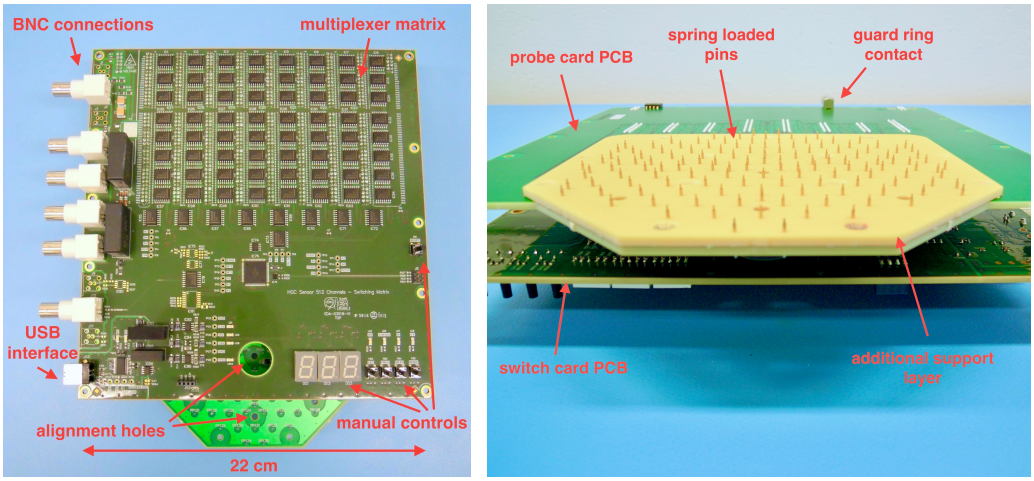
5.2.1 Overview of the ARRAY Sensor Testing System

The switching matrix is designed as a plug-in printed circuit board (PCB), called switch card. It is located directly on top of the probe card. Fig. 5.6 shows the system from (a) the top and (b) the bottom. The switch card is essentially a large array of multiplexers (MUXes) that controls the measurement, while the probe card is a passive device that provides the connectivity to the sensor. Dedicated probe cards are designed for each sensor geometry. Spring-loaded pins are used to connect the probe card to sensor pads.

The two cards are mechanically attached to a probe station, shown in Fig. 5.7(a), which itself is inside a light-tight box. The sensor is held in place by a vacuum chuck. Viewing holes inside the cards allow for x/y alignment of the probe card relative to the sensor. After alignment, the cards are lowered to contact the sensor. The contacting process of a single spring-loaded pin is shown in Fig. 5.7(b).

A simplified circuit diagram of the full system is shown in Fig. 5.8. Firmware and design files have been published under the CERN Open Hardware license and can be found in [60]. The production cost of the switch card are around 2500 CHF per PCB.[2] The cost of the probe card is typically about a factor two less than for the switch card but depends on the required number of spring loaded pins. The full system is roughly the size of an A4 paper stack.

[2]Prices refer to a production batch of five units including components, assembly, tooling and production in Switzerland in 2018.



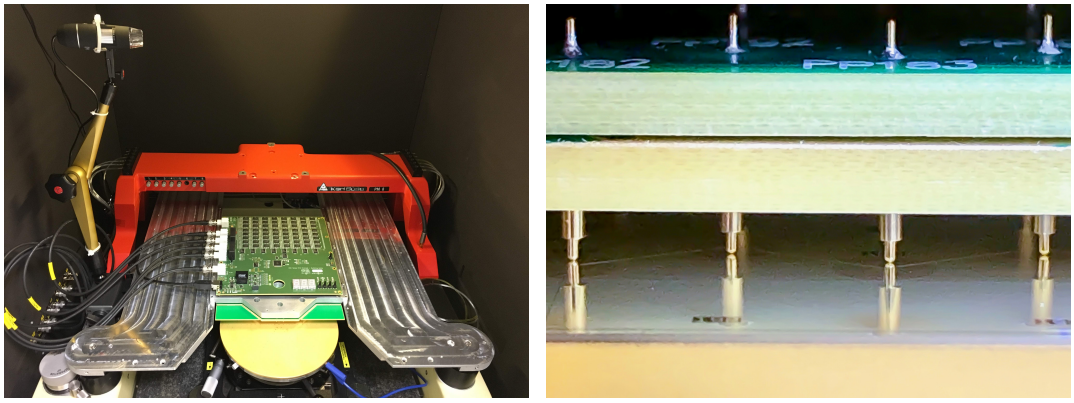
(a) (b)

Figure 5.6: Photographs of the assembled two card system, viewed from (a) the top and (b) the bottom.

5.2.2 The Probe Card

The probe card consists of an array of spring-loaded pins with 1.4 mm travel range that are used to contact all individual sensor pads (see Fig. 5.6(b)). The travel range avoids damaging the pads during contact and ensures a uniform contact over the full sensor area. The chosen pins are 4.8 mm long and have a rounded tip with a radius of 250 µm. The contact marks on the sensor have about twice the diameter of the marks from a conventional prober needle with 12 µm tip. To enable a good electrical contact to the aluminium pads on the sensor, a force of about 25 g per pin has to be applied. This force has to be applied via the mechanical integration into the probe station. Due to the spring-loaded nature of the pins, the positioning accuracy is limited to about ± 200 µm. For the guard ring with typical contact pads of few tens of µm, this precision is not sufficient. Therefore, the probe card provides an additional input that allows the use of an external probehead for the guard ring connection with needles.

The pins, as well as the guard ring input, are then routed to the connectors on the back side of the probe card, which plugs into the switch card. The routing avoids parallel tracks and maximises clearance in order to limit parasitic capacitances. Test capacitances are added to the probe card to allow for system calibration and cross checks. Humidity and temperature sensors are integrated into the card to monitor environmental conditions. For different sensor geometries, different probe cards are used, yet all cards are based on the same concept. The probe card is mechanically fixed to a steel mounting frame. The base PCB is 2.2 mm thick. In the contact area of the pins, an additional support layer of 2 mm thickness is used to provide mechanical support and to facilitate the pin placement. An array of standoffs between probe card and mounting frame prevents any bending of the PCB.



(a) (b)

Figure 5.7: (a) Photograph of the system integrated into the probestation. (b) A spring loaded pin contacting the sensor pad.

5.2.3 The Switch Card

The signal originating from the probe card is routed via a 512 channel multiplexer (MUX512) towards the readout instruments. To make the system compact and ensure a long lifetime, solid state multiplexers are used. Due to the limited number of channels available in such devices and in order to minimise the parasitic capacitance, the MUX512 is implemented in a hierarchical way, with three levels of 8-to-1 multiplexers. The output multiplexer (MUXOUT) enables switching between the IV and CV output circuits. Multiplexers of type MAX328 [61] have been selected due to their small series resistance of typically 1.5 kΩ and their leakage current of only a few pA. An additional 10 kΩ resistor offers protection to the multiplexers on the switch card in case the potential on the low side of the diode would drift too far away from ground potential. The total value of R_{series} from the multiplexers and this resistor is about 16 kΩ. Additionally, each channel can be shorted to ground using solid state switches to ensure proper bias for cells which are not being measured.

The switch card provides six coaxial BNC connections to a source measurement unit (SMU), LCR meter (LCR) and ammeter (AM). For the IV circuit, the output of MUXOUT is connected directly to the measurement device and then returned to ground. In the case of the CV circuit, a high value resistor R_{bias} of 10 MΩ to ground is added to ensure proper bias. This resistor raises the impedance in the parallel circuit and allows for a measurement of the capacitance of a single pad, rather than all pads on the sensor (see Sec. 5.4). Two 1 µF capacitors C_{dec} decouple the LCR meter from the DC voltage.

The high voltage (HV) is connected to the switch card, routed via the filter network of R_{HV} and C_{filter} to the probe card and applied to the sensor back side via the chuck. On the switch card, a microprocessor controls on-board components and provides the user interface. The settings of the switch card can be changed manually (using buttons) or remotely via a software interface. The switching scheme is hard-coded to always short all pads to ground first before switching the pad-under-test. In this way, switching under large loads is avoided. Since the resistance of the

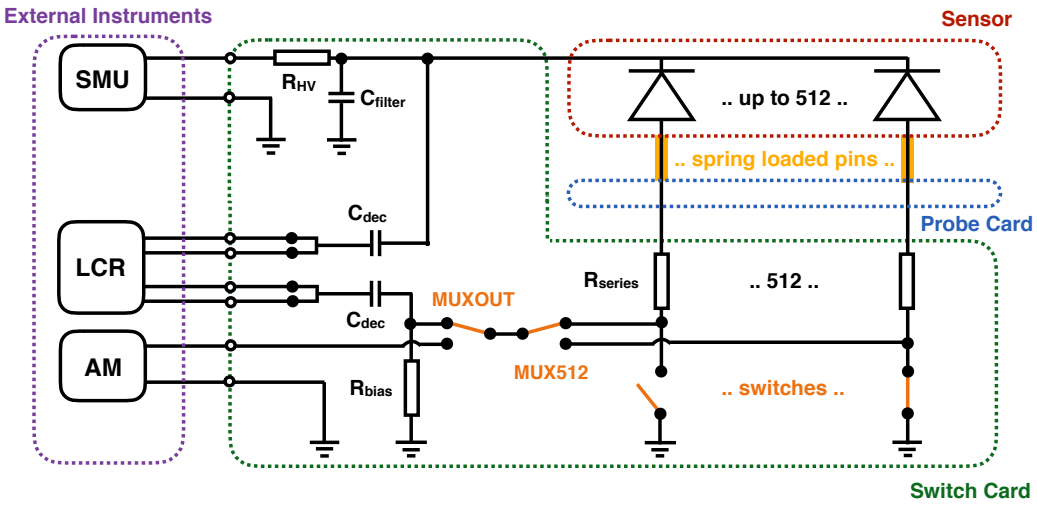


Figure 5.8: A simplified circuit diagram of the system. For CV measurements, the pad-under-test returns to ground via a variable bias resistor, and for IV measurements via the ammeter. All other pads are directly shorted to ground. A hard-coded switching scheme guarantees that all pads are biased at all times. Acronyms are introduced in Sec. 5.2.3.

diodes is typically O(100 GΩ), the change in bias voltage is negligible.

5.2.4 Data Acquisition and Controls

The switch card is controlled via a USB serial interface. Two libraries have been developed to interface the hardware, one in Python and one in LabVIEW™. They can be found in the design repository [60]. These libraries also provide the data acquisition and monitoring with controls for a few common measurement instruments.

The total switching time between channels is around 100 ms. The measurement of the current or capacitance per channel takes between 1 s and 2 s, depending on the external devices and integration times that are used. An additional delay of a few seconds is used per voltage step to account for the time constant introduced by C_{dec} and R_{bias}. In the author's setup, a typical IV or CV scan for a sensor with about 150 pads takes about one hour for 15 voltage steps. This is at least an order of magnitude faster than a semi-automatic probe station with multiple probe heads could achieve, due to the needed settling time after a change in voltage.

5.3 Current Measurement

The total current of the system drawn from a high voltage power supply (here Keithley 2410) at 1000 V as a function of time is shown in Fig. 5.9 for the bare switch card, for the assembly

of switch card and probe card as well as for the assembled system when the LCR meter is disconnected. After charging of the capacitors, a small current of about 10 nA remains for the bare switch card and for the assembled two-card system. This current disappears if the LCR meter is disconnected and can therefore be attributed to the leakage current of the C_{dec} on the switch card. This is in agreement with standalone measurements of the used capacitors.

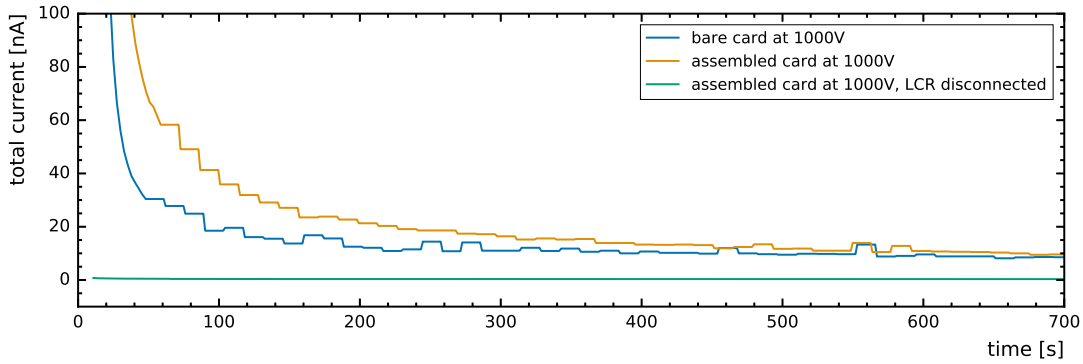


Figure 5.9: The system's total current over time at 1 kV in three partial configurations. A small leakage via the HV decoupling capacitors can be observed.

The per-cell current is measured with an external ammeter (here Keithley 6487). The leakage current per channel is less than 10 pA at 1000 V and is shown in Fig. 5.10. These measurements as well as all other measurements in this work have been recorded at room temperature and a humidity of about 50%.

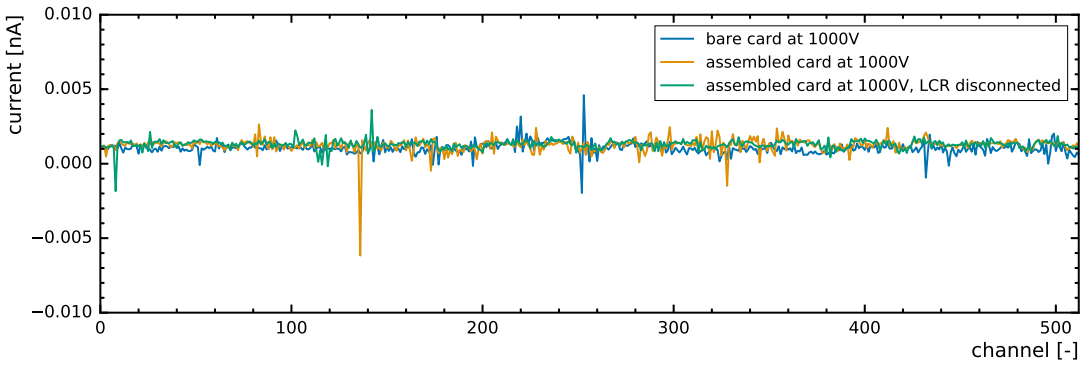


Figure 5.10: The system's leakage current per channel at 1 kV.

5.3.1 Verification of the Measurement

Fig. 5.11 shows the total current drawn by the SMU and the sum of all measured channels as well as the residual between them. At 1 kV, the total current exceeds the sum of all channels by about 20 nA. This difference is attributed to the leakage current of the three HV decoupling

capacitors on the board. In comparison to Fig. 5.9, the measurement of the total current is done after roughly 150 s.

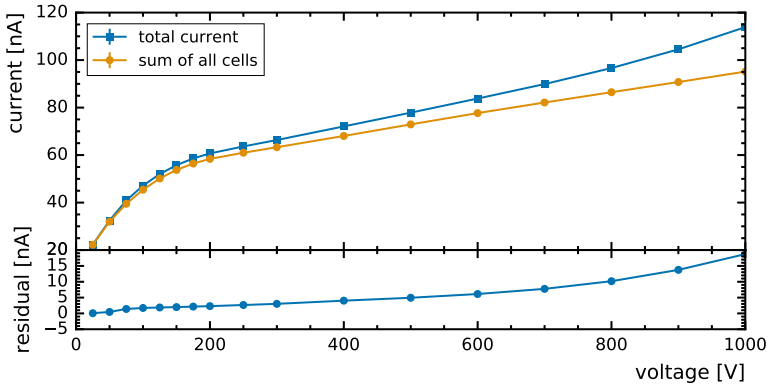


Figure 5.11: The total current drawn by the SMU and the sum of all channels that are measured. The residual value of about 20 nA is attributed to the leakage current of three HV capacitors in the circuit.

5.3.2 Comparison with Probe Head Measurement

An alternative to the probe card measurement is the measurement with micro-positioning probe heads and needle tips. For a hexagonal geometry, seven probe heads have to be used: one to contact the central pad-under-test and six to bias the immediate neighbours. If the latter are not biased, the depleted region of the central pad will punch through to the neighbouring pads and collect the current from them as well. This process will continue, with some time constant, until the voltage difference between a pad and its neighbour is smaller than the punch through voltage or until the full wafer is depleted. Therefore, this method would deliver very unstable as well as non-representative results. Fig. 5.12 shows the IV behaviour of the same pad once measured with a single probe head and floating neighbours and once with fully biased neighbours. A significantly larger current is observed for single probe head measurements.

Fig. 5.13 shows example IV curves for the same pad measured with the probe card and with a seven probe head setup for two different p-type sensors, one with common p-stop (a) and with atoll p-stop (b). Except for the p-stops geometry, the two sensors are identical. Several features can be observed from these figures. First, the current measured with a seven probe head setup is consistently larger than what is measured with the probe card. This indicates that the biasing of only the direct neighbours is not without effect on the pad-under-test. Second, a peak around 25 V is visible for probe head measurements with the atoll p-stop geometry. This effect is consistently visible in all probe head measurements with atoll p-stop sensors, although in various degrees of magnitude and with long time constants (several minutes). This is thought to be a charge up effect of the atoll p-stop implants and is not present when all pads are biased simultaneously.

In summary, measurements with a limited number of biased pads have shown limitations. As biasing more pads with probe heads is not practical, probe card measurements are preferred. Additionally, the guard ring, that is typically most susceptible to high fields and thus breakdown

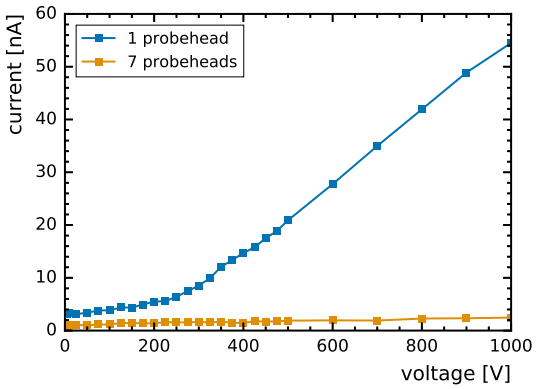


Figure 5.12: IV curve of the same pad once measured with a single probe head and once measured with seven probe heads. In the latter case, six of the seven probe heads are used to properly bias to the first neighbours of the pad-under-test.

below the operating voltage, cannot be measured by any reasonable means in a probe head setup with fully biased neighbours.

5.3.3 Results for Prototype Sensors

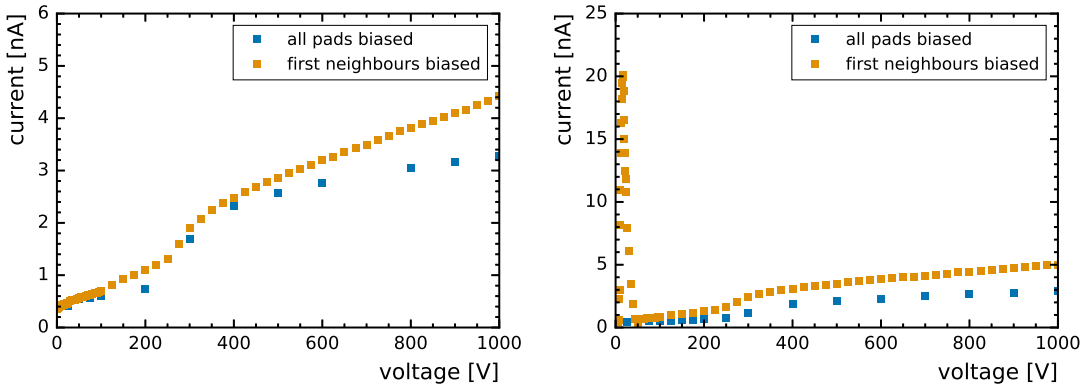
During the prototyping campaign of HGCAL, many sensors have been measured. Here, only a selection of the most relevant results is presented. All sensor types are listed in Tab. 5.1.

Results for 6 Inch Standard Diffused Sensors

Fig. 5.14 shows wafer views at 1000 V for an (a) n-type sensor, (b) p-type sensor with common p-stop and (c) p-type sensor with atoll p-stop and standard diffused backside (20 µm thick backside implant, see Sec. 5.1.1). All three sensor types are produced from 6 inch wafers.

For the n-type sensor, higher currents can be observed in the corner regions than in the centre. The p-type sensors show a significantly higher base current, though it is still very low in absolute terms. Due to a production error in this series, the atoll p-stop mask is slightly misaligned. In the 20 µm gap region (lower left quarter), this results in a contact between collection implant and p-stop implant and a few defect pads.

Fig. 5.15 shows the median and the inter-quartile range (IQR, see App. A) of the IV curves across the sensors. Only full sized pads are considered. These quantities are chosen because they are insensitive to outliers. The n-type sensor shows an almost textbook like behaviour. The current rises with $\sqrt{V}$ until full depletion at approximately 200 V due to the increase in volume and then follows a roughly linear rise. The two p-type sensors show a very different behaviour, after a slow linear rise, the current increases drastically just around the depletion voltage of approximately 300 V. This indicates that their leakage currents are not dominated by bulk contributions but rather by surface effects in connection to the p-stops.



(a) p-type common

(b) p-type atoll

Figure 5.13: Example IV curves of the same pad measured once on a fully biased sensor (with the probe card) and once with only the first neighbours biased (with probe heads). Generally, slightly lower currents are measured in the probe card configuration. For atoll geometries, significant discrepancies can be observed at lower bias voltages for the two configurations.

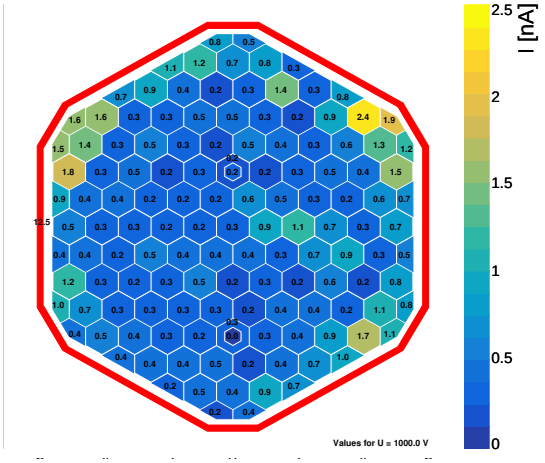
Results for 6 Inch Deep Diffused Sensors

As described in Sec. 5.1.1, deep diffused sensors have a backside implant of about 200 µm. Fig. 5.16 shows the wafer view at 1000 V for an example n-type sensor and a p-type sensor with common p-stop. A p-type sensors with atoll p-stop geometry has not been measured for this sensor type. The n-type sensor shows a radial dependence with higher currents at larger radii, similar to the STD sensor but with a much stronger trend. For the p-type sensor with common p-stop a trend to higher currents in the half with 30 µm gap distance is visible. Both sensors contain pads with the jumper geometry described in Sec. 5.1.2. At 1000 V, no degradation in HV stability can be observed with these pads.

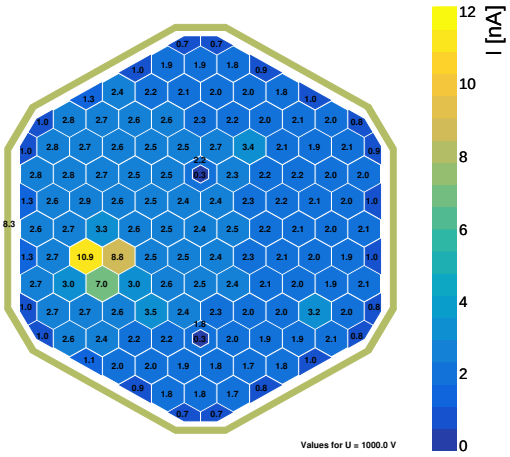
Fig. 5.17 shows the median and the inter-quartile range of the IV behaviour across the same sensors. Generally, the observed base current of the DD sensors is an order of magnitude larger than for the STD sensors. This is attributed to the larger diffusion current from the non-depleted volume in these sensors. The manufacturer approximates the full depletion voltage to be 55 V (n-type) and 65 V (p-type) respectively. However, due to their production process, these sensors never fully deplete. Rather, after a depletion depth of 120 µm the resistivity changes and one depletes further into the highly doped backside. This is more evident from the CV measurements presented in the following section.

5.4 Capacitance Measurement

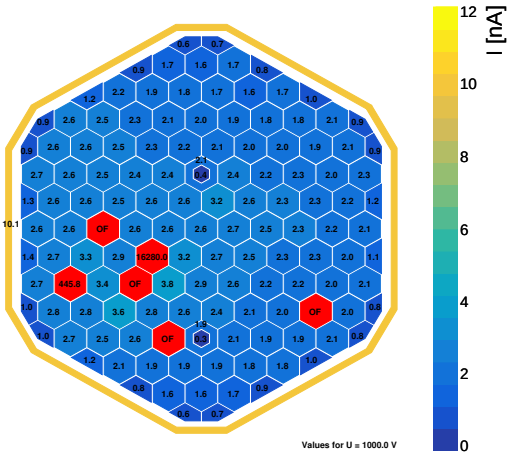
The capacitance is measured with an external LCR meter (Keysight E4980). It produces an AC voltage between its poles and measures the resulting AC current as well as the phase shift ϑ



(a) n-type

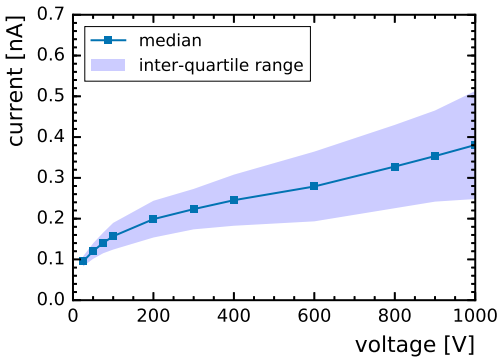


(b) p-type common

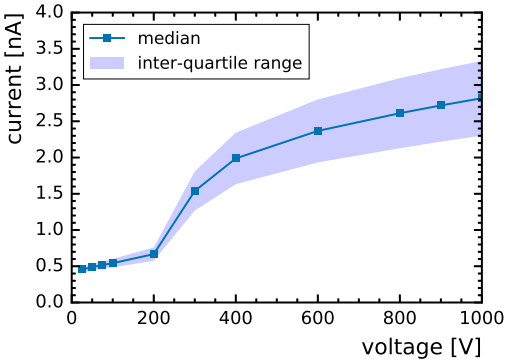


(c) p-type atoll

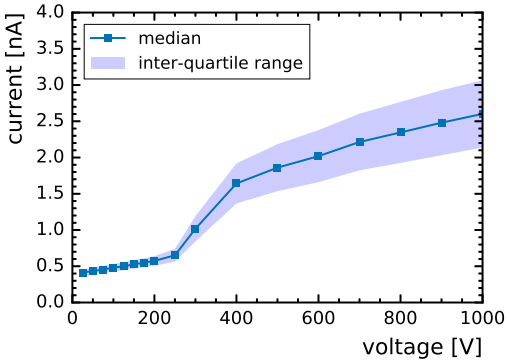
Figure 5.14: Wafer views at 1000 V of a set of standard diffused sensors. OF indicates values above 40 µA. Red coloured pads have values larger than the maximum of the colour scale.



(a) n-type

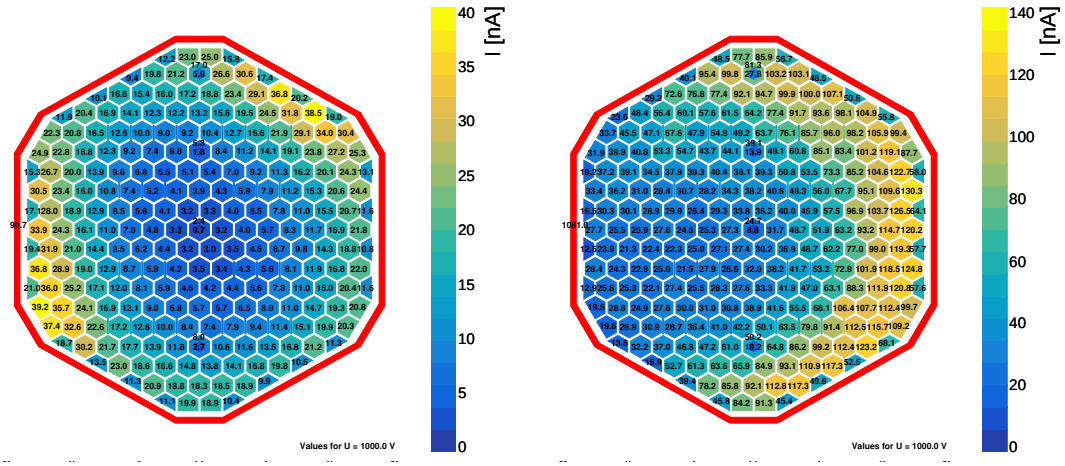


(b) p-type common



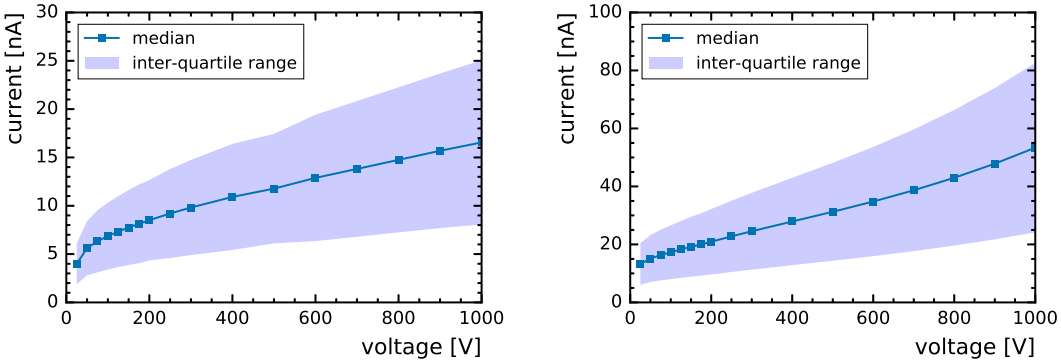
(c) p-type atoll

Figure 5.15: Median IV curves across all pads of a set of standard diffused sensors.



(a) n-type (b) p-type common

Figure 5.16: Example wafer views at 1000 V of a set of deep diffused sensors.



(a) n-type (b) p-type common

Figure 5.17: Median IV curves across all pads of a set of deep diffused sensors.

between the two. From these measurements, the complex impedance is calculated,

$$Z = \frac{V_0}{I_0} \exp^{-i\vartheta}. \tag{5.1}$$

Assuming a parallel equivalent circuit, i.e. a purely ohmic element R_p and a purely capacitive element C_p in parallel, the admittance Y, conductance G, and susceptance B can be calculated.

$$Y = \frac{1}{Z} \qquad G = Y\cos\vartheta \qquad B = Y\sin\vartheta. \tag{5.2}$$

The parallel capacitance is given by

$$C_\mathrm{p} = \frac{B}{2\pi f}. \tag{5.3}$$

To limit the influence of the cable impedance on the measurement, four wire sensing is used. The four wires are connected on the switch card at the closest possible point to the pad-under-test.

Due to the existence of stray capacitances in any real circuit, the parallel equivalent model is never fully accurate and corrections have to be applied. The simplest correction is to assume that all stray elements of the circuit can be modelled as one parallel impedance. The corrected capacitance is then

$$C_\mathrm{cor} = C_\mathrm{meas} - C_\mathrm{open}. \tag{5.4}$$

In literature, this is referred to as open compensation. Here, C_meas is the capacitance of the two-card system plus sensor and C_open the capacitance of the two-card system without sensor. The latter is measured with all pins lifted from the sensor. This is however not the exact compensation as it would require the removal of the single capacitance-under-test. With a full-size silicon pad sensor, this is not possible and all pads are disconnected. The systematic uncertainty introduced by this procedure has been analysed with a SPICE simulation and is less than 0.1 pF for frequencies between 1 and 100 kHz.

An alternative compensation model is the so-called open-short compensation, which assumes all stray elements can be modelled by one serial and one parallel impedance. The measured impedance is then corrected with open and shorted measurements,

$$Z_\mathrm{cor} = \frac{Z_\mathrm{meas} - Z_\mathrm{short}}{1 - (Z_\mathrm{msr} - Z_\mathrm{short})Y_\mathrm{open}}. \tag{5.5}$$

Here, Z_cor is the corrected impedance of the pad-under-test, Z_meas is the measured impedance with the pad-under-test connected, Z_short is the measured impedance with the pad-under-test shorted and Y_open the admittance with the pad-under-test disconnected. For more details on this compensation model, the reader is referred to [62, chapter 4.3].

Fig. 5.18 shows the capacitance of the bare switch card and after assembly of switch and probe card. For the bare switch card, a step structure and two sub structures, each with a period of 8, are visible. These structures are due to the three layers of multiplexers and can be correlated to the physical position of each channel on the board. The absolute values can change by O(10%) from PCB to PCB, however the structure remains unchanged. Once the probe card is attached to the switch card, the capacitance of the channels with a connected trace on the probe card is increased. The values reach from 2 to 20 pF and correspond to the trace lengths of the routing on the probe card. If an open compensation is used, this is the offset capacitance that is subtracted. For both compensation methods, all corrections are done per channel and voltage.

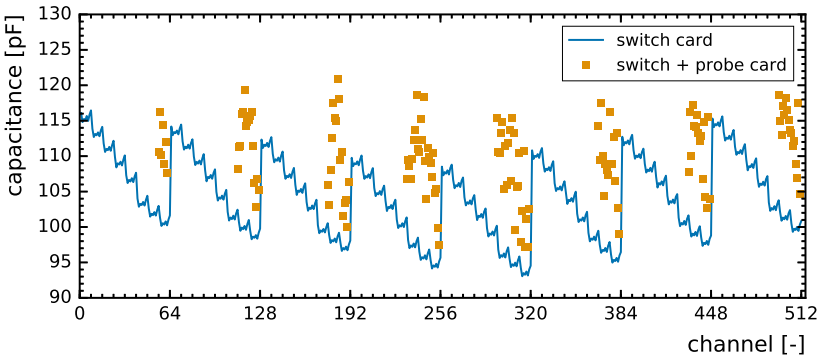


Figure 5.18: The capacitance per channel for the switch card without and with probe card connected. In the latter case, only the connected channels are measured.

5.4.1 Conceptual Design of the Capacitance Measurement

On the sensor, all pads are in parallel and have to be biased to obtain representative results. However, only the capacitance of an individual pad should be measured at a time. To achieve this, the parallel circuit formed by all other pads has to be removed from the measurement. Since it is not possible to physically remove the pads, their influence has to be suppressed by other means. This is achieved by exploiting the measurement principle of the LCR meter. If the impedance of the circuit in parallel to the pad-under-test is increased by introducing a large ohmic element R_{bias}, its influence on the measurement result can be suppressed.

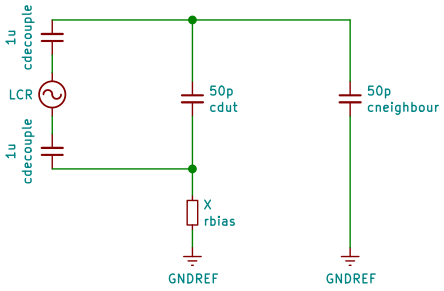
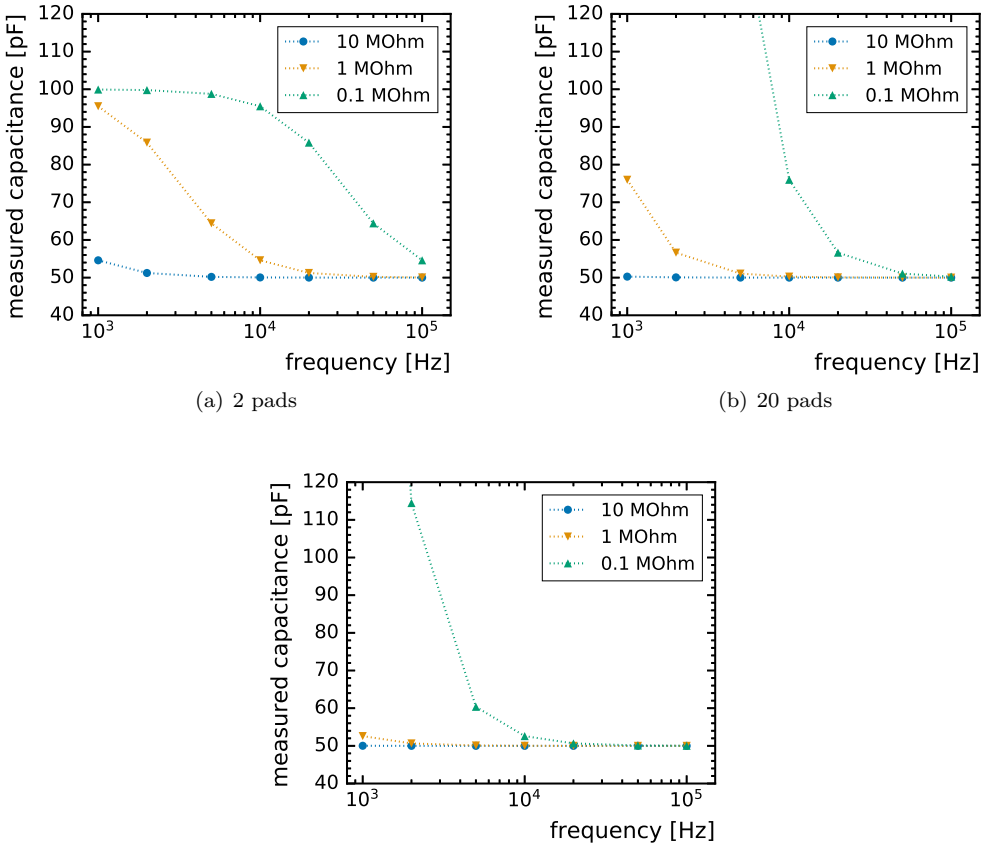


Figure 5.19: Example circuit with 2 parallel pads and the LCR meter. The resistor R_{bias} increases the impedance in the parallel path and allows a measurement of a single pad.

Consider the circuit in Fig. 5.19, representing two pads in parallel and the LCR meter. The effect of R_{bias} will depend on the ratio of the impedance between the two parallel circuits. At higher frequencies, the value of the resistor can be reduced while keeping the same effect. To investigate this further, the simple circuit of Fig. 5.19 was simulated in SPICE with multiple values for R_{bias} and the number of pads in the circuit. Fig. 5.20 shows the results of the simulated capacitance measurement for multiple pads of 50 pF and for different scenarios.

For the case of two pads and 100 kΩ, at low frequencies the measured capacitance is the total capacitance of both pads as the impedance of the pads themselves is significantly larger than



(a) 2 pads

(b) 20 pads

(c) 200 pads

Figure 5.20: Effect of R_{bias} on the measurement of multiple parallel capacitors of 50 pF each.

100 kΩ. For example, at 1 kHz, a 50 pF capacitance has an impedance of roughly 3.2 MΩ. If the frequencies or the value of R_{bias} are increased, the ratio of the two impedances changes and the measured capacitance approaches the value of 50 pF. For the case of 20 and 200 pads, the parallel capacitance is increased and hence the impedance is decreased. This amplifies the mechanism seen in the case of two pads and allows for a correct measurement also at lower frequencies and/or smaller values of R_{bias}.

For the initial design of the switch card, a value of 10 MΩ was chosen. As long as the leakage current of the pad-under-test is less than a few hundred nA, the DC voltage drop across this resistor is negligible. However, for irradiated sensors with leakage currents of several µA, this resistance will have to be reduced.

5.4.2 Simulation of the Full System

The full electrical circuit is more complex than the simplified model shown in Fig. 5.19. Each multiplexer (MUX) introduces a stray capacitance to ground and a series resistance. This will influence the capacitance measurement. Therefore the simple parallel equivalent model is not accurate. On the probe card, parallel routing of traces (though avoided where possible) introduce cross coupling between channels as can also be seen from Fig. 5.18. To evaluate the influence of these effects, a more detailed simulation model has been implemented in a SPICE simulation. Fig. 5.21 shows a partial visualisation of the simulated circuit. The pad-under-test is denoted as 'cdut'. Besides the global elements already discussed in Sec. 5.2, it includes multiple possible parasitic capacitances to ground ('cwire', 'cgnd') as well as to neighbouring channels ('ccross') and internal models for the switches ('rsw', 'csw') and multiplexers ('rmux', 'cmux'). The capacitive behaviour of the power supply is simulated by including a 10 mF bypass capacitor ('cbypass'). Appendix C lists the python function used to generate the netlist, including a description of all simulated circuit elements.

Tuning of Free Parameters in the Simulation Model

The values of the stray capacitances included in the model shown in Fig. 5.21 are a priori unknown. Some values (e.g. cgnd and ccross) can be measured. However, precise measurements have proven difficult and the exact values are likely to vary between channels. For others (e.g. cmux or rsw), average values are given in the data sheets. The values in Fig. 5.21 are found by tuning the simulation to measured data over a wide range of frequencies and are in agreement with manual measurements as well as vendor data. The result is shown in Fig. 5.22. The frequency response is well reproduced up to around 100 kHz. Here, the attention should be on the relative behaviour rather than the absolute values which may vary from channel to channel.

Verifying the Simulation Model

To validate the simulation model, the difference of two observables, the open and open-short compensated values, are compared between data and simulation for two different values of R_{series}. These observables are highly sensitive to details in the circuit and should therefore present a good

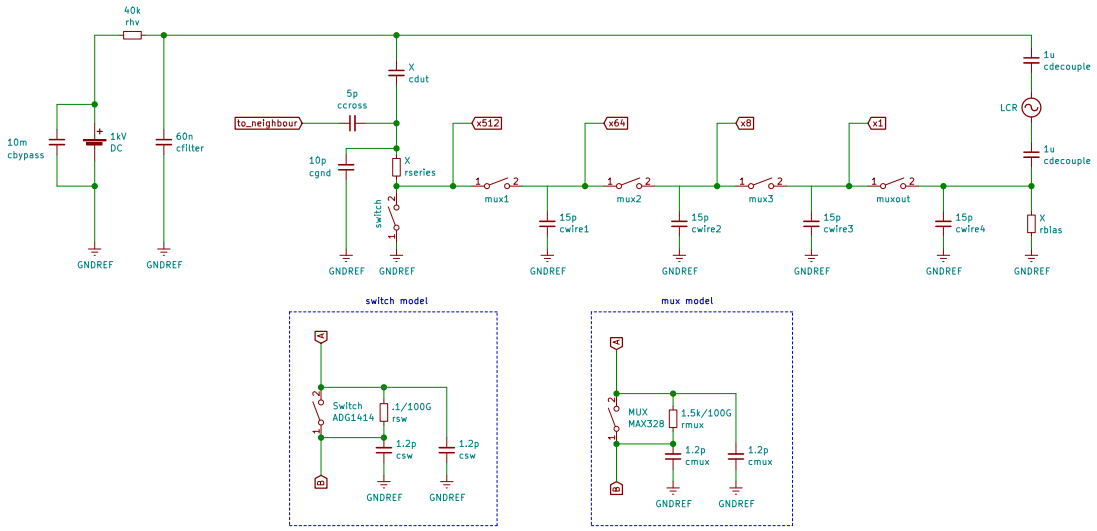


Figure 5.21: Partial visualisation of the SPICE simulation model used to simulated multiple parallel capacitors of 50 pF each. The sub circuits used to model the switches and multiplexers are shown at the bottom. The pad-under-test is denoted as 'cdut'.

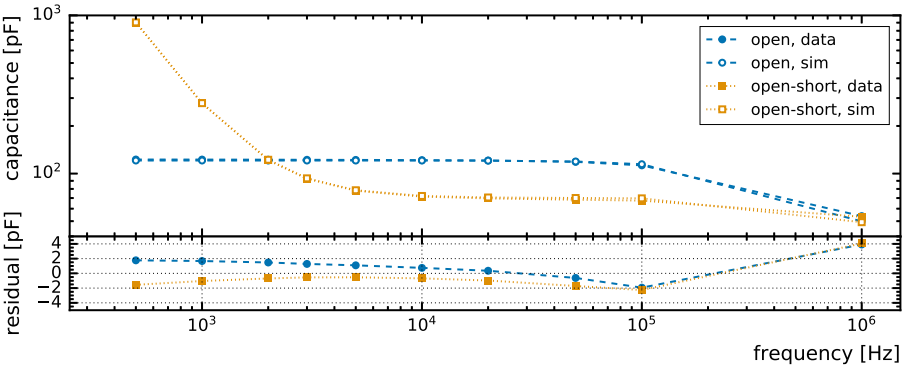


Figure 5.22: Simulated and measured data for open and open-short compensated measurements of the two card system.

benchmark for the simulation. Fig. 5.23 shows that the simulation is accurate to better than 0.1 pF in the range of 1 to 10 kHz and better than 0.2 pF at 20 kHz.

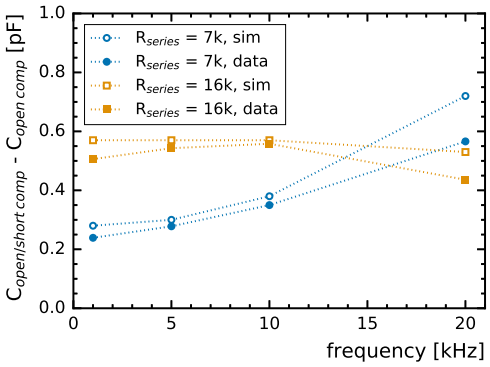


Figure 5.23: Comparison between simulation and measurement for open and open-short compensated cases as a function of frequency. The difference is sensitive to details in the circuit and is used to benchmark the simulation accuracy.

5.4.3 Accuracy of the Capacitance Measurement

The two main factors determining the precision of the capacitance measurement are R_{bias}, which increases the impedance of the circuit parallel to the pad-under-test, and the series resistance R_{series} of the multiplexer network and the protection resistor. To investigate their effect on the measurement accuracy, a full simulation with 270 parallel pads (corresponding to one specific probe card design) is used.

Fig. 5.24(a) shows the systematic measurement uncertainty as a function of the measurement frequency at a pad capacitance of 50 pF for different values of R_{bias} and for $R_{\text{series}} = 16\,\text{k}\Omega$. At low frequencies, the pad impedance becomes comparable to an O(1 MΩ) resistor. If the value of R_{bias} is too small, the impedance of the circuit parallel to the pad-under-test is not negligible anymore. At higher frequencies, the pad impedance decreases and eventually becomes comparable to O(10 kΩ) and therefore R_{series}. In this case, the parallel equivalent model used to calculate a capacitance from the measured impedance becomes invalid and introduces a systematic uncertainty. Fig. 5.24(b) shows the systematic uncertainty on the measurement as a function of the pad capacitance for different frequencies at $R_{\text{bias}} = 10\,\text{M}\Omega$. A systematic uncertainty below 0.2 pF can be maintained over the range from 1 to 100 pF in the 1 to 2 kHz frequency range.

For safe operation of the multiplexers, the voltage drop across R_{bias} should not exceed 10 V. This limits the maximum leakage current of pads to 1 µA for capacitance measurements. In irradiation studies for HGCAL, this current will easily be exceeded. For this purpose, a second version of the switch card has been designed with a variable bias resistor of eight discrete steps between 100 kΩ to 100 MΩ. As can be seen from Fig. 5.24(a), reducing R_{bias} to 1 MΩ has only a very small impact on the measurement at low frequencies. Further reduction of R_{bias} however limits the performance. To improve the accuracy for higher frequencies and lower values of R_{bias}, the protection resistor is reduced to 1 kΩ, for a total series resistance of $R_{\text{series}} = 7\,\text{k}\Omega$. The resulting accuracy is shown in Fig. 5.25.

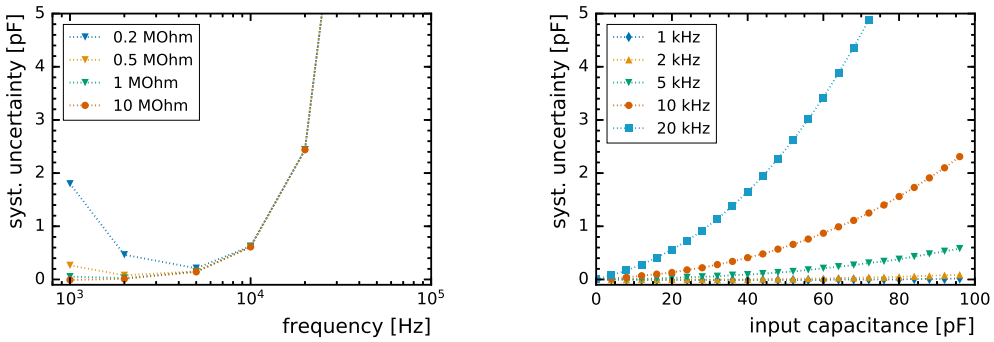


Figure 5.24: Predicted systematic uncertainty of the capacitance measurement from SPICE simulations with $R_{\text{series}} = 16\,\text{k}\Omega$ for (a) different values of R_{bias} at a pad capacitance of 50 pF and (b) at $R_{\text{bias}} = 10\,\text{M}\Omega$ for different frequencies as a function of the pad capacitance.

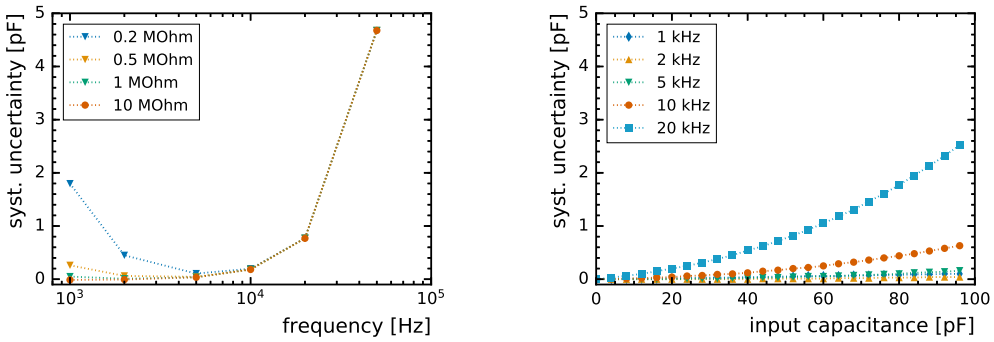


Figure 5.25: Predicted systematic uncertainty of the capacitance measurement from SPICE simulations with $R_{\text{series}} = 7\,\text{k}\Omega$ for (a) different values of R_{bias} at a pad capacitance of 50 pF and (b) at $R_{\text{bias}} = 1\,\text{M}\Omega$ for different frequencies as a function of the pad capacitance.

Open vs. Open-Short Compensation

Open-short compensation has also been investigated in simulation and measurement but did not yield superior results. Fig. 5.26 shows the achievable accuracy with (a) open compensation (Eq. 5.4) and (b) open-short compensation (Eq. 5.5) for the case of $R_{\text{bias}} = 10\,\text{M}\Omega$ and $R_{\text{series}} = 16\,\text{k}\Omega$. At high frequencies, the open-short compensation is superior to the open compensation while at low frequencies the opposite is the case. Generally, the results of the open-short compensation are more difficult to predict. In further simulations, the results have also been less robust against details of the simulation, e.g. the exact value of the cross coupling between traces on the probe card (ccross in Fig. 5.21). The simpler open compensation is therefore the preferred method and used throughout this work.

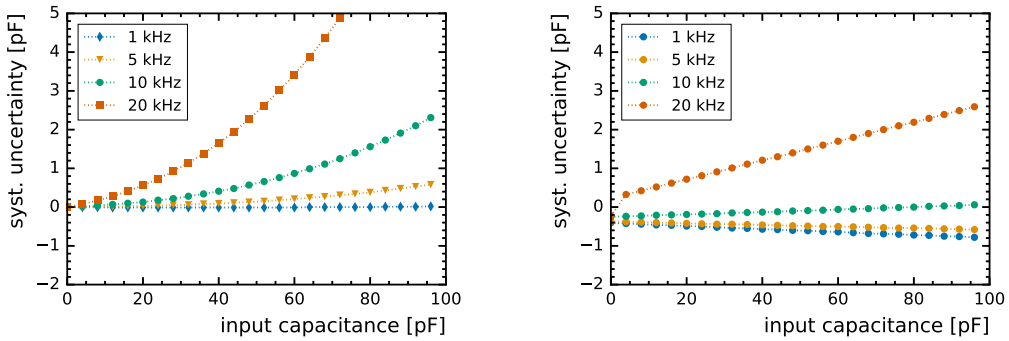
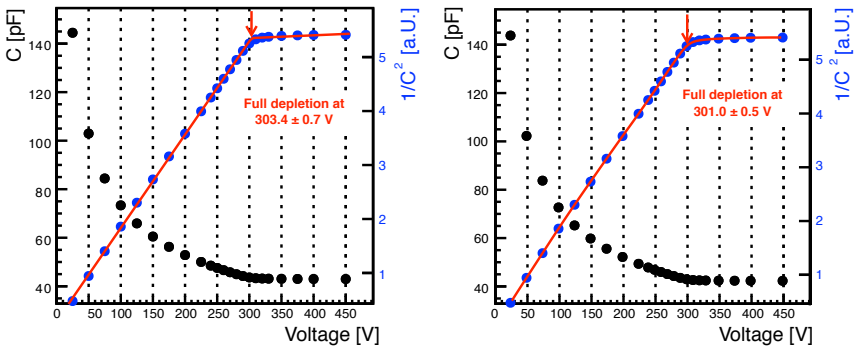


Figure 5.26: Comparison of the achievable accuracy with (a) open compensation and (b) open-short compensation with $R_{\text{bias}} = 10\,\text{M}\Omega$ and $R_{\text{series}} = 16\,\text{k}\Omega$.

5.4.4 Extraction of the Depletion Voltage

From the CV behaviour of a sensor, its full depletion voltage V_{fd} can be extracted (see also Eq. 2.4). By plotting $1/\text{C}^2$ as function of V_{bias}[3], two separate regimes can be observed. Before full depletion, the behaviour follows a linear rise $\propto V_{\text{bias}}$. For higher values of V_{bias}, the capacitance remains (nearly) constant. A common method to extract V_{fd} is to fit a straight line to each regime and extract V_{fd} as the intersection of the two lines. This approach requires a prior identification of the two regimes. Another approach tried in this work is to extract V_{fd} as the point of maximum curvature from a quintic spline fit to the $1/\text{C}^2$ curve. Generally, both methods give the same results within $\pm 3\,\text{V}$. However, the approach using spline fits has proven more robust, especially in automatic analysis of several hundred CV curves. An example curve analysed with both methods is shown in Fig. 5.27.

[3]Here, the internal voltage of the junction V_{j} is neglected and it is assumed that $V_{\text{bias}} = V_{\text{eff}}$.



(a) (b)

Figure 5.27: Two methods to extract the full depletion voltage V_{fd}: (a) using two line fits in two separate regimes and define V_{fd} as the intersection and (b) using a quintic spline fit and define V_{fd} as the point of maximum curvature.

5.4.5 Extraction of the Edge Capacitance

The measured pad capacitance is the sum of the bulk capacitance C_{bulk} and the inter-pad capacitance $C_{\text{inter-pad}}$ which are assumed to scale with the pad area A and the pad circumference S respectively,

$$C_{\text{tot}} = C_{\text{bulk}} + C_{\text{inter-pad}} = A \cdot C_{\text{planar}} + S \cdot C_{\text{edge}}, \tag{5.6}$$

with the planar contribution defined as a parallel plate capacitor with thickness t

$$C_{\text{planar}} = \frac{\epsilon_0 \epsilon_r}{t}. \tag{5.7}$$

A value of 11.9 is used for the relative permittivity ϵ_r of silicon [18].

Two methods are used to extract the inter-pad capacitance. The first method relies on knowledge of the active thickness of the sensor. The edge capacitance C_{edge} is then calculated via

$$C_{\text{edge}} = \frac{C_{\text{tot}} - A \cdot C_{\text{planar}}}{S} \tag{5.8}$$

This method is used for physically thinned sensors in which the active thickness equals the passive thickness minus the aluminium layers. The accuracy of the active thickness is quoted as ±3% by the manufacturer. However, in thickness measurements of test structures an accuracy of ±1% was found.

For STD and DD sensors, this method is not usable as the active thickness is not a priori known. Instead, a different method similar to the one described in [63] is used. Assuming two pads with same thickness t, same gap distance d but different area A, one can extract C_{planar} and C_{edge} as

$$C_{\text{planar}} = \frac{S_2 \cdot C_{\text{tot},1} - S_1 \cdot C_{\text{tot},2}}{S_2 \cdot A_1 - S_1 \cdot A_2}, \tag{5.9}$$

$$C_{\text{edge}} = \frac{A_2 \cdot C_{\text{tot},1} - A_1 \cdot C_{\text{tot},2}}{A_2 \cdot S_1 - A_1 \cdot S_2}. \tag{5.10}$$

On the 6 inch design of the STD sensors, the conditions mentioned above are fulfilled by any central pad from the 60 µm gap quadrant as well as the two calibration pads. The values are then extracted at 50 V above full depletion. The propagated uncertainties from the measurement uncertainties of $C_{\mathrm{tot},1}$, $C_{\mathrm{tot},2}$ on C_{planar}, C_{edge} are given by

$$\Delta C_{\mathrm{planar}} = \sqrt{\left(\frac{S_2 \cdot \Delta C_{\mathrm{tot},1}}{S_2 \cdot A_1 - S_1 \cdot A_2}\right)^2 + \left(\frac{S_1 \cdot \Delta C_{\mathrm{tot},2}}{S_2 \cdot A_1 - S_1 \cdot A_2}\right)^2} \tag{5.11}$$

$$\Delta C_{\mathrm{edge}} = \sqrt{\left(\frac{A_2 \cdot \Delta C_{\mathrm{tot},1}}{A_2 \cdot S_1 - A_1 \cdot S_2}\right)^2 + \left(\frac{A_1 \cdot \Delta C_{\mathrm{tot},2}}{A_2 \cdot S_1 - A_1 \cdot S_2}\right)^2}. \tag{5.12}$$

Here, it is assumed that the areas A_1, A_2 as well as the circumferences S_1, S_2 are known with a negligible uncertainty. With the lithography precision typically in the order of µm, this is certainly the case. Fig. 5.28 shows the propagated uncertainty on C_{bulk} and $C_{\mathrm{inter\text{-}pad}}$ for the full pad geometry. The statistical uncertainty of the pads is typically less than 0.1 pF. The systematic uncertainty is estimated from the simulation results shown in Fig. 5.24(b) which at 1 kHz predict a systematic uncertainty below 0.1 pF. In accordance with Sec. 5.4.2, the total uncertainty on the measurement of the pads is then estimated as 0.14 pF. The propagated uncertainty for $C_{\mathrm{inter\text{-}pad}}$ of the full sized pad is about 0.5 pF at 1 kHz.

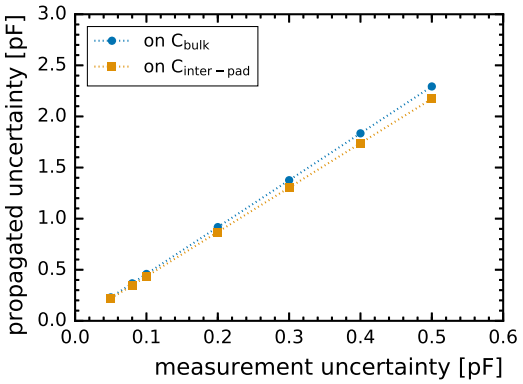


Figure 5.28: Uncertainty for the bulk and inter-pad contribution of the full pad as a function of the measurement uncertainty ΔC_{tot}.

5.4.6 Results for Prototype Sensors

The following results have been measured with a 10 MΩ bias resistor and 16 kΩ series resistance. Tab. 5.2 shows the detailed information on the geometry of the pads of the various sensors that is used for the calculation of the inter-pad capacitances.

Results for 6 Inch Standard Diffused Sensors

Fig. 5.29 shows the capacitance at approximatly 100 V over-depletion for the three different types of standard diffused sensors (STD), each with 300 µm in (nominal) active thickness. Due

Table 5.2: Detailed values of the frontside geometries of the measured sensors. All values are with respect to the metal implant (incl. overhang). Gap distances and pad areas are identical for p- and n-type sensors.

Common Name [-]	Gap Distance [µm]	Pad Area [mm^2]	Pad Circumference [mm]
6in 135ch Full Pad	80	108.22	38.72
	60	108.59	38.79
	40	109.01	38.87
	20	109.18	38.89
6in 135ch Cal. Pad	50	14.59	14.22
6in 239ch Full Pad	50	58.38	28.44
	30	58.57	28.49
8in 271ch Full Pad	50	96.47	36.56

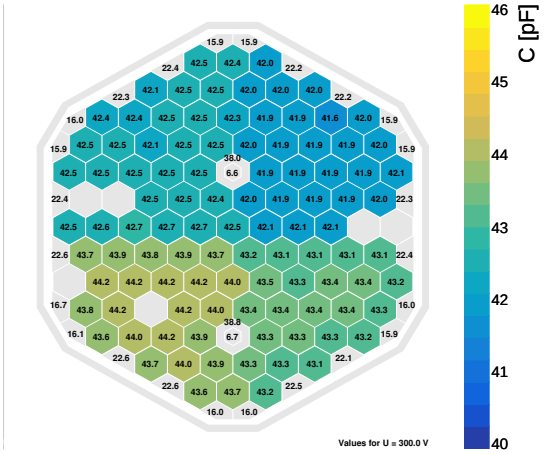
to the additional p-stop implant, p-type sensors show a slightly larger capacitance than the n-type sensor. The influence of the inter-pad gap distance in the different quadrants on the total capacitance is also resolved.

For the region with a 60 µm gap distance, the inter-pad capacitance for the fully depleted sensor can be separated by comparing the average total capacitance value of that region with the average of the two calibration pads and using Eq. 5.12. For the other quadrants, the inter-pad values are calculated by assuming the same planar capacitance, scaling it to the pad area and taking the difference to the measured capacitance. The result is shown in Fig 5.30 for the three sensor types. As shown in Fig. 5.28 the relative uncertainty on this measurement procedure is at the level of 10%. Nevertheless, a trend towards larger capacitances with p-stop implants and smaller gap distances can be observed. Below full depletion, the inter-pad capacitance decreases and due to the large uncertainty a valid measurement is not possible anymore.

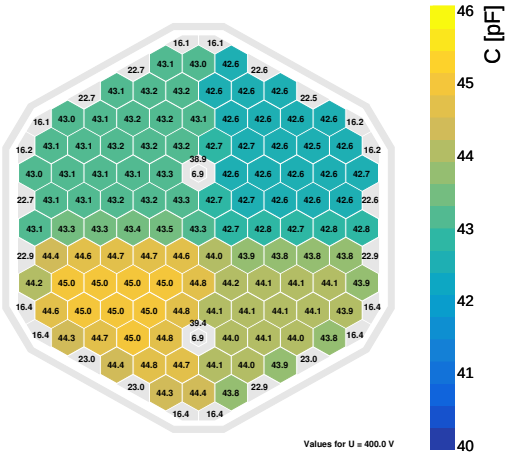
Fig. 5.31(a) shows the CV and $1/C^2$ behaviour for the same set of sensors, averaged over all full sized pads of the 60 µm region. The error bars correspond to the quadratic sum of standard deviation and systematic uncertainty. The n-type sensor is fully depleted at 200 V, while the two p-type sensors only deplete at 280 V (common p-stop) and 300 V (atoll p-stop). While a higher V_{fd} is expected for p-type sensors due to the lower mobility of holes, the difference between common and atoll p-stop is unexpected. However, while the bulk and backside processing is the same for both sensors, they are from different production batches. Further conclusions on this topic would require more statistics.

Results for 6 Inch Deep Diffused Sensors

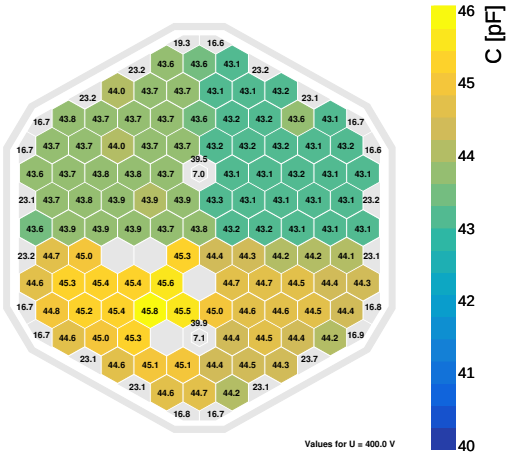
Fig. 5.32 shows the CV and $1/C^2$ behaviour for a deep diffused (DD) sensor. Average values for multiple pads with and without jumper geometry from the 50 µm gap region are shown. Due to the nature of this sensor, the capacitance does not reach a constant value at full depletion but decreases further as the depleted region extends into the deep diffused backside implant. From the different doping concentrations of bulk and backside, the nominal depletion depth of 120 µm



(a) n-type



(b) p-type common



(c) p-type atoll

Figure 5.29: Wafer views of the capacitance at approximate 100 V over-depletion for a set of 135-channel standard diffused sensors measured at a frequency of 1 kHz. Grey coloured pads have either values smaller than the minimum of the colour scale (e.g. half cells) or no proper measurement was achieved.

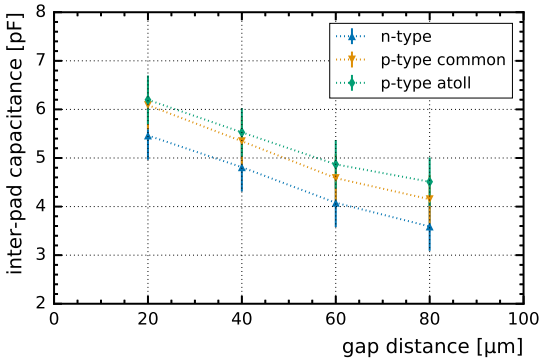
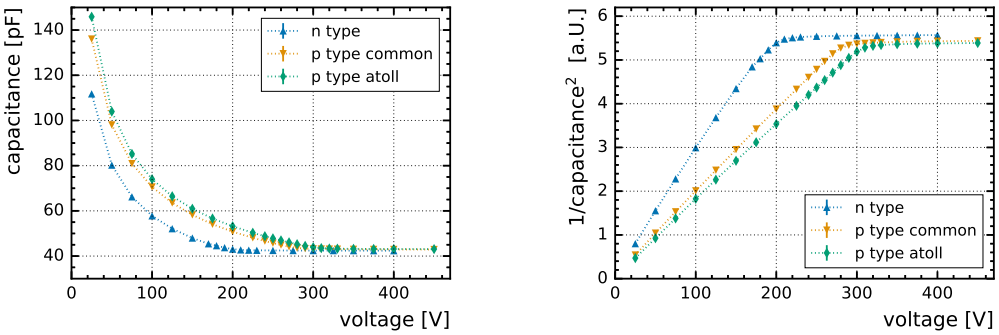


Figure 5.30: Inter-pad capacitance values for a set of standard diffused sensors at approximate 100 V over depletion.

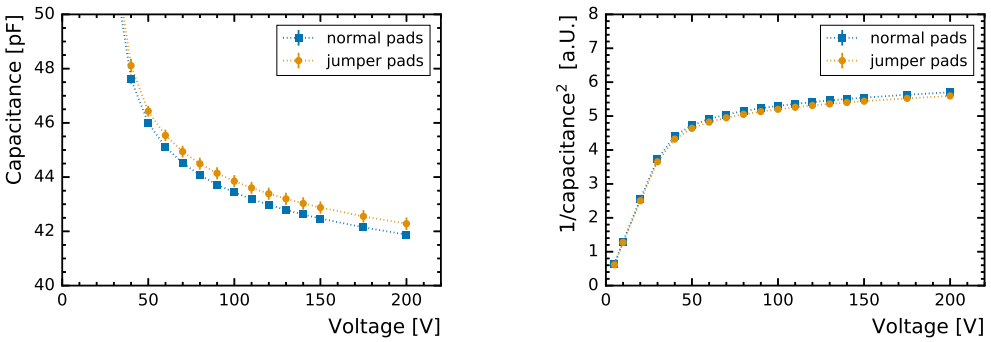


(a) (b)

Figure 5.31: Average CV and $1/C^2$ behaviour for a set of 135-channel standard diffused sensors (60 µm inter-pad gap distance).

can be extracted as approximately 45 V.

The jumper geometry increases the pad capacitance by about 0.5 pF. However, the fluctuations in capacitance of the measured pads with jumper geometry were significantly larger than for the pads without jumper geometry. This can be seen from the larger uncertainties on the corresponding measurement points that are dominated by the statistical uncertainty.



(a) (b)

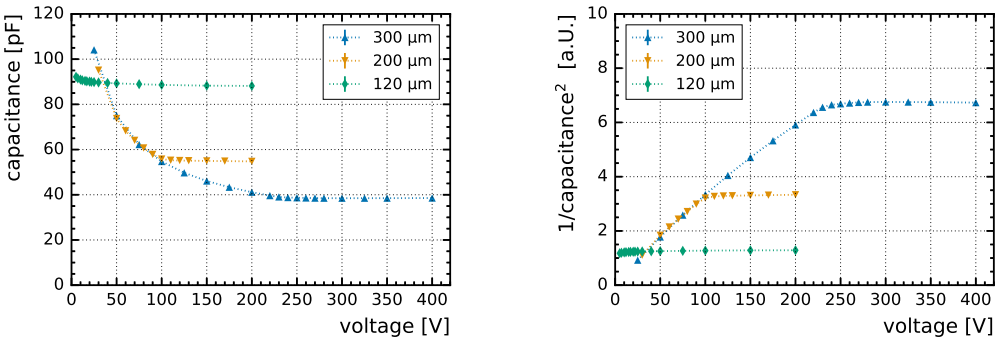
Figure 5.32: Average CV and $1/C^2$ behaviour for a deep diffused sensor (50 µm inter-pad gap distance) for pads with and without jumper geometry. The active thickness does not reach a constant value but increases beyond the nominal value of 120 µm.

Results for 8 Inch Sensors

For the 8-inch sensors, prototypes have been produced with three different thicknesses of 120 µm, 200 µm and 300 µm but identical frontside masks. The 200 µm and 300 µm thick sensors are physically thinned, while the 120 µm sensor is an epitaxial layer on a carrier wafer. All full pads are slightly smaller than for the 6 inch prototypes (see Tab. 5.2). The inter-pad gap distances are 50 µm and only atoll p-stop geometries are used.

Fig. 5.31 shows the CV and $1/C^2$ behaviour, averaged over multiple pads. The three sensor thicknesses are fully depleted at 15 V, 95 V and 235 V respectively.

For these sensors, the inter-pad capacitance is calculated by assuming knowledge of their active thickness. Multiple measurements of the total thickness on two separate test structures per wafer have been conducted with a precision vernier caliper. Values of 306 µm and 204 µm with an uncertainty of about ±1% was found. From the measured thickness, 4 µm are subtracted to account for passivation, metallisation and non-depleted implant on front- and backside. The result of 302 µm and 200 µm is in excellent agreement with the nominal values of the manufacturer. The expected bulk capacitance is then subtracted from the measured total capacitance, analogous to Eq. 5.12. To compare these values with the results from the STD sensors, they have to be normalised to unit circumference.



(a) (b)

Figure 5.33: Average CV and $1/C^2$ behaviour for a set of 8 inch sensors with identical frontside but different thicknesses.

5.4.7 Edge Capacitance

Fig. 5.34 shows all inter-pad capacitances normalised to a unit circumference, i.e. the edge capacitance. Here, closed markers represent data points for 300 µm thickness and open markers represent data points for 200 µm thickness. The 300 µm thickness values at 50 µm gap distance from the 8 inch sensor fit nicely with the trend of the data from the 6 inch sensors. A decrease of the edge capacitance with gap spacing is observed. At 200 µm thickness, the edge capacitance is smaller compared to 300 µm. This behaviour is also confirmed by TCAD simulation, as discussed in the following.

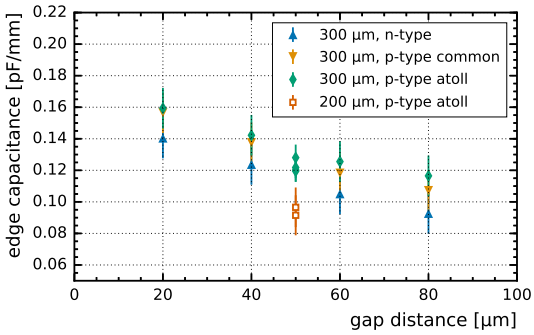


Figure 5.34: Edge capacitance values for all tested sensor types.

Fig. 5.35 compares the measured values from the 300 µm n-type sensor to TCAD simulations of the same design. The simulated data points are taken from [41, Fig. 7.2] and normalised to a unit circumference. At 20 µm gap distance, they agree well within the uncertainty. At larger gap distances however, the simulated values appear to systematically underestimate the edge capacitance.

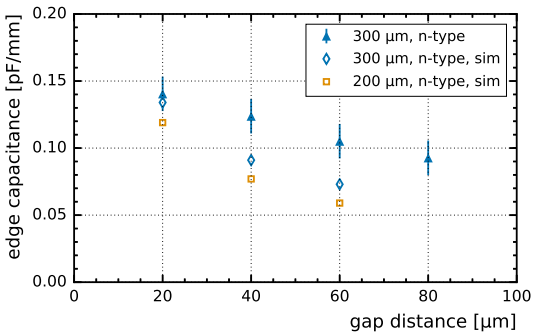


Figure 5.35: Edge capacitance for a 300 µm n-type sensor compared to TCAD simulations taken from [41, Fig. 7.2].

5.5 Summary and Further Implications

With ARRAY, a sensor testing system has been developed that is well adapted to the high granularity of future particle flow based calorimeters. The system consists of two plugin PCBs: an active switching matrix with 512 input channels that holds all controls and a passive probe card that connects to the sensor. While the switch card is universal, a dedicated probe card is developed for each sensor geometry. The system delivers accurate measurements for a large range of cell currents and capacitances. Currents ranging from 500 pA to 5 µA and capacitances between 5 pF and 100 pF have been measured so far. A precision of better than 0.2 pF on capacitance measurements in that range can be achieved. With this system, several prototype sensors for HGCAL have been measured and design choices have been evaluated.

Scaling to CLIC parameters of 0.25 cm^2 pad size and 500 µm thickness, the total pad capacitance will be around 7.1 pF (assuming a 60 µm gap distance), well within the measurement range of the system. If 8 inch sensors are used, the number of pads per sensor will surpass 512 input channels. In this case, the sensor either has to be measured in two steps or the switching matrix has to be redesigned to add another layer of multiplexers. This will add another 1.5 kΩ of series resistance and additional stray capacitances but should not degrade the accuracy of the capacitance measurement significantly. Designing a probe card with such a high channel density and without introducing additional cross channel capacitances appears to be a more prominent challenge.

Chapter 6

Precision Timing Capabilities of Silicon Pad Sensors

At the HL-LHC, one of the two main driving factors in the detector design is the unprecedented level of pile-up. As discussed in Sec. 3.3, the pile-up will increase to an average 140 proton-proton interactions per bunch crossing with peak vertex densities of $1.2\,\mathrm{mm}^{-1}$. The main counter measure besides the high granularity, is precise time information. A resolution of 30 ps is envisaged at the single cell level. For charged particles, this resolution will be provided by a separate MIP timing detector [42]. In the endcaps, it is based on LGAD detectors. For photons, the ECAL is used.

A preliminary study on the impact of precision timing on event reconstruction and a selection of benchmark physics observables can be found in [42]. As an example, the statistical uncertainty on the cross section measurement of $H \to \gamma\gamma$ decays is predicted to improve by 25% if a time resolution of 30 ps per photon is available in the calorimeter. Following the discussion in Sec. 2.4, several factors have to be fulfilled to achieve such a time resolution. The electronics has to enable the resolution in terms of noise and digitisation performance. Then, all channels have to be synchronised to a precision better than the envisaged resolution. And finally, the contributions to the measurement uncertainty from the signal development itself have to be small enough. The latter is measured in this chapter.

To evaluate the intrinsic timing capabilities of the silicon pad sensors and establish the contributions of the silicon sensor itself, beam tests were conducted at the CERN North Area. The experimental setup used for these tests is described Sec. 6.1. In Secs. 6.2 and 6.3 the basic signal properties and the calibration of the time response are discussed. The results on the time resolution achievable with a single silicon pad are presented in Sec. 6.4 and results on cluster timing and uniformity are shown in Sec. 6.5. Finally, section 6.6 gives an outlook on the challenges to translate these test beam results to a full system and summarises the results.

Some results of this chapter have also been published in [41] and [64].

6.1 Experimental Setup and Methodology

The experimental setup is shown in Fig. 6.1. Incident electrons of 32 GeV to 250 GeV are sent into a closely stacked set of absorber plates with 5.7 X_0 of tungsten (W). An HGCAL prototype sensor is used to sample the electromagnetic shower immediately downstream of the absorber stack. It is readout by custom electronics, here denoted as timing module. A delay wire chamber [65] is used to measure the incident beam profile. The system is triggered by a $40\times40\,\mathrm{mm}^2$ scintillating screen in front of the absorber plates. The size of the trigger counter allows to illuminate several pads of the sensor without moving the system.

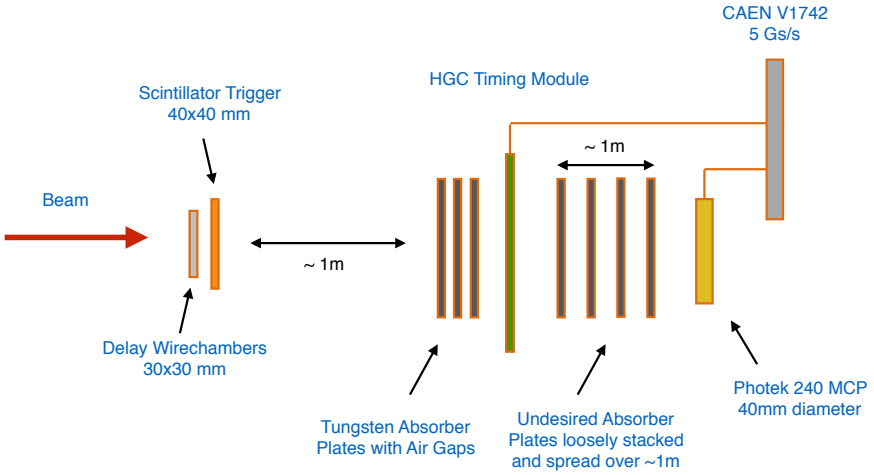


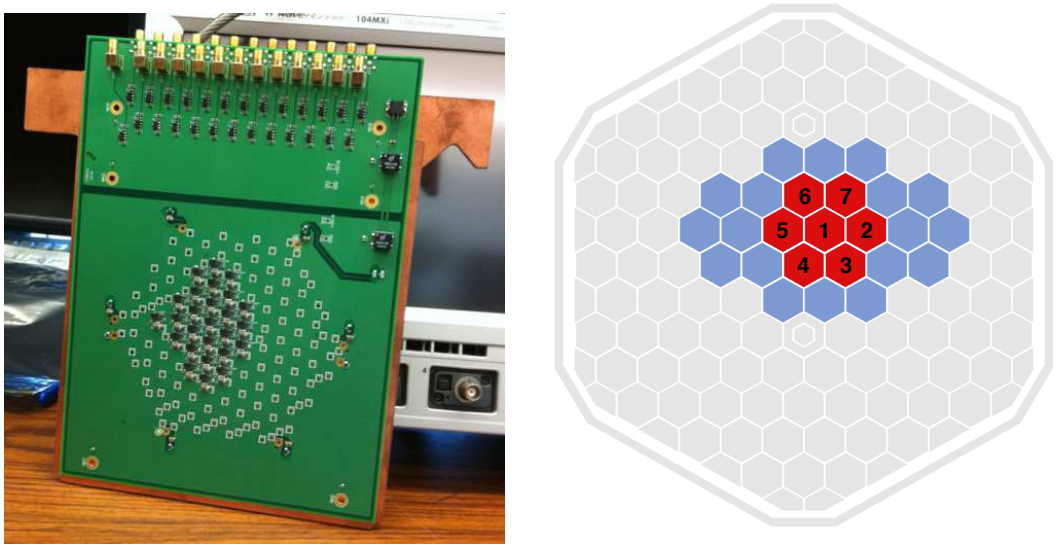
Figure 6.1: The experimental setup. The electromagnetic showers produced by incident electrons of 32 GeV to 250 GeV are sampled by a silicon pad sensor equipped with a fast electronics and digitised at 5 Gs/s.

The reference timestamp is obtained by a Photek 240 micro-channel plate (MCP) with a 40 mm diameter, placed further downstream of the setup. The resolution of this detector has been measured previously to be around 7 ps [66] when placed immediately downstream of the main detector module. In the presented setup however, the MCP is placed about a metre away from the silicon sensor, with about 10 X_0 of additional, loosely stacked absorber material in between.

6.1.1 The Timing Module

The timing module can be seen in Fig. 6.2(a). It consists of a 6in 135ch n-type HGCAL prototype sensor of 300 µm active thickness and a pad size of approximately 1.1 cm^2 (see Sec. 5.1). The detector was operated at 300 V, around 100 V above full depletion. Of the 135 sensor pads, 25 are read out via a custom amplifier chain with a total gain of 100. The other pads are not connected and are therefore on floating potential. Of the 25 read out pads (marked blue in Fig. 6.2(b)), only 7 are used for most of the analysis due to limited statistics. These pads are marked in red

on Fig. 6.2(b). The voltage signal is sampled at 5 Gs/s by a CAEN V1742 digitiser [67] with a 12 bit ADC. The total dynamic range of the system is 1 V.



(a) (b)

Figure 6.2: The dedicated liming module designed for fast readout of an HGCAL sensor is shown in (a). The silicon pads are wire-bonded to the PCB. The sensor layout is shown in (b), facing the beam. All blue and red pads are read out. Red indicates the central seven pads that are used for large parts of the analysis.

The CAEN module comprises four DRS4 ASICs [68] with 8 channels each and an analogue bandwidth of 500 MHz. The intrinsic time resolution between two channels on the same DRS4 ASIC has previously been measured to be 5 to 7 ps [66]. The resolution between two ASICs is significantly worse. Therefore, the MCP signal was fed into channel 0 of each ASIC to synchronise them with the time resolution of the MCP itself. The central 7 silicon pads that are mainly used in this analysis have all been read out by the same ASIC.

The schematic circuit diagram of a single readout channel is shown in Fig. 6.3. Each pad is equipped with an amplifier placed very close to the wire-bonded readout. A 1:2 transformer is used to lower the effective input impedance and therefore provide a faster rise time of the signals. An additional amplifier and bandpass with 200 MHz is placed at the periphery of the PCB.

The readout chain was originally designed for a previous testbeam at FNAL with up to 32 GeV electron beams. At the CERN H2 beamline electrons with up to 250 GeV are provided. To avoid saturation of the digitiser, 10 dB attenuators were placed between the output of the readout board and the digitiser input.

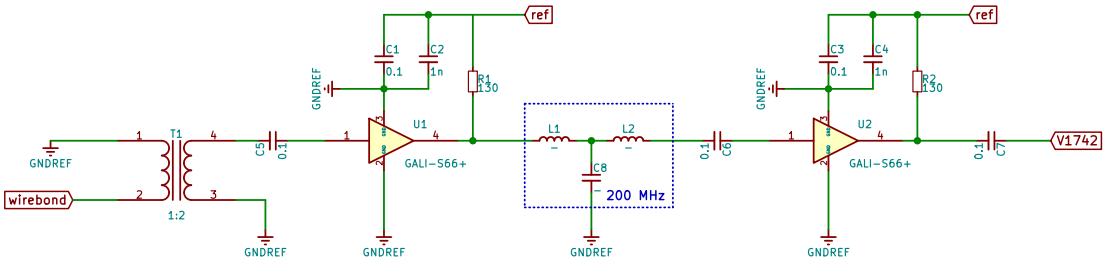


Figure 6.3: The schematic circuit diagram of the readout electronics for one channel of the timing module.

6.1.2 Waveform Analysis

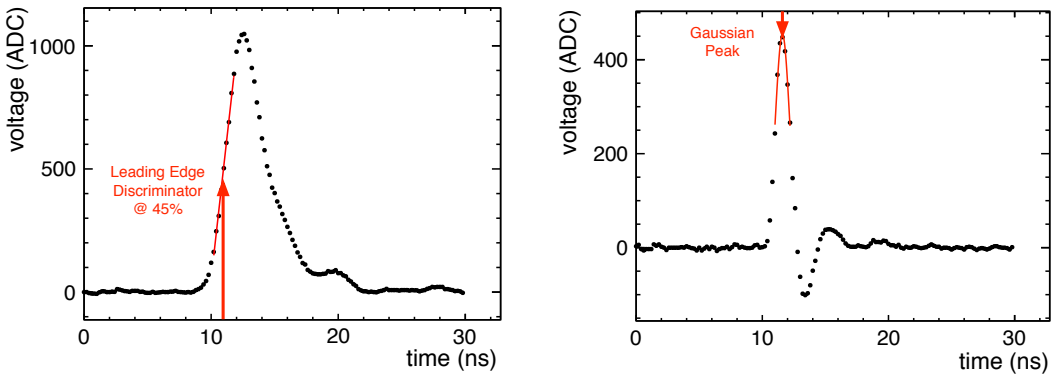
Examples of the recorded waveforms are shown in Fig. 6.4. After an external trigger is received, 1024 time samples corresponding to 200 ns are recorded. The recorded waveforms from the silicon sensor and MCP are then analysed per channel and event with the following procedure.

1. Find the global maximum of the waveform.
2. Establish the peak region as the 30 samples (6 ns) before and 70 samples (14 ns) after the maximum.
3. Establish the baseline and noise of the waveform as the mean and standard deviation of all recorded samples before the peak region.
4. Subtract the baseline on event-by-event and channel-by-channel basis.
5. Extract the pulse amplitude as the pulse maximum.
6. Extract the pulse integral as the sum of the signal in the window of 20 samples (4 ns) before and 25 samples (5 ns) after the maximum.
7. For the MCP waveforms, fit a Gaussian function between three samples (0.6 ns) before and three samples after the peak. Take the mean value as timestamp.
8. For the silicon waveforms, fit a linear function between 20% and 90% of the rising edge. Extract the rise time as 10% to 90% of the leading edge and take a constant fraction discrimination at 45% as timestamp.

6.1.3 Beam Profiles and Event Definition

In total, six electron beam energies between 32 and 250 GeV are used. The beam profiles measured with the wire chamber for all energies can be seen in Fig. 6.5. In all cases, the beam spot is roughly focused on the central pad of the silicon sensor (pad 1, see Fig. 6.2(a)).

A valid event is defined as a signal in the scintillator that triggered the readout, a reconstructed MCP waveform and at least one reconstructed silicon waveform. Individual waveforms are reconstructed if their amplitude surpasses a noise threshold of 0.01 mV. In addition, it is required that at least 50% of the event energy is deposited in the central seven pads that are used for the timing analysis. This condition is necessary as the trigger scintillator extends beyond the coverage of the seven central pads.

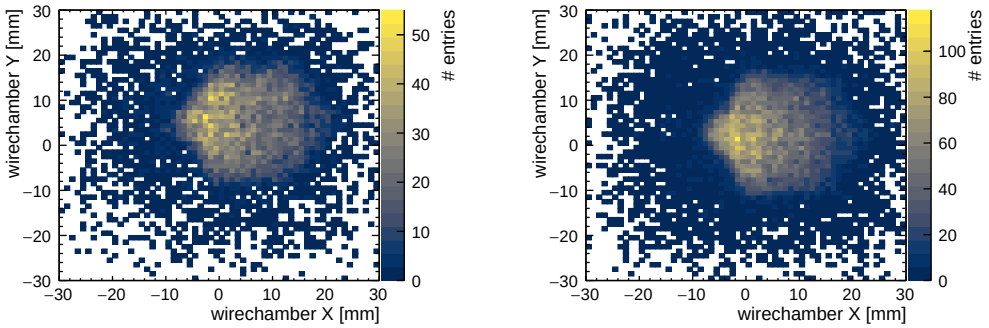


(a) (b)

Figure 6.4: Example waveforms from the silicon sensor (a) and the MCP (b). For the silicon sensor, the reference timestamp is extracted as 45% of the peak value. For the MCP, a Gaussian function is fitted around the peak and the mean is taken as timestamp.

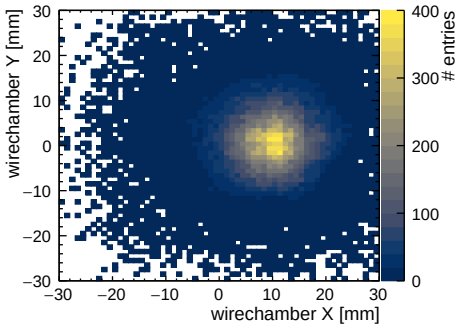
Table 6.1: Number of reconstructed events per beam energy.

Particle ID [-]	Beam Energy [GeV]	Absorber Setting [-]	Events [-]
Electrons	32	5.7 X_0 W	15350
Electrons	50	5.7 X_0 W	30939
Electrons	100	5.7 X_0 W	70406
Electrons	150	5.7 X_0 W	50850
Electrons	200	5.7 X_0 W	103898
Electrons	250	5.7 X_0 W	70291

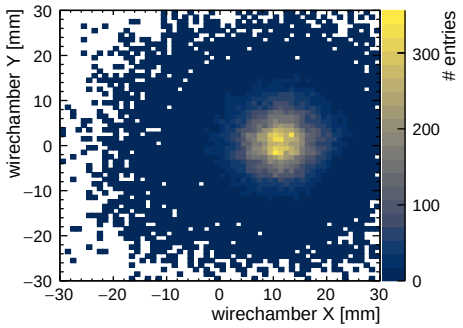


(a) 32 GeV

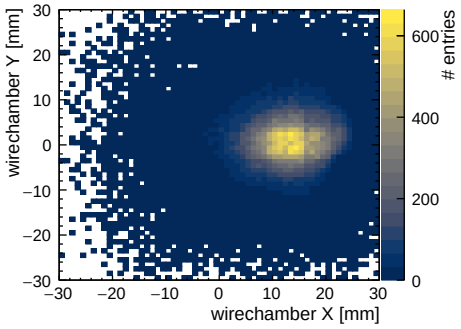
(b) 50 GeV



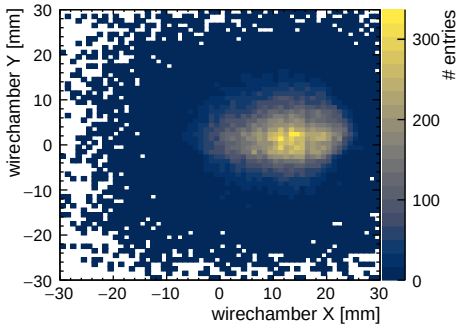
(c) 100 GeV



(d) 150 GeV



(e) 200 GeV



(f) 250 GeV

Figure 6.5: Beam profiles from the wire chamber.

Fig. 6.6 shows the position of the pad with the largest energy deposit as a function of the X/Y measurement of the wire chamber. In this reference system, the central pad is approximately centred at (x,y) = (+8,+6) mm.

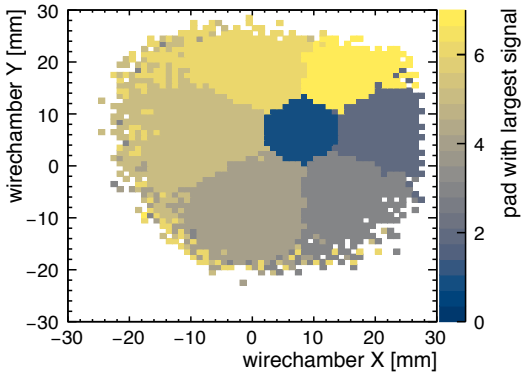


Figure 6.6: Position of the pad with the largest energy deposit in the reference frame of the wire chambers.

6.2 Basic Signal Properties

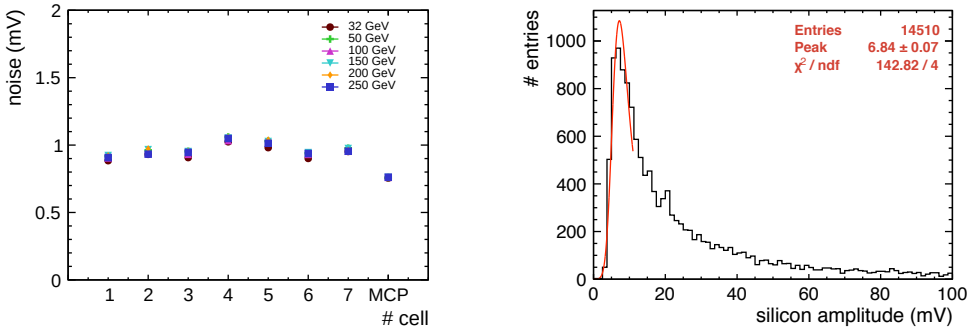
6.2.1 Noise and MIP response

The extracted noise values for all pads and beam energies can be seen in Fig. 6.7(a). They are used to determine the S/N ratio for each event. The average value is about 1 mV per pad. Common mode noise of around 4 mV is present but it is not effecting the analysis as the baseline is established and subtracted on an event per event basis.

The energy deposition of 120 GeV pions is shown in Fig. 6.7(b). Unfortunately, these data were taken with 9.5 X_0 of Pb upstream of the setup and the distribution differs somewhat from the typical Landau distribution due to the possibility of early showering. Nevertheless, a rough energy calibration can be attempted. The Landau peak is at 6.8 mV, resulting in a S/N ratio of around 6.8 for MIPs. Taking the active thickness of 300 µm and a most probable energy loss of 78 e- per µm (see Tab. 2.1), the noise N of the system is approximately 3500 e-. However, due to the large systematic uncertainty on this value, the absolute energy calibration is not used further in this work and all results are quoted in units of S/N.

6.2.2 Rise Time

The rise time of silicon signals is around 1.8 ns as can be seen in Fig. 6.9(a) for the 100 GeV data set. The MCP signals have shorter rise times of around 700 ps as shown in Fig. 6.9(b). Using



(a) noise per pad

(b) response of 120 GeV pions

Figure 6.7: (a) Noise levels per channel. Values are extracted as the standard deviation of all recorded waveform samples before the peak region. (b) Energy deposition of 120 GeV pions after 9.5 X_0 of Pb. A Landau function convoluted with a Gaussian is fitted to the beginning of the spectrum to limited the influence of early showering.

(see Sec. 2.2.1)

$$t_{\text{rise}} = 2.2\tau = \frac{0.35}{f_{\text{u}}}, \tag{6.1}$$

this is in excellent agreement with the expectations from $f_u = 200$ MHz of the silicon readout board and the $f_u = 500$ MHz bandwidth of the CAEN module.

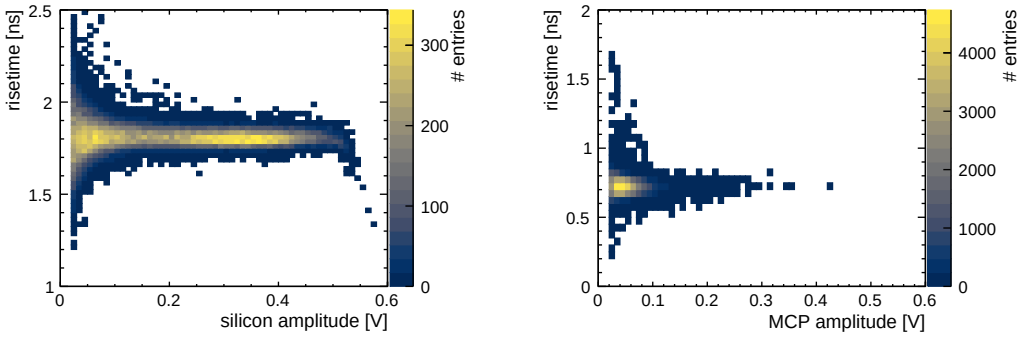
6.2.3 Saturation

Fig. 6.9 shows the waveform amplitude as a function of the waveform integral of the reconstructed waveform for 150 GeV incident electron energy after 5.7 X_0 of W. It can be seen that even though the dynamic range of the digitiser is 1 V, the amplitude-to-integral ratio becomes non-linear above an amplitude of 0.45 V. This indicates a saturation at the pre-amplifier level. Saturated signals are cut off at the top and when extracting the timestamp at 45% of the amplitude, they are sampled earlier than non-saturated pulses. Such signals are observed for beam energies of 150 GeV and above.

6.2.4 Time Measurement

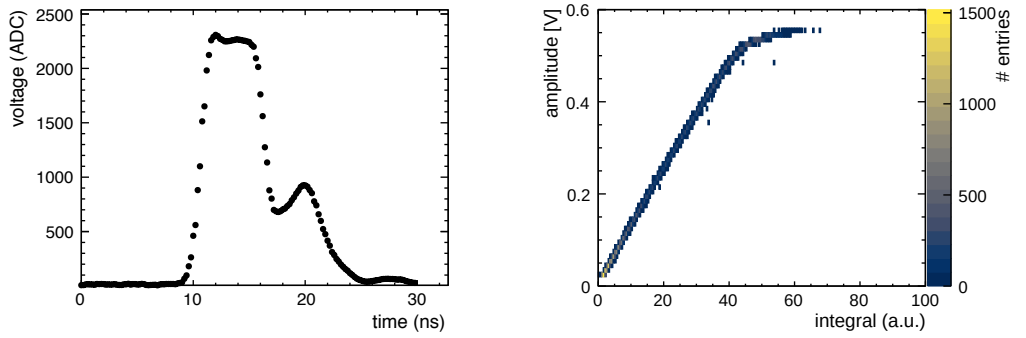
After the timestamps of the timing module and the MCP are extracted by the procedure described in section 6.1.2, two time differences are calculated: the time difference relative to the MCP, Δt_{mcp}, and the time difference between adjacent pads within a cluster on the silicon sensor, Δt_{si}. For the latter, the pad with the largest signal (that is not saturated) inside a cluster is taken as reference.

$$\Delta t_{\text{mcp}} = t_{\text{mcp}} - t_{\text{pad,i}} \tag{6.2}$$



(a) central silicon pad (b) MCP

Figure 6.8: Rise time as a function of the amplitude for the central most pad (a) and the Photek MCP (b) at 100 GeV beam energy. The shorter rise time of the MCP is expected as it is directly sampled with a bandwidth of 500 MHz while the silicon signal passes through an additional 200 MHz filter.

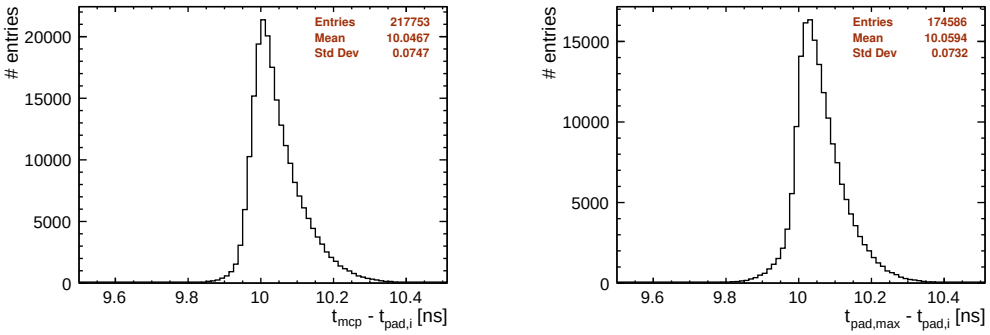


(a) (b)

Figure 6.9: Saturation of the amplifier chain for (a) a single example waveform and (b) the waveform amplitude as a function of the waveform integral for the central pad at 150 GeV after 5.7 X_0 of W. Saturation is visible above 0.45 V which corresponds to an ADC value of $\approx$ 1850.

$$\Delta t_{\mathrm{si}} = t_{\mathrm{pad,max}} - t_{\mathrm{pad,i}} \tag{6.3}$$

In the following, these two quantities are analysed separately. Example distributions for the central most pad at 100 GeV beam energy are shown in Fig. 6.10. Deviations from a Gaussian shape are discussed in the following section. When using Δt_{si}, no entry can be filled to the histogram for the pad that is used as reference. Therefore, fewer entries are available for these histograms.



(a) Δt_{mcp} (b) Δt_{si}

Figure 6.10: Examples of single pad time distributions relative to (a) the MCP and (b) the pad with the largest signal within a cluster. For 100 GeV beam energy after 5.7 X_0 of W.

6.3 Calibration of the Time Response

6.3.1 Delay Calibration

Each silicon pad has an offset in time. Extracting this parameter with picosecond precision is crucial for the inter-pad synchronisation. Here, a data driven approach is taken and the delay values are extracted from the single pad time distributions using Δt_{mcp}.

To avoid bias from an amplitude dependent time response (see Sec. 6.3.2), only signals with an amplitude between 0.25 V and 0.45 V are used to obtain the delay. For each beam energy of 100 GeV and above, a Gaussian is fitted to the Δt_{mcp} distribution. The resulting peak positions of all data sets are then combined to estimate the delay value. The method and threshold are optimised to obtain the smallest standard deviation between the data sets. The extracted delay values and their uncertainty can be found in Tab. 6.2. All timestamps are corrected with these values to a common reference value. An absolute time-of-flight calibration is not attempted.

For pads outside of the first ring, the limited number of events with significant large signals prevents a reliable extraction of the delay values. These pads have therefore been discarded from further analysis.

Pad Nr. [-]	Delay [ns]	Uncertainty [ns]
Pad 1	10.1171	0.0018
Pad 2	10.3047	0.0012
Pad 3	10.0458	0.0011
Pad 4	10.1529	0.0010
Pad 5	10.1076	0.0016
Pad 6	10.0160	0.0016
Pad 7	10.1916	0.0036

Table 6.2: Delay values of the different pads. The values are extracted by fitting a Gaussian to the Δt_{mcp} distributions including all signals between 0.25 V and 0.45 V per pad. Each beam energy of 100 GeV and above is treated separately. The results are then averaged to extract a common delay value.

6.3.2 Amplitude Dependence of the Time Response

Compared to the situation with Timepix3 in Sec. 4.2, no constant value threshold is used in this analysis. Therefore no timewalk due to the frontend electronics is expected. However, effects of lateral shower evolution as well as the observed signal saturation result in a similar effect.

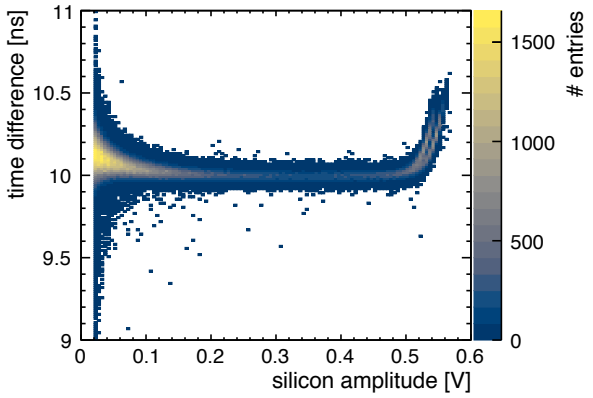
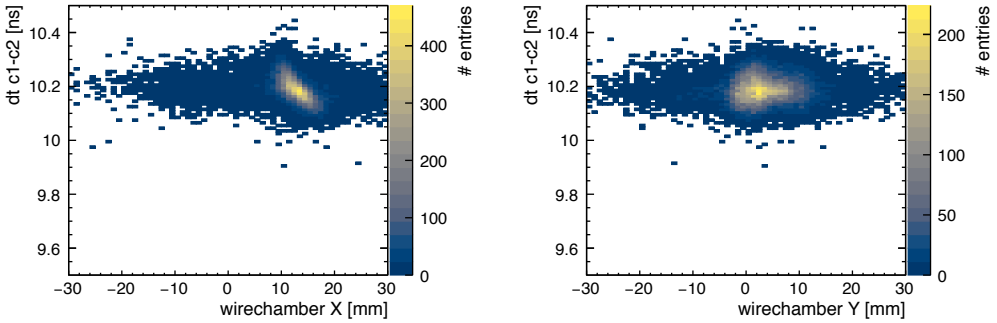


Figure 6.11: Amplitude dependence of the time difference including all pads for Δt_{mcp}. For 200 GeV beam energy.

Fig. 6.11 shows the measured time difference Δt_{mcp} as a function of the signal amplitude for 200 GeV. At high amplitudes above approximately 0.45 V, saturation of the signal (see in Sec. 6.2.3) results in an amplitude dependent delay. At low amplitudes, time-of-flight delays in the lateral shower evolution cause a similar deviation. That this effect is indeed caused by time-of-flight can be seen if one looks at the time difference between two pads as a function of the distance to the central shower axis. Fig. 6.12 shows the time difference between pad 1 and pad 2, which are located on similar Y coordinates but different X coordinates (see Fig. 6.2(a)),

as a function of the wire chamber impact position. A clear dependence on X can be seen while the time difference is independent of Y. Analogous behaviour can be observed for other pads.



(a) (b)

Figure 6.12: Delay between pad 1 and pad 2 in (a) X direction and (b) Y direction. Both pads are located on similar Y coordinates but different X coordinates.

The full amplitude dependence $\delta_{\text{amp}}(x)$ is extracted as the mean time difference per amplitude bin. For the high amplitude region, the pads can be separated in two groups with different saturation behaviour. This is also visible in Fig. 6.11. For the low amplitude region, no difference between the pads is observed as long as the delay correction is applied beforehand. The corrections are then applied separately for the two groups of pads via

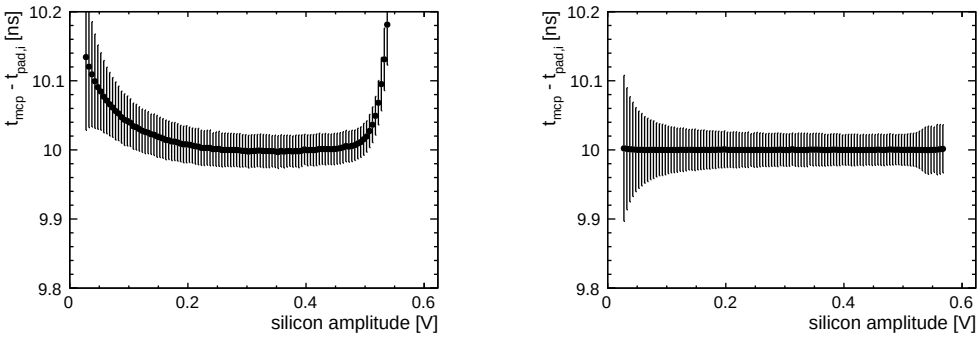
$$\Delta t'_{\text{mcp/si}} = \Delta t_{\text{mcp/si}} - \delta_{\text{amp}}(x) + \delta_{\text{common}}, \tag{6.4}$$

where x is the signal amplitude and δ_{common} a common shift. The two quantities, Δt_{mcp} and Δt_{si}, are treated independently. All beam energies are also treated separately.

Fig. 6.13 displays the effect of the correction. The mean and standard deviation of the time difference per amplitude bin before and after δ_{amp} correction are shown. With applied corrections, the mean time difference is independent of the amplitude. The standard deviation decreases with increasing amplitude until the regime of saturated signals is reached, at which point the standard deviation increases. Most likely this is due to an imperfect correction and even finer binning would be required. Due to limited statistics, this is however not feasible.

While the saturation correction should not be necessary in a final system, one might argue that the method of correcting for the time-of-flight effects cannot be applied outside of test beam. However, these corrections could also be applied based on a shower shape analysis that determines correction as a function of the distance d to the shower core, $\delta_{\text{amp}}(d)$. This in turn is not possible with this setup due to the limited lateral coverage. A correction based on Monte Carlo simulation would is also possible.

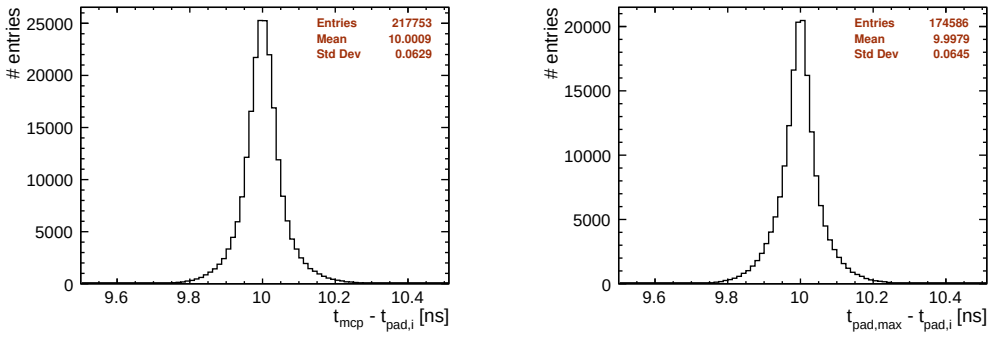
The effect of the corrections can be seen in Fig. 6.14 which shows the same single pad time distributions as in Fig. 6.10 after both delay and amplitude corrections. After correction, the standard deviations are reduced and the distributions are well represented by a Gaussian model.



(a) without δ_{amp} correction

(b) with δ_{amp} correction

Figure 6.13: Profiles of the time difference as a function of the amplitude before and after δ_{amp} correction. The data points correspond to the mean value per bin while the bars correspond standard deviation per bin. For 200 GeV beam energy.



(a) Δt_{mcp}

(b) Δt_{si}

Figure 6.14: Examples of single pad time distributions with applied corrections relative to (a) the MCP and (b) the most energetic pad in a cluster. For 100 GeV beam energy.

6.4 The Time Resolution of a Single Silicon Pad

As discussed in Sec. 2.4, the time resolution holds contributions from jitter and digitisation but also from the signal formation in the sensor. The latter can arise from variations in the energy deposition throughout the sensor as well non-uniform coupling of the charge to the readout electrodes. For example, in Low Gain Avalanche Detectors, such contributions often limit the achievable time resolution [23]. In silicon pad sensors, this contribution has yet to be established.

6.4.1 Effective Signal-to-Noise

To parametrise the time resolution, the signal-to-noise ratio (S/N) is a good figure of merit. For a large area silicon diode in a calorimeter environment, one can try to describe the uncertainty on a single time measurement with a two parameter model

$$\sigma_t = \sigma_\mathrm{j} \oplus \sigma_\mathrm{c} = \frac{a}{S/N} \oplus c. \tag{6.5}$$

The parameters a should be dominated by the contribution from noise and the finite signal rise time while c represents constant contributions from e.g. digitisation and the signal formation effects. A priori, this is an assumption and whether this model delivers an accurate description of the data has to be shown.

Consider now a time measurement between two generic detectors. Taking into account possible contributions from synchronisation, σ_synchr, and shower fluctuations, σ_fluct, the uncertainty on the measurement of the time difference is given by the square root quadratic sum of the two individual contributions,

$$\sigma^2_\mathrm{msr} = \sigma^2_{t1} + \sigma^2_{t2} = \sigma^2_\mathrm{j1} + \sigma^2_\mathrm{c1} + \sigma^2_\mathrm{j2} + \sigma^2_\mathrm{c2} + \sigma^2_\mathrm{synchr} + \sigma^2_\mathrm{fluct}. \tag{6.6}$$

As mentioned previously, the synchronisation across the 8 channels of one DRS4 chip is assumed to be better than 7 ps [66]. In addition, the uncertainty on the delay value has to be considered.

The first two terms of Eq. 6.6 can be rewritten to

$$\sigma^2_\mathrm{j1} + \sigma^2_\mathrm{j2} = \frac{a_1^2}{(S/N)_1^2} + \frac{a_2^2}{(S/N)_2^2} = a_1^2 \frac{(S/N)_2^2 + a_2^2/a_1^2 \cdot (S/N)_1^2}{(S/N)_1^2 \cdot (S/N)_2^2}. \tag{6.7}$$

By introducing an effective signal-to-noise quantity

$$(S/N)_\mathrm{eff} = \frac{(S/N)_1 \cdot (S/N)_2}{\sqrt{(S/N)_2^2 + \alpha^2 \cdot (S/N)_1^2}}, \tag{6.8}$$

where α is the ratio of the a parameters that can be approximated as the ratio of the rise times,

$$\alpha = \frac{a_2}{a_1} \approx \frac{t_\mathrm{rise,2}}{t_\mathrm{rise,1}}, \tag{6.9}$$

one can then rewrite Eq. 6.6 to a fit-able function

$$\sigma_\mathrm{msr} = \sqrt{\frac{A^2}{(S/N)^2_\mathrm{eff}} + C^2}. \tag{6.10}$$

where $A = a_1$ and $C = \sqrt{c_1^2 + c_2^2 + \sigma_{\text{synchr}}^2 + \sigma_{\text{fluct}}}$.

For each time difference, an effective S/N quantity is calculated according to Eq. 6.8. The time differences are then sorted into S/N bins and the standard deviation is calculated. A minimum of 300 events per bin is requested. In the case of two silicon pads, one can assume equal rise times and set $\alpha = 1$. For the measurement relative to the MCP, this factor is calculated from the measured rise time on event-by-event basis. Finally, the data are fitted with Eq. 6.10.

From Fig. 6.13(b) it is visible that the measurement resolution degrades for saturated signals. To avoid entanglement of this effect with the contributions from the system of frontend electronics and sensor, signals above 0.45 V are excluded for the results in this section. Effectively, this reduces the available number of high S/N signals but does not introduce any bias on the result as it is quoted as a function of S/N.

6.4.2 Time Difference Between Adjacent Pads

The S/N dependence of the time difference between adjacent pads in a cluster, Δt_{si} as defined in Eq. 6.3, is shown in Fig. 6.15. For S/N values above 200, uncertainties below 15 ps are reached. The four data sets agree well with each other and the two parameter model appears to be an excellent representation of the data. A slight trend towards smaller A and larger C values for higher beam energies can be observed. This trend is likely to be introduced by the fitting procedure due to the reduced lever arm for lower beam energies. The data points of 32 and 50 GeV have not been fitted due to the missing number of events at high S/N values.

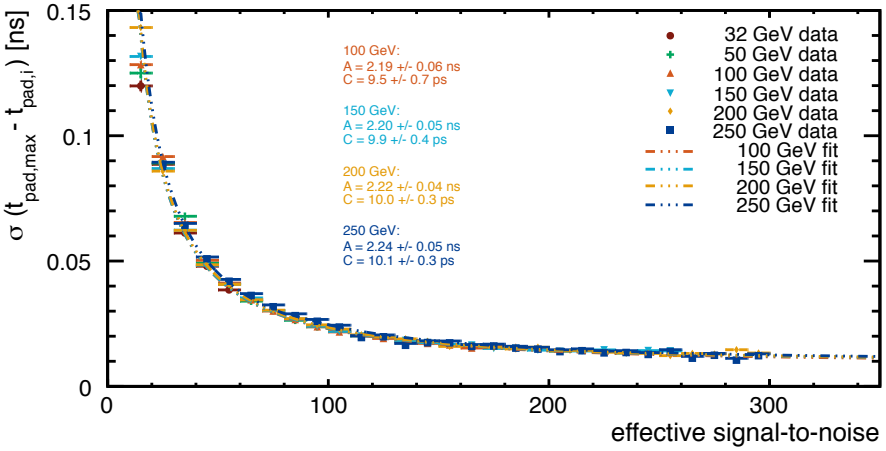


Figure 6.15: Measurement uncertainty of the time difference between adjacent pads, Δt_{si}, as a function of the effective S/N. The four different data sets are fitted with the function described in Eq. 6.10.

The averages over all four data sets are

$$A = (2.21 \pm 0.06) \text{ ns}$$

$$C = (9.9 \pm 0.5)\ \text{ps}$$

Assuming equal constant terms for all pads, the single pad constant term can be at most $C/\sqrt{2}$ or

$$c_{\text{si}} \leq (7.0 \pm 0.4)\ \text{ps}.$$

Considering also the expected synchronisation uncertainty of 5-7 ps, one can conclude that lateral shower fluctuations and non-uniform signal formation in the silicon sensor contribute at most a few picoseconds to the time resolution of silicon pads of $\sim 1\,\text{cm}^2$ size when sampling electromagnetic showers.

6.4.3 Time Difference Relative to the MCP

The time difference relative to the MCP, Δt_{mcp} from Eq. 6.2, is shown in Fig. 6.16 as function of the effective S/N. The constant term C of around 21 ps is worse than what is observed for Δt_{si}. Averaging over the four data sets and subtracting c_{si} in quadrature, the constant contribution to the time resolution of the MCP is

$$c_{\text{mcp}} \leq (19.6 \pm 0.3)\ \text{ps}.$$

Thus the intrinsic resolution of the MCP is less precise than that of the silicon. It is also worse than what is expected from previous measurements [66]. A possible explanation could be the additional, sparsely stacked absorber material between the silicon sensor and the MCP, which was not present in the setup reported in [66]. The lateral shower fluctuations arising from this configuration could limit the time resolution of the MCP as not all channels on the MCP are hit simultaneously.

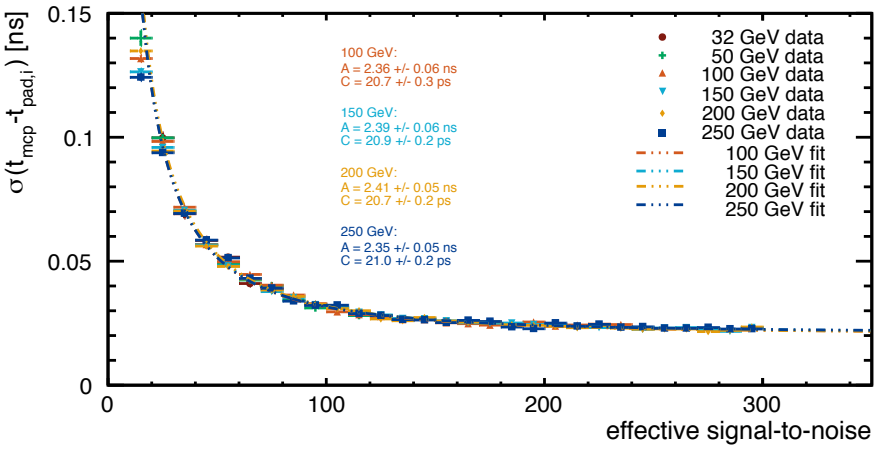


Figure 6.16: Measurement uncertainty of the time difference relative to the MCP, Δt_{mcp}, as a function of the effective S/N.

In consequence, a time measurement relative to the MCP does not allow to access the best achievable timing performance with silicon pad sensors. However, the MCP acts in a similar way as the digitisation with a TDC of O(50 ps) binning and is thus an interesting test case for a realistic system.

6.5 The Time Resolution of a Cluster

From the last section one can conclude that the uncertainty of the time measurement depends on the S/N ratio. The question then arises whether one can improve the time resolution per calorimeter layer by combining multiple timestamps within a cluster.

6.5.1 Combining Timestamps from Multiple Pads

Per event, direct neighbouring pads with a S/N value larger than 10 are considered a cluster. A timestamp t_{cluster} is associated with the cluster by combining all timestamps from neighbouring pads that are above a certain S/N value. The individual timestamps are then weighted by the inverse variance according to the results from Fig. 6.13(b).

$$t_{\text{cluster}} = \frac{\sum_i t_{\text{pad,i}} \cdot 1/\sigma_i^2}{\sum_i 1/\sigma_i^2} \tag{6.11}$$

Other weighting methods have been tried but delivered inferior results. Should no pad be above the threshold, the timestamp from the most precise pad is used. Then, the time difference to the MCP is calculated.

$$\Delta t_{\text{mcp,cluster}} = t_{\text{mcp}} - t_{\text{cluster}} \tag{6.12}$$

6.5.2 Dependence on Cluster Threshold

The condition whether a pad contributes to a cluster timestamp or not, is established via a S/N threshold. A lower threshold incorporates more pads but may include pads with low S/N ratio while a high threshold only uses information from a few high resolution pads. If no pad passes the threshold, the timestamp of the pad with the highest S/N is used. It can be seen from Fig. 6.17, that there is no significant dependence of the cluster time resolution on the S/N threshold. A cluster time uncertainty of better than 30 ps is achieved for beam energies above 100 GeV. For all further results, a S/N threshold of 30 is used.

6.5.3 Dependence on Cluster Energy

The dependence of cluster timing on the cluster energy is shown in Fig. 6.18. It is shown in units of effective cluster S/N. This quantity is the sum of all effective S/N values in the cluster, independent of the cluster threshold.

$$(S/N)_{\text{eff,cluster}} = \sum_i (S/N)_{\text{eff,i}} \tag{6.13}$$

For the cluster timestamp, uncertainties below 30 ps are accessible for all cluster S/N values above 200. To put this value into perspective, at the noise level that is foreseen for HGCAL of 2500 e- and a sensor thickness of 300 µm, a deposition of approximately 21 MIPs is necessary to reach that value. If the charge collection efficiency should drop to 50% due to irradiation damage, this increases accordingly to 42 MIPs - assuming the noise remains at a constant level.

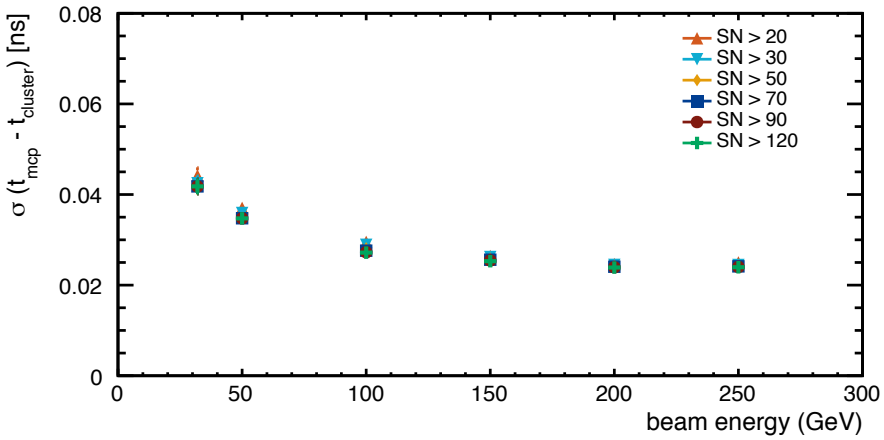


Figure 6.17: Cluster timing for different S/N threshold conditions. Only pads passing the threshold contribute to the cluster timestamp. If no pad passes the threshold, the timestamp of the pad with the highest S/N is used.

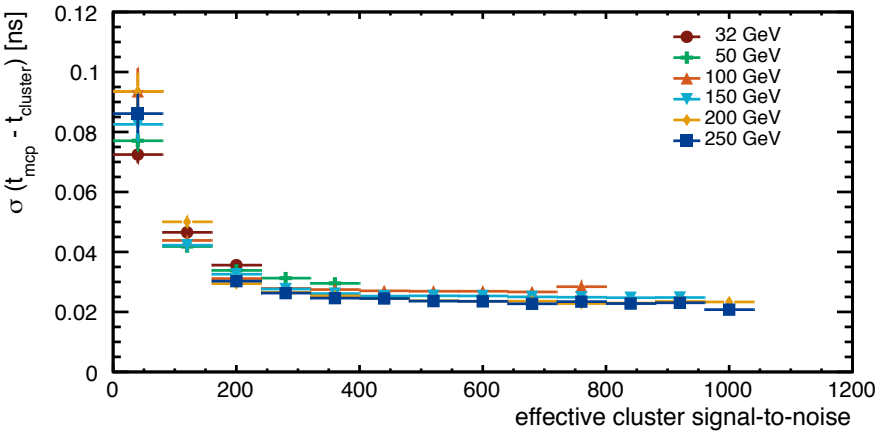


Figure 6.18: Cluster time uncertainty as a function of the effective cluster signal-to-noise. Uncertainties of 30 ps or better are accessible for all cluster S/N values above 200.

6.5.4 Dependence on Cluster Size

If the fluctuation of the lateral shower evolution is negligible, one could naively expect that a shower covering multiple pads would improve the uncertainty by a factor close to $1/\sqrt{n}$, with n being the cluster size. However, this is only true if all measurements are done with the same precision which is not the case. Fig. 6.19 shows the cluster time uncertainty for different cluster sizes. Only slight improvements with the larger cluster sizes are observed.

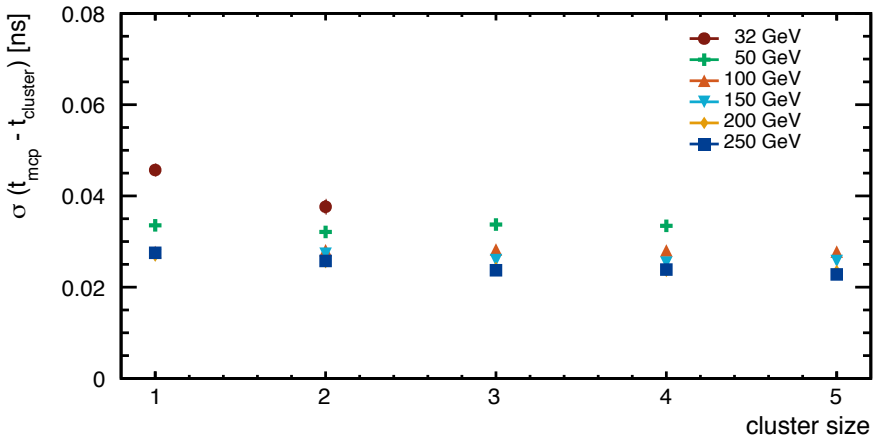


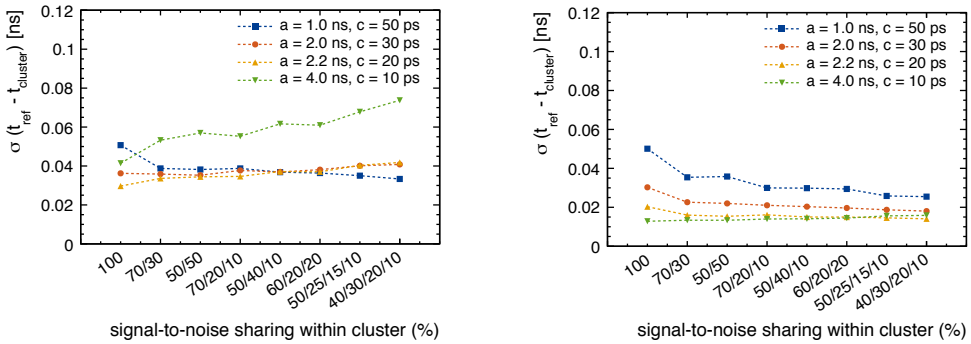
Figure 6.19: Cluster time uncertainty as a function of the cluster size.

Whether the cluster time resolution is dominated by the pad with the largest signal or does indeed gain from larger cluster sizes, depends on the a and c parameters from Eq. 6.5 and the S/N distribution between pads. The latter is in turn determined by beam energy, absorber configuration and sensor geometry. In the following, this effect is illustrated with a small toy model.

A Toy Model of the Cluster Timing

Assume a set of pads with various S/N values and simultaneous hit time t_0. The hit time is then convoluted with an uncertainty according to Eq. 6.10 and the resulting timestamps recombined according to Eq. 6.11. This is repeated 10000 times for each data point.

The effects for different energy sharing ratios between cells within a cluster can be seen in Fig. 6.20 for four different sets of a and c parameters and two different values of cluster S/N. For large a terms and small c terms, the pad with the largest signal dominates and the time resolution degrades with larger cluster sizes at the same cluster energy. On the opposite, a small a term and large c term profits from larger cluster sizes and equal sharing between the cells. In the case of $a = 2.2\,\text{ns}$ and $c = 20\,\text{ps}$ that corresponds to the Δt_{mcp} case, larger cluster sizes at a given S/N ratio slightly degrade the time resolution at smaller S/N values. At larger S/N values slight improvements can be observed.



(a) cluster S/N 100

(b) cluster S/N 500

Figure 6.20: Simulation results of the cluster time resolution for different cluster sizes and energy sharing ratios between cells for two different cluster energies.

6.5.5 Uniformity of the Cluster Timing

Another interesting property is the uniformity of the time resolution. Fig. 6.21 shows the uncertainty on the cluster time in relation to the wire chamber position for the 200 GeV data set. The central silicon pad is approximately centred at (+8,+4) mm. No dependence on the impact position can be observed.

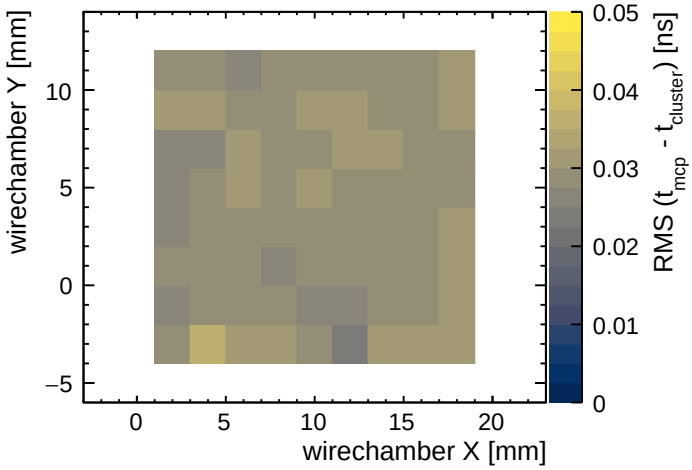


Figure 6.21: Cluster time uncertainty relative to the impact position for 200 GeV beam energy.

6.6 Towards a Full System Timing

The previous sections have shown that neither the silicon sensor itself nor lateral shower fluctuations have a significant contribution to the time resolution at the level of 30 ps. However, the readout system used in the previous sections was designed for the purpose to study the achievable time resolution of HGCAL on a single module level. It is unlikely to have any application in a high luminosity collider experiment, mainly due to its high power consumption of about 40 mW per channel, limited rate capabilities of less than 1 kHz and limited dynamic range of 12 bit. For HGCAL, the frontend will require a 15 bit dynamic range that records data at 40 MHz. A power budget of 14 mW per channel is foreseen [41].

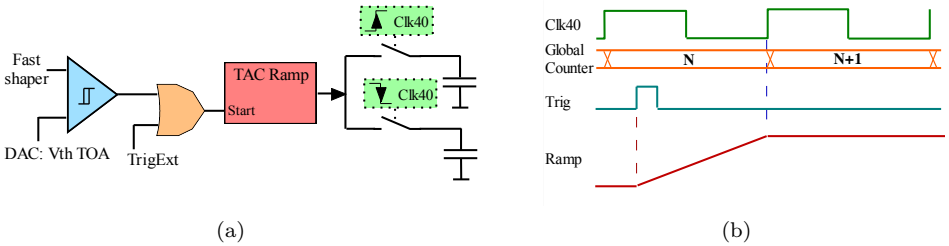
To achieve this, the final ASIC, called HGCROC, will employ a time-of-arrival (ToA) method for the time measurement. While the implementation will likely be different from Timepix3 (see Sec. 4.1) - due to the much finer resolution that is required - the method will be based on a constant value threshold. Nevertheless, the method will show similar non-linearities and require a careful calibration. Compared to a waveform sampling method, a ToA approach does not allow common mode correction. It is therefore essential, that the magnitude of the (coherent) common mode noise is reduced to values well below the incoherent noise. Furthermore, to reach 30 ps also on a system level, the synchronisation of all frontend channels has to be significantly lower than the envisaged resolution. On the one hand, this depends on the system wide calibration of the clock itself. The ongoing effort to achieve a long term stable precision of 10 to 15 ps is described in [41, Sec. 8.4]. On the other hand, the clock delay of all channels has to be determined with similar accuracy.

A first test towards a system timing has been done with prototype HGCAL modules and the Skiroc2CMS ASIC.

6.6.1 Time Calibration of the Skiroc2CMS ASIC

Skiroc2CMS is the first prototype of an HGCAL specific ASIC and is used for test beam purposes. It comprises 64 frontend channels implemented in AMS 0.35 µm SiGe technology. Each channel can measure energy and arrival time of incident particles. Due to the large dynamic range of energy deposits that is expected per pad in HGCAL, the energy measurement is realised in three stages: two are based on waveform sampling with 40 MS/s and one is based on time-over-threshold (ToT). Additionally, a fast shaping stage dedicated for time measurements based on a time-of-arrival (ToA) method is provided. More details on this ASIC can be found in [69]. For the following discussion however, only the ToA stage is relevant.

The principe of the ToA implementation is shown in Fig. 6.22. The ASIC is clocked at 40 MHz. If the input signal rises above a constant threshold value, the global counter value is recorded and a time-to-amplitude converter (TAC) starts to charge two capacitors. The charging of the capacitors is stopped on the next falling and the next rising clock edge, respectively. In this process, the first clock edge after the crossing of the threshold is ignored to increase the charging time to a minimum of 12.5 ns. The capacitors are consequently read out by an 12 bit ADC. The ADC values then correspond to a charging time between 12.5 ns and 37.5 ns (0.5 to 1.5 clock periods) and represent the clock phase. The two ToA values, called *ToA Rise* and *ToA Fall*, are foreseen to reduce the uncertainty introduced by the jitter of the 40 MHz clock which provides



(a) (b)

Figure 6.22: Principle of the time measurement in Skiroc2CMS. Two measured values called ToA rise and ToA fall are provided. See the text for further explanations. Taken from [70].

the stop signal. For this work, only the ToA Rise value has been used.

The full range of 25 ns charging time corresponds to typically a bit less than 1500 ADC values, with the exact ADC ranges varying from channel to channel. In principle, this should allow to measure the hit time with a binning of approximately 17 ps. However, the ADC response shows a non-linear dependence on the clock phase. To fully calibrate the ToA response from ADC values to time, the following characteristics have to be established for each channel:

- The maximum and minimum ADC values that correspond to a charging time of 12.5 ns and 37.5 ns.
- The non-linear relationship between measured ADC value and clock phase.
- The non-linear relationship between measured ADC value and signal amplitude (timewalk).
- The clock delay.

Trying to build upon the success of the calibration procedure for Timepix3 shown in Chapter 4, a first calibration was attempted using test pulse injection. This is shown in Fig. 6.23 for a module consisting of four different ASICs that read out one prototype HGCAL sensor (see Sec. 5.1). Pulses of fixed amplitude are injected into one channel per ASIC with a programmable delay in steps of 3.125 ns. The delay is synchronous to the clock. The dependence of the clock phase on the ToA value is fitted with four free parameters a, b, c, d and the function

$$f(x) = a \cdot x + b - \frac{c}{x - d}. \tag{6.14}$$

The problem with this approach is that only the relative position of the injected pulses on the Y axis are known while the delay for each channel, and thus the absolute position, is not. Also, the minimum and maximum values of the ADC can not be established this way.

A possible solution is the use of an asynchronous pulse with a delay that is random and continuous with respect to the ASIC clock and covers the full range of 25 ns. The clock phase is then uniformly sampled and, due to non-linear relationship, large ToA values will be more frequent than lower ones. The resulting frequency of ToA values gives the derivative distribution of the ToA response function. This approach has been proposed by A. Lobanov [71] and is illustrated in Fig. 6.24. In Fig. 6.24(a) the frequency distribution of the ToA values is shown. The integral of this distribution represents the sought-after response curve. The distribution also gives the minimum and maximum ADC values, i.e. the lowest and highest measured ADC value above a certain noise threshold. With those values in hand, the ADC vales can be projected into a range

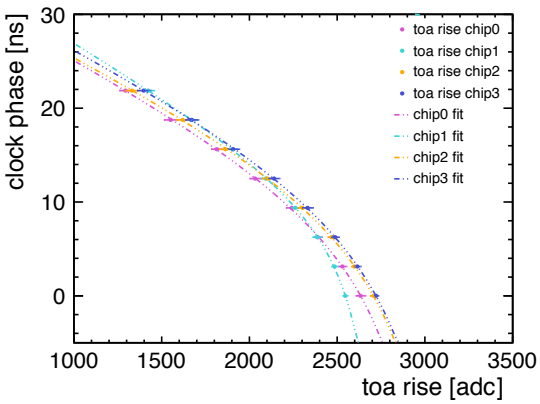


Figure 6.23: First calibration attempt of the time response of Skiroc2CMS with test pulses. The finite and synchronous delay settings of the injected pulse prohibit a precise calibration.

of [0,1] and the integrated frequency values into a range of [0,25]. The response function $f(x)$ can then be reduced to three free parameters

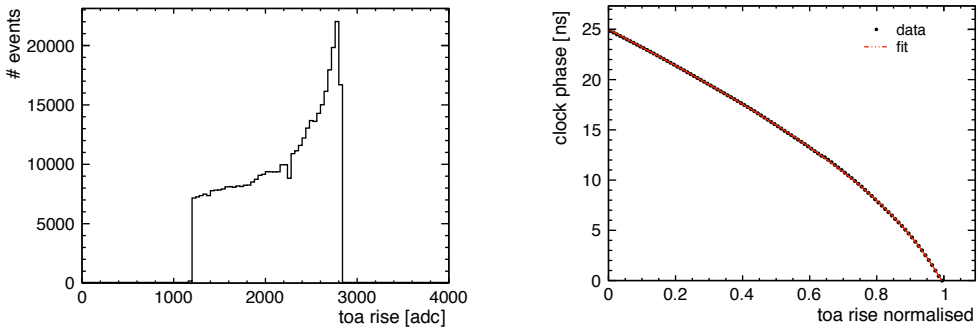
$$f(x) = a \cdot x + 25 - c\Big(\frac{1}{x-d} + \frac{1}{d}\Big). \tag{6.15}$$

The resulting fit is shown in Fig. 6.24(b). Similar methods are commonly used in computational physics, e.g. for numerical integration with Monte Carlo methods. It should be remarked that the non-linear timewalk does not interfere with this procedure if the hit time is truly uniform over the 25 ns time window. Timewalk will simply shift the threshold crossing to later times without affecting the uniformity. This method is also independent of the origin of the pulse or the amplitude, assuming it is above the ToA threshold. It does however require a large number of events to be accurate.

Such an asynchronous signal can be acquired in test beam, e.g. if the particle spills are uncorrelated with the readout system. In beam tests at the CERN North Area, enough data have been accumulated for a few channels to be calibrated with this method. In a setup very similar to Fig. 6.1, an MCP gives a reference timestamp and a set of HGCAL prototype modules equipped with Skiroc2CMS readout is used to sample electromagnetic showers from incident electrons. Using the calibration method described above, the global timestamp can then be constructed via

$$t_{\text{pad}} = t_{\text{course}} + \text{clock phase} - \text{timewalk}, \tag{6.16}$$

where t_{course} refers the global counter value of the 40 MHz clock. The energy calibration, ranging over the three gain stages, is still being optimised. The procedure is similar to the description in [64]. The timewalk is calibrated relative to the MCP. To limit time-of-flight effects far from the shower axis (see Sec. 6.3.2), a 2 by 2 mm selection criterion around the shower axis is used to establish the timewalk. Several beam energies between 100 and 300 GeV electrons are then averaged. The result is shown in Fig. 6.25. The data points are fitted with Eq. 4.1. Compared to Sec. 4.1, it can be seen that even at the largest measured signals of around 700 MIPs, no plateau is reached and the timewalk is still significant.



(a) (b)

Figure 6.24: Principle of the ToA calibration: The frequency distribution of the ToA values (a) is integrated and projected into a range of [0,1]. The integrated frequency is then projected into a range [0,25] and fitted with Eq. 6.15 (b).

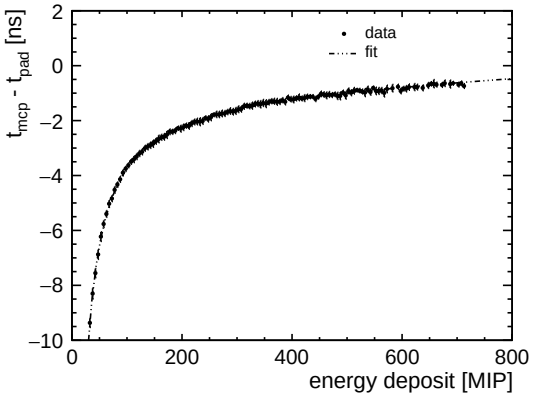


Figure 6.25: The timewalk measured with a set of HGCAL prototype modules equipped with Skiroc2CMS readout.

The achieved time resolution for 100 GeV electrons as a function of the energy deposit is shown in Fig. 6.26. The ToA threshold is approximately 20 MIPs. Each data point is an average of five channels on different layers along the shower axis. The data points are fitted with Eq. 6.10, similar to Sec. 6.4. The jitter term A is limited to approximately 13 ns/MIP, corresponding to 46 ns/fC for the 300 µm silicon sensor, while the constant term C reaches around 70 ps, most likely limited by the jitter of the 40 MHz clock. The MCP resolution is expected to be below 20 ps and is not unfolded. The results are worse than the 10 ns/fC and 50 ps expected from laboratory measurements described in [69]. It is thought that this difference arises from the fact that in a test pulse environment timewalk correction is not required.

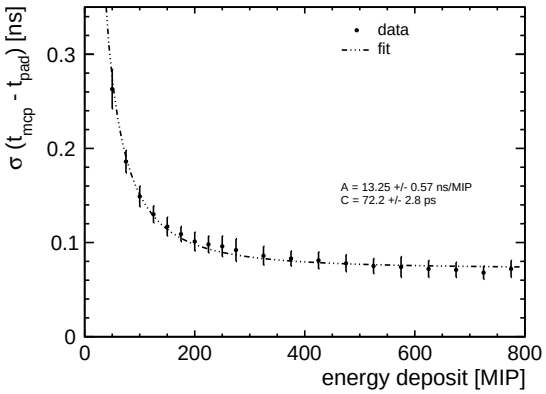


Figure 6.26: The achieved resolution measured with a set of HGCAL prototype modules equipped with Skiroc2CMS readout.

The calibration of this ASIC is an ongoing effort. The results shown in Fig. 6.26 are the averaged response of a few channels that had enough statistics to attempt the calibration procedure. With an improved energy and timewalk calibration, as well as the inclusion of ToA Fall information, it is likely that the achievable resolution can be improved beyond what is presented in this section. However, due to the limited energy resolution especially at low energy deposits, the imperfect timewalk correction will remain a significant contribution to the time resolution.

For studies of the time structure of electromagnetic and hadronic shower evolution, significantly more pads have to be calibrated. In beam tests, the low data rate of the readout and the limited time have so far prohibited the calibration of a larger amount of pads. A more flexible solution for the calibration would be a pulse generator triggered by a random external event, e.g. a radioactive source, or a pulse generator with a delay binning that is finer than the TDC binning of about 20 ps. This way, only the clock delay per channel has to be established in beams, while the ADC range, clock phase and timewalk can be established with test pulses.

6.7 Summary and Discussion

Single pad timing uncertainties of below 15 ps for S/N ratios above 200 have been reached with an HGCAL prototype sensor coupled to a custom DRS4 based frontend. The uncertainty is

well modelled by a two parameter function. Neither the silicon sensor itself nor lateral shower fluctuations have shown a significant contribution to the constant term of the time resolution.

For the cluster timing, no significant improvement from integrating timestamps from additional pads inside the cluster compared to taking the most precise timestamp in the cluster has been found. However, it allows for a more uniform performance with no dependency on the impact position on the sensor. These conclusions however, apply only to the specific geometry and electronics used. For example, if a smaller pad size or a larger Molière radius is used, individual S/N values would decrease while the cluster sizes would increase. For such a system, a high precision TDC has a reduced impact.

From the first tests with the Skiroc2CMS ASIC it can be seen that at this point the most difficult challenges to reach 30 ps on a system level are neither the TDC nor the jitter of the preamplifier but the jitter of the clock as well as the timewalk calibration. The final HGCROC ASIC will likely outperform Skiroc2CMS significantly. However, provided the synchronisation of channels across multiple layers of the calorimeter can be done with a precision below 30 ps, then the combination of multiple layers with a system close to the presented performance in Sec. 6.6 should be sufficient to reach a resolution of at least 50 ps for high-energy photons and electrons.

It has to be clearly stated, that the full impact of such a time resolution in combination with the granularity is not fully understood. As described in Sec. 3.3 and at the beginning of this chapter, the time resolution can be used to improve the vertex identification by 'slicing' the bunch crossing event into shorter periods and thus reducing the effective vertex density. More complex approaches that include the time information into clustering and shower reconstruction have only started and few results have been publicly presented. In [41, Sec. 10.4] results on spatial clustering in combination with a selection window on the hit time in the presence of pile-up are shown. While the results on the time distribution and cluster identification are promising, no further results on higher level reconstruction parameters are shown. It has been suggested [72], that a time resolution of 1 ns allows the identification of elastic neutron scattering events and can be used to improve the hadronic energy resolution. Whether inelastic neutron reactions could also be identified if the time resolution approaches 30 ps, remains to be seen. Similarly, it has yet to be established whether CLIC could profit from a time resolution on this level. The current calorimeter design envisages a resolution of 1 ns per pad. Further knowledge on the time evolution of hadronic showers and optimised reconstruction algorithms are certainly required. However, a fully functional HGCAL will present an excellent tool to conduct such studies.

Chapter 7

Summary

While LHC has completed the standard model and confirmed many of its predictions, HL-LHC and CLIC will explore phenomena beyond the standard model. Complementary physics programmes will allow for further exploration of the mass scale above 1 TeV as well as precision measurements to confirm or discard competing theoretical models. In both accelerators and their corresponding experiments, silicon detectors will play a vital role to reach the necessary precision and cope with the difficult experimental environments.

For a future vertex detector at CLIC, it has been shown that planar silicon pixel sensors of 50 µm substrate thickness can comfortable meet the requirements on the time resolution but lack any significant amount of charge sharing to improve beyond the binary pixel resolution. Therefore, new bonding techniques for pixel pitches below 20 µm together with a smaller transistor feature size are likely needed to make this technology the primary choice for CLIC.

Silicon sensors will also play a vital role as active elements in future highly granular calorimeters as are foreseen for CLIC and HGCAL. Here, HGCAL will be the first calorimeter of this kind and will give vital experience for CLIC. Many design sensor features have been evaluated and can feed into the design for CLIC. It has also been shown that the contribution of silicon pad sensors to the time resolution is limited to the capacitance. Intrinsic contributions are below 10 ps.

For CLIC as well as HGCAL, the granularity will surpass any previous detector system. While improving performance, it creates challenges for characterisation and calibration. The ARRAY system has been developed to characterise and test silicon pad sensors. For the calibration, it was shown that individual calibration of each pixel or pad is essential to reach the full performance potential. To do so, the use of a combined method utilising electrical test pulses and particle beams has proven very useful for the calibration of the Timepix3 ASIC. For Skiroc2CMS ASIC, such a system is in development and will help to improve beyond the results shown in this work.

List of References

[1] The ATLAS Collaboration. 'Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC'. In: *Phys. Lett. B* 716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020.

[2] The CMS Collaboration. 'Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC'. In: *Phys. Lett. B* 716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021.

[3] D. Hanneke, S. Fogwell and G. Gabrielse. 'New Measurement of the Electron Magnetic Moment and the Fine Structure Constant'. In: *Phys. Rev. Lett.* 100 (2008), p. 120801. DOI: 10.1103/PhysRevLett.100.120801.

[4] The LHCb Collaboration. 'Observation of the Resonant Character of the $Z(4430)^-$ State'. In: *Phys. Rev. Lett.* 112 (2014), p. 222002. DOI: 10.1103/PhysRevLett.112.222002.

[5] The LHCb Collaboration. 'Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays'. In: *Phys. Rev. Lett.* 115 (2015), p. 072001. DOI: 10.1103/PhysRevLett.115.072001.

[6] The CMS Collaboration. 'Measurement of the $B_s^0 \to \mu^+\mu^-$ Branching Fraction and Search for $B^0 \to \mu^+\mu^-$ with the CMS Experiment'. In: *Phys. Rev. Lett.* 111 (2013), p. 101804. DOI: 10.1103/PhysRevLett.111.101804.

[7] The LHCb Collaboration. 'First Evidence for the Decay $B_s^0 \to \mu^+\mu^-$'. In: *Phys. Rev. Lett.* 110 (2013), p. 021801. DOI: 10.1103/PhysRevLett.110.021801.

[8] G. Apollinari et al. *High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report*. CERN-2017-007-M. 2017. DOI: https://dx.doi.org/10.23731/CYRM-2017-004.

[9] P. Burrows et al. *The Compact Linear e+e- Collider (CLIC) - 2018 Summary Report*. CERN-2018-005-M. 2018. DOI: 10.23731/CYRM-2018-002.

[10] Wikipedia article on box-plots. Last accessed 21/11/2019. URL: https://en.wikipedia.org/wiki/Box_plot.

[11] H. Spieler. *Semiconductor Detector Systems, 2nd ed.* Oxford University Press, 2005. ISBN: 9780198527848.

[12] W. R. Leo. *Techniques for Nuclear and Particle Physics Experiments, 2nd ed.* Springer Verlag Berlin Heidelberg, 2005. ISBN: 9783540572800.

[13] G. F. Knoll. *Radiation Detection and Measurement, 4th ed.* John Wiley and Sons, 2010. ISBN: 9780470131480.

[14] H. Kolanoski and N. Wermes. *Teilchendetektoren, 1st ed.* Springer-Verlag Berlin Heidelberg, 2016. ISBN: 9783662453490.

[15] G.W. Fraser et al. 'The X-ray energy response of silicon Part A. Theory'. In: *Nucl. Instr. Meth. A* 350 (1994), pp. 368–378. DOI: 10.1016/0168-9002(94)91185-1.

[16] A. Chilingarov. 'Temperature dependence of the current generated in Si bulk'. In: *J. Instr.* 8 (2013), P10003. DOI: 10.1088/1748-0221/8/10/p10003.

[17] Wikipedia article on Float-Zone-Silicon. Last accessed 21/11/2019. URL: https://en.wikipedia.org/wiki/Float-zone_silicon.

[18] S. M. Sze and K. K. Ng. *Physics of Semiconductor Devices, 3rd ed.* John Wiley and Sons, 2007. ISBN: 9780471143239.

[19] L. Rossi et al. *Pixel Detectors: From Fundamentals to Applications, 1st ed.* Springer-Verlag Berlin Heidelberg, 2006. ISBN: 9783540283324.

[20] Presentation by Sami Vahanen, VERTEX Conference 2017. Last accessed 21/11/2019. URL: https://indico.cern.ch/event/627245/contributions/2676713.

[21] R. Wigmans. *Calorimetry, 2nd ed.* Oxford University Press, 2017. ISBN: 9780198786351.

[22] H. Bichsel. 'Straggling in thin silicon detectors'. In: *Rev. Mod. Phys.* 60 (1988), pp. 663–699. DOI: 10.1103/RevModPhys.60.663.

[23] H. Sadrozinski. '4-Dimensional Tracking with Ultra-Fast Silicon Detectors'. In: *Rev. Mod. Phys.* 81 (2018), p. 026101. DOI: 10.1103/RevModPhys.60.663.

[24] Presentation by Stefan Ritt, PSEC Meeting 2014. Last accessed 21/11/2019. URL: https://psec.uchicago.edu/library/chipdesign/ritt_timing_calibration_method.pdf.

[25] B. Hyams et al. 'A silicon counter telescope to study short-lived particles in high-energy hadronic interactions'. In: *Nucl. Instr. Meth. A* 205 (1983), pp. 99–105. DOI: 10.1016/0167-5087(83)90177-1.

[26] F. Pitters. 'The CLIC detector concept'. In: *CERN Proceedings* 1 (2017), p. 37. DOI: 10.23727/CERN-Proceedings-2017-001.37.

[27] F. Pitters. 'The CMS High-Granularity Calorimeter for Operation at the High-Luminosity LHC'. In: *Springer Proceedings in Physics* 213 (2018). DOI: 10.1007/978-981-13-1316-5_2.

[28] The CMS Collaboration. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software.* CERN-LHCC-2006-001. 2006. URL: http://cds.cern.ch/record/922757.

[29] The CMS webpage. Last accessed 21/11/2019. URL: http://cms.cern/detector.

[30] The CLICdp Collaboration. *A detector for CLIC: main parameters and performance.* CLICdp-Note-2018-005. 2018. URL: https://cds.cern.ch/record/2649437.

[31] H. Abramowicz et al. 'Higgs physics at the CLIC electron-positron linear collider'. In: *Eur. Phys. J. C.* 77 (2017), p. 475. DOI: 10.1140/epjc/s10052-017-4968-5.

[32] The CLICdp Collaboration. *Top-Quark Physics at the CLIC Electron-Positron Linear Collider.* 2018. arXiv: 1807.02441.

[33] J. de Blas et al. *The CLIC Potential for New Physics.* 2018. DOI: 10.23731/CYRM-2018-003. arXiv: 1812.02093.

[34] T. Abe et al. *The International Large Detector: Letter of Intent.* DESY-2009-87. 2010. arXiv: 1006.3396.

[35] H. Aihara et al. *SiD Letter of Intent.* SLAC-R-989. 2009. arXiv: 0911.0006.

[36] M. A. Thomson. 'Particle Flow Calorimetry and the PandoraPFA Algorithm'. In: *Nucl. Instr. Meth. A* 611 (2009), p. 25. DOI: 10.1016/j.nima.2009.09.009.

[37] J. S. Marshall, A. Münnich and M. A. Thomson. 'Performance of Preport Flow Calorimetry at CLIC'. In: *Nucl. Instr. Meth. A* 700 (2012), p. 153. DOI: 10.1016/j.nima.2012.10.038.

[38] The CLICdp Collaboration. *CLICdet: The post-CDR CLIC detector model*. CLICdp-Note-2017-001. 2017. URL: https://cds.cern.ch/record/2254048.

[39] C. Adloff et al. 'Hadronic energy resolution of a highly granular scintillator-steel hadron calorimeter using software compensation techniques'. In: *J. Instr.* 7 (2012), P09017. DOI: 10.1088/1748-0221/7/09/p09017.

[40] D. Contardo et al. *Technical Proposal for the Phase-II Upgrade of the CMS Detector*. CERN-LHCC-2015-010. 2015. URL: https://cds.cern.ch/record/2020886.

[41] The CMS Collaboration. *The Phase-2 Upgrade of the CMS Endcap Calorimeter*. CERN-LHCC-2017-023. 2017. URL: https://cds.cern.ch/record/2293646.

[42] L. Gray and T. Tabarelli de Fatis. *Technical Proposal for a MIP Timing Detector in the CMS Experiment Phase-2 Upgrade*. LHCC-P-009. 2017. URL: https://cds.cern.ch/record/2296612.

[43] E. Curras et al. 'Radiation hardness and precision timing study of silicon detectors for the CMS High Granularity Calorimeter'. In: *Nucl. Instr. Meth. A* 845 (2017), pp. 60–63. DOI: 10.1016/j.nima.2016.05.008.

[44] *List of OMEGA ROCs*. Last accessed 21/11/2019. URL: https://portail.polytechnique.edu/omega/en/products/products-presentation.

[45] T. Poikela et al. 'Timepix3: a 65K channel hybrid pixel readout chip with simultaneous ToA/ToT and sparse readout'. In: *J. Instr.* 9 (2014), p. C05013. DOI: 10.1088/1748-0221/9/05/C05013.

[46] M. Campbell et al. 'A readout chip for a 64x64 pixel matrix with 15-bit single photon counting'. In: *IEEE Trans. Nucl. Sci.* 45 (1998), pp. 751–753. DOI: 10.1109/23.682629.

[47] X. Llopart et al. 'Timepix, a 65k programmable pixel readout chip for arrival time, energy and/or photon counting measurements'. In: *Nucl. Instr. Meth. A* 581 (2007), pp. 485–494. DOI: 10.1016/j.nima.2007.08.079.

[48] F. Pitters et al. 'Time and Energy Calibration of Timepix3 Assemblies with Thin Silicon Sensors'. In: CLICdp-Note-2018-006 (2018). URL: https://cds.cern.ch/record/2649493.

[49] F. Pitters et al. 'Time resolution studies of Timepix3 assemblies with thin silicon pixel sensors'. In: *J. Instr.* 14 (2019), P05022. DOI: 10.1088/1748-0221/14/05/p05022.

[50] D. Dannheim et al. 'Detector Technologies for CLIC'. In: CERN Yellow Reports: Monographs Vol. 1 (2019). DOI: 10.23731/CYRM-2019-001.

[51] F. Krummenacher. 'Pixel detectors with local intelligence: an IC designer point of view'. In: *Nucl. Instr. Meth. A* 305 (1991), pp. 527–532. DOI: 10.1016/0168-9002(91)90152-G.

[52] M. De Gaspari et al. 'Design of the analog front-end for the Timepix3 and Smallpix hybrid pixel detectors in 130 nm CMOS technology'. In: *J. Instr.* 9 (2014), p. C01037. DOI: 10.1088/1748-0221/9/01/C01037.

[53] J. Visser et al. 'SPIDR: a read-out system for Medipix3 and Timepix3'. In: *J. Instr.* 10 (2015), p. C12028. DOI: 10.1088/1748-0221/10/12/C12028.

[54] N. Alipour Tehrani. 'Test-beam measurements and simulation studies of thin pixel sensors for the CLIC vertex detector'. PhD thesis. ETH Zürich, 2017. DOI: 10.3929/ethz-b-000164813.

[55] Lawrence Berkeley National Laboratory. *The X-Ray data booklet*. Last accessed 21/11/2019. URL: https://xdb.lbl.gov.

[56] R. Turchetta. 'Spatial resolution of silicon microstrip detectors'. In: *Nucl. Instr. Meth. A* 335 (1993), pp. 44–58. DOI: 10.1016/0168-9002(93)90255-G.

[57] Xavi Llopart. Private communication.

[58] H. Jansen. *Lateraler-Drift-Halbleitersensor*. DE Patent App. DE201,510,116,270. 2017.

[59] F. Pitters et al. 'ARRAY: An Open Source, Modular and Integrated Probe-Card based System for Characterisation of Large Area Silicon Pad Sensors'. In: *Nucl. Instr. Meth. A* 940 (2019), pp. 168–173. DOI: 10.1016/j.nima.2019.06.007.

[60] *CERN Open Hardware Repository*. Last accessed 21/11/2019. URL: https://www.ohwr.org/project/array.

[61] *MAX328 datasheet*. URL: https://datasheets.maximintegrated.com/en/ds/MAX328-MAX329.pdf.

[62] Keysight Technologies. *Keysight Impedance Measurement Handbook, 6th ed.* 2016. URL: https://literature.cdn.keysight.com/litweb/pdf/5950-3000.pdf.

[63] E. Fretwurst et al. 'The influence of edge effects on the determination of the doping profile of silicon pad diodes'. In: *Nucl. Instr. Meth. A* 867 (2017), pp. 231–236. DOI: 10.1016/j.nima.2017.04.020.

[64] N. Akchurin et al. 'First beam tests of prototype silicon modules for the CMS High Granularity Endcap Calorimeter'. In: *J. Instr.* 13 (2018), P10023. DOI: 10.1088/1748-0221/13/10/P10023.

[65] J. Spanggaard. *Delay Wire Chambers Users Guide*. 1998. URL: http://cds.cern.ch/record/702443.

[66] A. Ronzhin et al. 'Study of the timing performance of micro-channel plate photomultiplier for use as an active layer in a shower maximum detector'. In: *Nucl. Instr. Meth. A* 795 (2015), pp. 288–292. DOI: 10.1016/j.nima.2015.06.006.

[67] *CAEN V1742 Datasheet*. URL: http://www.caen.it/csite/CaenProd.jsp?parent=11&idmod=661.

[68] *DRS4 Datasheet*. URL: https://www.psi.ch/drs/DocumentationEN/DRS4_rev09.pdf.

[69] J. Borg et al. 'SKIROC2_CMS an ASIC for testing CMS HGCAL'. In: *J. Instr.* 12 (2017), p. C02019. DOI: 10.1088/1748-0221/12/02/c02019.

[70] Presentation by Stephane Callier, HGCAL WG Meeting 2017 (restricted access). Last accessed 21/11/2019. URL: https://indico.cern.ch/event/609269/contributions/2456363.

[71] Artur Lobanov. Private communication.

[72] Presentation by Christian Graf, CALICE Collaboration Meeting Shanghai 2018. Last accessed 21/11/2019. URL: https://agenda.linearcollider.org/event/7799/contributions/42206.

Appendices

Appendix A

Box Plot Conventions Used in this Work

Many results in the presented work are shown as box plots. For these plots, the median of the distribution is represented by a solid line, the mean by a dashed line. The box boundary represents the inter-quartile range, i.e. the middle 50% of the distribution (which corresponds to $1.349\,\sigma$). The whiskers are at 2.5 times the inter-quartile range and contain $5.396\,\sigma$ or 98.95% of all data points. For more details on this form of representation the reader is referred to [10].

Appendix B

Derivation of Timewalk and Time-over-Threshold Response Function

Consider an event with amplitude x occurring at time t_0. The time t_0' when the amplitude crosses a constant threshold value x_{thr} is recorded. Due to the finite and constant rise time t_{rise} of the system, this timestamp is shifted in time by an amount that depends on the signal amplitude

$$t_0' = t_0 + \Delta t$$

This is sketched in fig. B.1.

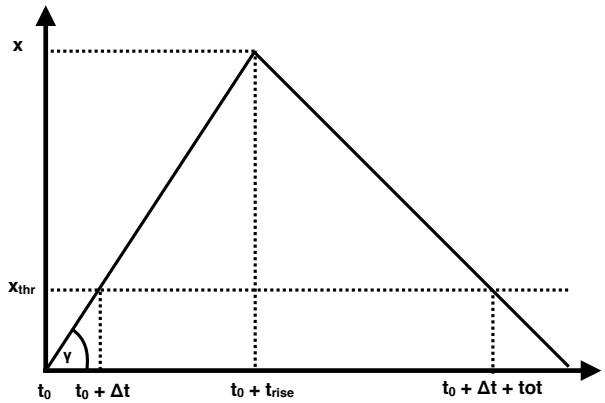


Figure B.1: Sketch of the timewalk and time over threshold principle.

Assuming a linear rise and fall, the following relations holds.

$$\tan\gamma = \frac{x_{\text{thr}}}{\Delta t} = \frac{x}{t_{\text{rise}}}$$

Both x_{thr} and t_{rise} are system constants. With the possibility of a constant delay and a shifted amplitude scale x′, the time shift can then be expressed as

$$\Delta t = \frac{c_0}{x' - c_1} + c_2 \tag{B.1}$$

The time-over-threshold before the peaking time is then $t_{\mathrm{rise}} - \Delta t$. Assuming also a constant discharge rate, the time to discharge the capacitor is linear and total time over threshold

$$\mathrm{tot} = \frac{c_0}{x' - c_1} + c_3 x' + c_4 \tag{B.2}$$

In reality, the signal shape is not a perfect triangle. The rising edge is not fully linear and the discharge already starts as soon as there is a potential difference across the CSA. However, for Timepix3 the rise time of 25 ns is much faster than the discharge and the functions derived here fit quite well.

Appendix C

Netlist Generation

The following python function has been used to generate the netlist that is used to simulate the full system of switch card, probe card and sensor in Chapter 5.

```
def genNetlist(self, freq, calibration=""):

  ## Simulated Conditions
  self.freq = freq
  self.csensor = 40e-12
  self.protectionResistor = 1e3
  self.biasResistor = 1e7
  self.number_of_channels = 270

  ## HV filter values
  self.filterCapacitor = 60e-9
  self.filterResistor = 40e3
  self.splitResistor = 0

  ## Multiplexers & Switches
  self.switches_to_ground = 1
  self.rmux = 1.5e3
  self.csw1 = 1.2e-12
  self.csw2 = 1.2e-12

  ## Cross Coupling
  self.include_coupling_capacitance = 1
  self.cwire1 = 15.8e-12  # parasitic C from after MUX 1 to GND
  self.cwire2 = 15.8e-12  # parasitic C from after MUX 2 to GND
  self.cwire3 = 15.8e-12  # parasitic C from after MUX 3 to GND
  self.cwire4 = 15.8e-12  # parasitic C from after MUX 4 to GND
  self.ccsensor = 10e-12  # parasitic C between traces
  self.ccgnd = 5e-12      # parasitic C between each protection resistor and GND

  ## LCR Decoupling
  self.include_LCR_coupling = 1

```

```
## Basic netlist
netlist = ""

netlist += """
*.model switchMAX328 sw vt=1 vh=0.1 ron=1 roff=1e12

*.subckt swMAX328 gnd A B en
*CA A gnd 20p
*CB A gnd 0p
*Switch1 A B en gnd switchMAX328
*.ends

.subckt swShortON gnd A B
CA A gnd %.5e
CB B gnd %.5e
R0 A B 10
.ends

.subckt swShortOFF gnd A B
CA A gnd %.5e
CB B gnd %.5e
R0 A B 10e12
.ends

.subckt swON gnd A B
CA A gnd %.5e
CB B gnd %.5e
R0 A B %.5e
.ends

.subckt swOFF gnd A B
CA A gnd %.5e
CB B gnd %.5e
R0 A B 1e12
.ends
""" % (self.csw1, self.csw1, self.csw1, self.csw1, self.csw2, \
       self.csw2, self.rmux, self.csw2, self.csw2)

include_sensor_cap = 0

if calibration == "OPEN":
    include_sensor_cap = 0

elif calibration == "SHORT":
    include_sensor_cap = 0
        netlist += "* short to ground\n"
        for i in range(self.number_of_channels):
            netlist += "Rshort%03d SenBackPlane Vin%03d 1\n" % (i, i)
elif calibration == "OPENONE":
  include_sensor_cap = 0
  for i in range(self.number_of_channels):
    if i == 0:
      pass
```

```
86        else:
87          netlist += "Csensor%03d␣SenBackPlane␣Vin%03d␣%.3e\n" % (i, i, self.csensor)
88    elif calibration == "SHORTONE":
89      include_sensor_cap = 0
90      netlist += "*␣short␣to␣ground\n"
91      for i in range(self.number_of_channels):
92        if i == 0:
93          netlist += "Csensor%03d␣SenBackPlane␣Vin%03d␣1\n" % (i, i)
94        else:
95          netlist += "Csensor%03d␣SenBackPlane␣Vin%03d␣%.3e\n" % (i, i, self.csensor)
96    else:
97      include_sensor_cap = 1
98
99    ## Adding the filter network
100   include_hv_filter = 1
101   if include_hv_filter:
102     netlist += "VHV␣HVP␣gnd␣10\n"
103     netlist += "R4␣HVP␣filterCapacitor␣%.3e\n" % self.filterResistor
104     netlist += "R5␣filterCapacitor␣SenBackPlane␣%.3e\n" % self.splitResistor
105     netlist += "C4␣filterCapacitor␣gnd␣%.e\n" % self.filterCapacitor
106   else:
107     netlist += "R3␣SenBackPlane␣gnd␣0\n"
108
109   ## Adding all diodes
110   if include_sensor_cap:
111     netlist += "*␣sensor␣capacitances\n"
112     for i in range(self.number_of_channels):
113       netlist += "Csensor%03d␣SenBackPlane␣Vin%03d␣%.3e\n" % (i, i, self.csensor)
114
115   ## Adding cross coupling between diodes
116   if self.include_coupling_capacitance:
117     netlist += "*␣sensor␣cross␣coupling␣capacitances\n"
118     for i in range(self.number_of_channels):
119       netlist += "Ccoupling%03d␣Vin%03d␣Vin%03d␣%.3e\n" % (i, i, (i+1)%self.
      number_of_channels, self.ccsensor)
120
121   chn = 0
122
123   ## Adding the protection resistor
124   netlist += "*␣protection␣resistors\n"
125   for i in range(self.number_of_channels):
126     netlist += "Rprotect%03d␣Vin%03d␣Vinprot%03d␣%.2e\n" % (i, i, i, self.
      protectionResistor)
127
128   ## Adding parasitic capacitance on input line
129   netlist += "*␣parasitic␣capacitance␣on␣input␣line\n"
130   for i in range(self.number_of_channels):
131     netlist += "Cconnector%03d␣gnd␣Vin%03d␣%.3e\n" % (i, i, self.ccgnd)
132
133   ## Close all switches except the channel under test
134   if self.switches_to_ground:
135     netlist += "\n*␣switches␣to␣ground\n"
136     for i in range(self.number_of_channels):
```

```
137       en = "ON"
138       if i == chn:
139         en = "OFF"
140       netlist += "Xgndsw_0%03d␣gnd␣Vinprot%03d␣gnd␣swShort%s\n" % (i, i, en)
141
142   ## Adding capacitance between layers of MUXes
143   netlist += "\n*␣L0\n"
144   for i in range(self.number_of_channels):
145     en = "OFF"
146     if i == chn:
147       en = "ON"
148     netlist += "XL0_0%03d␣gnd␣Vinprot%03d␣VL1_%03d␣sw%s\n" % (i, i, i/8, en)
149
150   netlist += "\n*␣L1\n"
151   for i in range(self.number_of_channels/8):
152     en = "OFF"
153     if i == int(chn/8):
154       en = "ON"
155     netlist += "XL1_0%03d␣gnd␣VL1_%03d␣VL2_%03d␣sw%s\n" % (i, i, i/8, en)
156     netlist += "CL1_0%03d␣gnd␣VL1_%03d␣%.5e\n" % (i, i, self.cwire1)
157
158   netlist += "\n*␣L2\n"
159   for i in range(self.number_of_channels/64):
160     en = "OFF"
161     if i == int(chn/64):
162       en = "ON"
163     netlist += "XL2_0%03d␣gnd␣VL2_%03d␣VL3␣sw%s\n" % (i, i, en)
164     netlist += "CL2_0%03d␣gnd␣VL2_%03d␣%.5e\n" % (i, i, self.cwire2)
165
166   netlist += "CL3␣gnd␣VL3␣%.5e\n" % (self.cwire3)
167   netlist += "XoutMux␣gnd␣VL3␣VLCR␣swON\n"
168   netlist += "CVLCR␣gnd␣VLCR␣%.5e\n" % (self.cwire4)
169   netlist += "Rbias␣gnd␣VLCR␣%.5e\n" % (self.biasResistor)
170   netlist += "Rin␣gnd␣0␣0\n"
171   netlist += "Vin␣VLCR_P␣VLCR_N␣␣DC␣0␣AC␣1\n"
172
173   ## Adding LCR decoupling capacitors
174   if self.include_LCR_coupling:
175     netlist += "CLCRP␣VLCR␣VLCR_P␣1e-6\n"
176     netlist += "CLCRN␣SenBackPlane␣␣VLCR_N␣1e-6\n"
177     netlist += "RLCRP␣VLCR␣VLCR_P␣10G\n"
178     netlist += "RLCRN␣SenBackPlane␣␣VLCR_N␣10G\n"
179   else:
180     netlist += "RLCRP␣VLCR␣VLCR_P␣0\n"
181     netlist += "RLCRN␣gnd␣VLCR_N␣0\n"
182
183   netlist += ".ac␣dec␣1␣%.3e␣%.3e" % (self.freq,self.freq)
184   netlist += """
185     .control
186     set filetype=ascii
187     run
188     write output.txt i(vin)
189
```

```
190       exit
191       .endc
192       .end """
193
194   return netlist
```